

frontiers RESEARCH TOPICS

NEAR-INFRARED SPECTROSCOPY: RECENT ADVANCES IN INFANT SPEECH PERCEPTION AND LANGUAGE ACQUISITION RESEARCH

Topic Editor
Judit Gervain



frontiers in
PSYCHOLOGY



frontiers

FRONTIERS COPYRIGHT STATEMENT

© Copyright 2007-2015
Frontiers Media SA.
All rights reserved.

All content included on this site, such as text, graphics, logos, button icons, images, video/audio clips, downloads, data compilations and software, is the property of or is licensed to Frontiers Media SA ("Frontiers") or its licensees and/or subcontractors. The copyright in the text of individual articles is the property of their respective authors, subject to a license granted to Frontiers.

The compilation of articles constituting this e-book, wherever published, as well as the compilation of all other content on this site, is the exclusive property of Frontiers. For the conditions for downloading and copying of e-books from Frontiers' website, please see the Terms for Website Use. If purchasing Frontiers e-books from other websites or sources, the conditions of the website concerned apply.

Images and graphics not forming part of user-contributed materials may not be downloaded or copied without permission.

Individual articles may be downloaded and reproduced in accordance with the principles of the CC-BY licence subject to any copyright or other notices. They may not be re-sold as an e-book.

As author or other contributor you grant a CC-BY licence to others to reproduce your articles, including any graphics and third-party materials supplied by you, in accordance with the Conditions for Website Use and subject to any copyright notices which you include in connection with your articles and materials.

All copyright, and all rights therein, are protected by national and international copyright laws.

The above represents a summary only. For the full conditions see the Conditions for Authors and the Conditions for Website Use.

Cover image provided by lbbl sarl, Lausanne CH

ISSN 1664-8714

ISBN 978-2-88919-415-5

DOI 10.3389/978-2-88919-415-5

ABOUT FRONTIERS

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

FRONTIERS JOURNAL SERIES

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing.

All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

DEDICATION TO QUALITY

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.

Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view.

By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

WHAT ARE FRONTIERS RESEARCH TOPICS?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area!

Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: researchtopics@frontiersin.org

NEAR-INFRARED SPECTROSCOPY: RECENT ADVANCES IN INFANT SPEECH PERCEPTION AND LANGUAGE ACQUISITION RESEARCH

Topic Editor:

Judit Gervain, Psychology, CNRS - Universite Paris Descartes, Paris, France

Near-Infrared Spectroscopy (NIRS) is a novel and increasingly popular optical imaging technique that has revolutionarized brain research in the youngest developmental populations. After nearly a decade of technological development, NIRS has become a reliable, easy-to-use and efficient tool to explore the linguistic and cognitive abilities of neonates and young infants, opening new vistas for the investigation of language acquisition and cognitive development. This Research Topic covers the latest advances in these areas brought about by NIRS imaging. The main focus is to highlight innovative and foundational studies that go beyond methodological issues and advance our theoretical understanding of infant and child development. Contributions from the pioneers of this method are selected, illustrating how NIRS has allowed developmental researchers to ask theoretically relevant questions that more traditional methods couldn't address. These works further our understanding of language and cognitive development and bring us closer to bridging the gap between brain, mind and behavior at the very beginning of life.

Table of Contents

05 *Near-Infrared Spectroscopy: Recent Advances in Infant Speech Perception and Language Acquisition Research*

Judit Gervain

I: The Use of NIRS in Developmental Cognitive Neuroscience: Theoretical Issues and Methodological Principles

07 *Linking Behavioral and Neurophysiological Indicators of Perceptual Tuning to Language*

Eswen Fava, Rachel Hull and Heather Bortfeld

21 *Studying Neonates' Language and Memory Capacities with Functional Near-Infrared Spectroscopy*

Silvia Benavides-Varela, David M. Gómez and Jacques Mehler

II: Hemispheric Specialization for Language

26 *Language and the Newborn Brain: Does Prenatal Language Experience Shape the Neonate Neural Response to Speech?*

Lillian May, Krista Byers-Heinlein, Judit Gervain and Janet F. Werker

35 *Acoustic Processing of Temporally Modulated Sounds in Infants: Evidence From a Combined Near-Infrared Spectroscopy and EEG Study*

Silke Telkemeyer, Sonja Rossi, Till Nierhaus, Jens Steinbrink, Hellmuth Obrig and Isabell Wartenburger

49 *Assessing Signal-Driven Mechanism in Neonates: Brain Responses to Temporally and Spectrally Different Sounds*

Yasuyo Minagawa-Kawai, Alejandrina Cristià, Inga Vendelin, Dominique Cabrol and Emmanuel Dupoux

61 *Functional Hemispheric Specialization in Processing Phonemic and Prosodic Auditory Changes in Neonates*

Takeshi Arimitsu, Mariko Uchida-Ota, Tatsuhiko Yagihashi, Shozo Kojima, Shigeru Watanabe, Isamu Hokuto, Kazushige Ikeda, Takao Takahashi and Yasuyo Minagawa-Kawai

71 *Functional Lateralization of Speech Processing in Adults and Children who Stutter*

Yutaka Sato, Koichi Mori, Toshizo Koizumi, Yasuyo Minagawa-Kawai, Akihiro Tanaka, Emi Ozawa, Yoko Wakaba and Reiko Mazuka

III: Language Processing and Comprehension Mechanisms

81 *Large-Scale Brain Networks Underlying Language Acquisition in Early Infancy*

Fumitaka Homae, Hama Watanabe, Tamami Nakano and Gentaro Taga

95 *Neural Processing of Repetition and Non-Repetition Grammars in 7- and 9-Month-Old Infants*

Jennifer B. Wagner, Sharon E. Fox, Helen Tager-Flusberg and Charles A. Nelson

103 *The Role of Orbitofrontal Cortex in Processing Empathy Stories in 4-8 Year-Old Children*

Tila Tabea Brink, Karolina Urton, Dada Held, Evgeniya Kirilina, Markus J. Hofmann, Gisela Klann-Delius, Arthur M. Jacobs and Lars Kuchinke



Near-infrared spectroscopy: recent advances in infant speech perception and language acquisition research

Judit Gervain *

Laboratoire Psychologie de la Perception, CNRS - Université Paris Descartes, Paris, France

*Correspondence: judit.gervain@parisdescartes.fr

Edited and reviewed by:

Manuel Carreiras, Basque Center on Cognition, Brain and Language, Spain

Keywords: language learning, language developmental, speech perception, brain specialization for language, near-infrared spectroscopy, developmental cognitive neuroscience

Near-Infrared Spectroscopy (NIRS) is a relatively novel and increasingly popular optical imaging technique that has revolutionized brain research in the developmental populations (Villringer and Chance, 1997; Lloyd-Fox et al., 2009; Gervain et al., 2011). After more than a decade of technological development, NIRS has become a reliable, easy-to-use and efficient tool to explore the linguistic and cognitive abilities of neonates and young infants, opening new vistas for the investigation of language acquisition and cognitive development. This Research Topic covers the latest advances in these areas brought about by NIRS imaging. The main focus is to highlight innovative and foundational studies that go beyond methodological issues and advance our theoretical understanding of infant and child development. Contributions from the pioneers of this method are selected, illustrating how NIRS has allowed developmental researchers to ask theoretically relevant questions that more traditional methods couldn't address.

The first two contributions, by Fava et al. (2011) and Benavides-Varela et al. (2011), cover general theoretical issues and methodological principles. They provide a critical, but constructive overview of theoretical questions about linguistic and cognitive development that have been asked, outline challenges that the NIRS community still needs to face and offer recommendations for optimal experimental designs and data interpretations practices.

These general contributions are followed by a series of empirical papers exploring a key issue in the study of the neural correlates of language learning and development, the nature and origins of the brain specialization for speech and language. While it is well established that in the majority of right-handed adults, language is preferentially processed in the left hemisphere (e.g., Friederici, 2005), the reasons for and the ontogenetic origins of this left lateralization have so far been less well understood, partly because the field lacked a safe, fully non-invasive, participant-friendly brain imaging method with which to probe the infant brain. NIRS has filled this gap, opening up the way for exciting new discoveries about the brain specialization for speech and language in young babies (e.g., Pena et al., 2003; Sato et al., 2012). Five experimental articles in the current volume contribute to this exciting inquiry. May et al. (2011) compare newborn infants' brain responses to the native language, spoken by the mother during pregnancy, and to an unknown language, in an attempt

to investigate how prenatal experience with speech might shape the brain specialization for language. Telkemeyer et al. (2011), Arimitsu et al. (2011) as well as Minagawa-Kawai et al. (2011) take a different approach, seeking to identify the acoustic, spectro-temporal properties of the speech signal might underlie brain specialization. In adults, it has been shown that fast-changing sounds or sounds modulated in time preferentially recruit areas in the left hemisphere that are part of the language network, while slowly changing sounds or sounds modulated spectrally tend to engage the right hemisphere (Zatorre et al., 2002; Hickok and Poeppel, 2007). This offers a potential explanation for why most language stimuli, with their fast phoneme and syllable transitions, activate the left hemisphere, with prosody being the only aspect of language that is processed in the right hemisphere. However, adults have extensive experience with language, leaving open the issue of causation. Telkemeyer et al. (2011), Arimitsu et al. (2011), and Minagawa-Kawai et al. (2011) now test these hypotheses on newborns and young infants using different temporally and spectrally modulated tone stimuli, asking whether the observed hemispheric specializations are the causes or the results of lateralized language processing. As an innovative extension of the research on early brain specialization for speech, Sato et al. (2011) investigate whether, and if yes, how this specialization might be different in an atypical population, stuttering children and adults.

The last three contributions inquire into more advanced or higher level mechanisms of language processing and comprehension. Homae et al. (2011) used a new method of NIRS data analysis to explore functional connectivity and networks in 3-month-old infants at rest and while they listen to speech stimuli, identifying a large-scale brain network engaged in language processing. Wagner et al. (2011) explore the neural correlates of learning abstract linguistic rules at 7 and at 9 months of life and show important developmental changes signaling infants increased specialization for and attunement to language structure. Tabea Brink et al. (2011) study the brain mechanisms underlying the understanding of empathy in verbal and picture-based stories in pre-school children, an age which is believed to be crucial for the development of emotional and cognitive empathy.

It is my hope that these NIRS studies further our understanding of language and cognitive development and bring us closer to bridging the gap between brain, mind and behavior at the very beginning of life.

REFERENCES

- Arimitsu, T., Uchida-Ota, M., Yagihashi, T., Kojima, S., Watanabe, S., Hokuto, I., et al. (2011). Functional hemispheric specialization in processing phonemic and prosodic auditory changes in neonates. *Front. Psychol.* 2:202. doi: 10.3389/fpsyg.2011.00202
- Benavides-Varela, S., Gómez, D. M., and Mehler, J. (2011). Studying neonates' language and memory capacities with functional near-infrared spectroscopy. *Front. Psychol.* 2:64. doi: 10.3389/fpsyg.2011.00064
- Fava, E., Hull, R., and Bortfeld, H. (2011). Linking behavioral and neurophysiological indicators of perceptual tuning to language. *Front. Psychol.* 2:174. doi: 10.3389/fpsyg.2011.00174
- Friederici, A. D. (2005). Neurophysiological markers of early language acquisition: from syllables to sentences. *Trends Cogn. Sci.* 9, 481–488. doi: 10.1016/j.tics.2005.08.008
- Gervain, J., Mehler, J., Werker, J. F., Nelson, C. A., Csibra, G., Lloyd-Fox, S., et al. (2011). Near-infrared spectroscopy: a report from the McDonnell infant methodology consortium. *Dev. Cogn. Neurosci.* 1, 22–46. doi: 10.1016/j.dcn.2010.07.004
- Hickok, G., and Poeppel, D. (2007). The cortical organization of speech processing. *Nat. Rev. Neurosci.* 8, 393–402. doi: 10.1038/nrn2113
- Homae, F., Watanabe, H., Nakano, T., and Taga, G. (2011). Large-scale brain networks underlying language acquisition in early infancy. *Front. Psychol.* 2:93. doi: 10.3389/fpsyg.2011.00093
- Lloyd-Fox, S., Blasi, A., and Elwell, C. E. (2009). Illuminating the developing brain: the past, present and future of functional near infrared spectroscopy. *Neurosci. Biobehav. Rev.* 34, 269–284. doi: 10.1016/j.neubiorev.2009.07.008
- May, L., Byers-Heinlein, K., Gervain, J., and Werker, J. F. (2011). Language and the newborn brain: does prenatal language experience shape the neonate neural response to speech? *Front. Psychol.* 2:222. doi: 10.3389/fpsyg.2011.00222
- Minagawa-Kawai, Y., Cristià, A., Vendelin, I., Cabrol, D., and Dupoux, E. (2011). Assessing signal-driven mechanisms in neonates: brain responses to temporally and spectrally different sounds. *Front. Psychol.* 2:135. doi: 10.3389/fpsyg.2011.00135
- Pena, M., Maki, A., Kovacic, D., Dehaene-Lambertz, G., Koizumi, H., Bouquet, E., et al. (2003). Sounds and silence: an optical topography study of language recognition at birth. *Proc. Natl. Acad. Sci. U.S.A.* 100, 11702–11705. doi: 10.1073/pnas.1934290100
- Sato, H., Hirabayashi, Y., Tsubokura, H., Kanai, M., Ashida, T., Konishi, I., et al. (2012). Cerebral hemodynamics in newborn infants exposed to speech sounds: a whole-head optical topography study. *Hum. Brain Mapp.* 33, 2092–2103. doi: 10.1002/hbm.21350
- Sato, Y., Mori, K., Koizumi, T., Minagawa-Kawai, Y., Tanaka, A., Ozawa, E., et al. (2011). Functional lateralization of speech processing in adults and children who stutter. *Front. Psychol.* 2:70. doi: 10.3389/fpsyg.2011.00070
- Tabea Brink, T., Urton, K., Held, D., Kirilina, E., Hofmann, M., Klann-Delius, G., et al. (2011). The role of orbitofrontal cortex in processing empathy stories in 4- to 8-year-old children. *Front. Psychol.* 2:80. doi: 10.3389/fpsyg.2011.00080
- Telkemeyer, S., Rossi, S., Nierhaus, T., Steinbrink, J., Obrig, H., and Wartenburger, I. (2011). Acoustic processing of temporally modulated sounds in infants: evidence from a combined near-infrared spectroscopy and EEG study. *Front. Psychol.* 1:62. doi: 10.3389/fpsyg.2011.00062
- Villringer, A., and Chance, B. (1997). Non-invasive optical spectroscopy and imaging of human brain function. *Trends Neurosci.* 20, 435–442. doi: 10.1016/S0166-2236(97)01132-6
- Wagner, J. B., Fox, S. E., Tager-Flusberg, H., and Nelson, C. A. (2011). Neural processing of repetition and non-repetition grammars in 7- and 9-month-old infants. *Front. Psychol.* 2:168. doi: 10.3389/fpsyg.2011.00168
- Zatorre, R. J., Belin, P., and Penhune, V. B. (2002). Structure and function of auditory cortex: music and speech. *Trends Cogn. Sci.* 6, 37–46. doi: 10.1016/S1364-6613(00)01816-7

Conflict of Interest Statement: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 12 July 2014; accepted: 01 August 2014; published online: 15 August 2014.

Citation: Gervain J (2014) Near-infrared spectroscopy: recent advances in infant speech perception and language acquisition research. *Front. Psychol.* 5:916. doi: 10.3389/fpsyg.2014.00916

This article was submitted to Language Sciences, a section of the journal *Frontiers in Psychology*.

Copyright © 2014 Gervain. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Linking behavioral and neurophysiological indicators of perceptual tuning to language

Eswen Fava¹, Rachel Hull¹ and Heather Bortfeld^{2,3*}

¹ Department of Psychology, Texas A&M University, College Station, TX, USA

² Department of Psychology, University of Connecticut, Storrs, CT, USA

³ Child Language Studies, Haskins Laboratories, New Haven, CT, USA

Edited by:

Judit Gervain, CNRS – Université Paris Descartes, France

Reviewed by:

Judit Gervain, CNRS – Université Paris Descartes, France

Krista Byers-Heinlein, Concordia University, Canada

*Correspondence:

Heather Bortfeld, Department of Psychology, University of Connecticut, Storrs, CT 06269-1020, USA.
e-mail: heather.bortfeld@uconn.edu

Little is known about the neural mechanisms that underlie tuning to the native language(s) in early infancy. Here we review language tuning through the lens of type and amount of language experience and introduce a new manner in which to conceptualize the phenomenon of language tuning: the relative speed of tuning hypothesis. This hypothesis has as its goal a characterization of the unique time course of the tuning process, given the different components (e.g., phonology, prosody, syntax, semantics) of one or more languages as they become available to infants, and biologically based maturational constraints. In this review, we first examine the established behavioral findings and integrate more recent neurophysiological data on neonatal development, which together demonstrate evidence of early language tuning given differential language exposure even *in utero*. Next, we examine traditional accounts of sensitive and critical periods to determine how these constructs complement current data on the neural mechanisms underlying language tuning. We then synthesize the extant infant behavioral and neurophysiological data on monolingual, bilingual, and sensory deprived tuning, thereby scrutinizing the effect of these three different language profiles on the specific timing, progression, and outcome of language tuning. Finally, we discuss future directions researchers might pursue to further understand this aspect of language development, advocating our relative speed of tuning hypothesis as a useful framework for conceptualizing the complex process by which language experience works together with biological constraints to shape language development.

Keywords: near-infrared spectroscopy, perceptual tuning, language development, sensory deprivation, monolingual/bilingual

INTRODUCTION

Infants tune to the specific language(s) in their environment very quickly. Much evidence suggests that an individual infant's early language exposure is critical to this tuning process, but we know relatively little about the underlying neural mechanisms that facilitate it. For example, do neural pathways that support tuning to a single language differ from those that support tuning to two (or more) languages? If the timing of language exposure matters, does early tuning to a single native language limit the neural mechanisms available for later acquisition of languages? This review examines data from behavioral and neurophysiological research, following the developmental timeline, to examine how biological maturation interacts with language experience/exposure to influence the underlying neural mechanisms that support language tuning.

In what follows, we consider when experience with language truly begins, and how that experience impacts the development of neural pathways in support of language learning. We then review classic and more recent research on neonates, young infants, and older infants to better understand what is known (and unknown) about the different stages of postnatal language learning. And, since substantial evidence indicates that language learning is influenced by an infant's particular *language profile* (i.e., single versus multiple language experience), we also explore how perceptual tuning specifically to language (i.e., "*language tuning*") fits into the oft debated concept of critical (or sensitive) periods. Finally, we explore how

data collected using neurophysiological methodologies can add to what is currently known about the processes of infant perceptual tuning, in general, and language development, in particular. Based on our synthesis of these diverse sources of data, we conclude with the introduction of a new manner in which to conceptualize language tuning: the relative speed of tuning hypothesis.

EARLY EXPERIENCE AND ITS INFLUENCE ON NEURAL AND BEHAVIORAL CORRELATES OF LANGUAGE TUNING

Strict interpretations of language development as completely biologically endowed or completely experience-based have loosened in recent years. Although researchers accept that the biological basis for learning language is in place and at work at birth, there is also substantial appreciation for the fact that changes in environment have profound effects on language outcome. Indeed, there is considerable evidence that critical interactions between neural biases and environmental shaping are at work *in utero*.

For example, external sound stimuli are "heard" through tissue and liquid barriers of the womb, which excludes frequencies greater than 5000 Hz (Jardri et al., 2008). In spite of sound filtration by the womb, numerous researchers have used behavioral and physiological measures – such as heart rate and movement – to extrapolate information about when auditory processing begins *in utero* and what sorts of distinctions the developing fetus can make at various points in gestational time. Lecanuet et al. (1995) provided some of

the first physiological (heart rate) data suggesting that fetal hearing occurs before 28 gestational weeks. In fact, the fetus responds to sound at 22 gestational weeks (Hepper and Shabidullah, 1994) and habituates to repeated sound stimuli at 32 gestational weeks (Morokuma et al., 2004). Moreover, as babies near term, their sensitivity to more complex auditory stimuli improves, allowing them to perceive variations in music (Kisilevsky et al., 2004) and differentiate between prosodic cues in familiar and novel rhymes (Decasper et al., 1994). Although these studies illustrate a suprasegmental level of language exposure and processing, our concept of true language “experience” should not begin only at birth, as other aspects of language are likewise capable of passing through the uterine wall. This implies a currently unknown threshold in prenatal auditory processing. To better understand the true nature of the intrauterine experience of speech, further research is needed.

The domain of fetal development highlights how biology and environment have already combined prenatally to set the process of language learning in motion. Our position is that the development of neural pathways supporting language begins *in utero*, and is shaped by the interaction of biology and language experience even then. Understanding the extent and implication of that interaction will require data, both behavioral and neurophysiological, from multiple sources. However, in order to begin to examine language tuning, we must first understand to which aspects of language infants are initially sensitive. Fortunately, we have a wealth of data that can inform us about the neural mechanisms already available to support language development once a baby is born.

BEHAVIORAL AND NEUROPHYSIOLOGICAL EVIDENCE OF NATIVE/NON-NATIVE SPEECH PERCEPTION IN NEONATES

The most compelling initial evidence about the precocity of infants’ language abilities came from a series of behavioral studies showing that neonates are capable of performing various cross-language discriminations (Mehler et al., 1978, 1988; Moon et al., 1993; Nazzi et al., 1998; Ramus et al., 2000). Specifically, these studies revealed that neonates prefer their native language over another language with a dissimilar rhythmic structure. Furthermore, this preference – for the prosodic pattern of the native tongue – has been shown to stem from prenatal experience with native speech (Moon et al., 1993). The preference remains constant as well, including when babies are exposed to two languages (English and Tagalog) *in utero*. In the case of two languages, infants display equal preference for each when tested as newborns (Byers-Heinlein et al., 2010). Moreover, the data show that Chinese–English bilingually exposed neonates showed intermediate patterns of preference for English compared with Tagalog, but importantly, these newborns are able to discriminate English from Tagalog. Thus, given prenatal experience with more than one language, neonates appear capable of discriminating between the two.

Neonatal near-infrared spectroscopy (NIRS) studies have added support to these behavioral findings, demonstrating distinct hemodynamic response patterns given different syntactic structures (Gervain et al., 2008). Specifically, Gervain et al. (2008) observed increased hemodynamic activity to a repeated syllable sequence in both temporal and left frontal regions in contrast to an unchanging hemodynamic response to control sequences. Differences between the two auditory conditions were observed in

the first few trials of the study, as well as across the course of the experiment. The increased hemodynamic response to the repetition sequence in comparison with the control sequences suggests that neonates possess a (perhaps automated) neural mechanism responsible for detecting repetitions. Furthermore, a familiarity effect was inferred by the researchers from the increased hemodynamic response observed during later trials in response to repetition and not to control sequences. As the authors concluded, these data may demonstrate an early neural sensitivity to configurations of auditory stimuli that are often heard in speech.

Near-infrared spectroscopy has also been used to test neonates between 2 and 5 days of age as they listened to recorded speech samples in their native language. Peña et al. (2003) found that native language processing elicited focal regions of activation, including in the dorsolateral prefrontal cortex, the primary and auditory association cortices, and the supramarginal gyrus (a portion of Wernicke’s area). Results from this study demonstrated that monolingual neonates already show increased left relative to right temporal activation in response to forward relative to reversed speech. More recently, Saito et al. (2007) used NIRS to demonstrate that neonates discriminate between infant- and adult-directed speech, attending more to and showing greater hemodynamic responses to the former than to the latter in bilateral frontal regions.

Exciting recent findings (May et al., under review), in which NIRS was used to record monolingual neonates’ neural responses during exposure to auditory-only, low-pass filtered sentences in forward and backward native (or familiar) English, and non-native (or unfamiliar) Tagalog, showed that similar channels are activated for forward but not to backward speech in these infants. Moreover, there was no difference in lateralization observed for the two languages, with both eliciting bilateral hemodynamic functions. These findings suggest that similar regions are used to process native and non-native speech in monolingual neonates. However, infants in this study also showed no significant difference in response to forward and backward English conditions, a finding that contrasts with previous results obtained using unfiltered speech (Peña et al., 2003). The authors posit that the focused prosodic cues available in the low-pass filtered speech may have been driving the bilateral patterns of activation observed, as well as the atypical results concerning forward and backward English. Notably, the same stimuli were utilized in a previous study, which showed that bilingual exposure *in utero* resulted in bilingual neonates distinguishing between English and Tagalog (Byers-Heinlein et al., 2010).

Overall, these results add support to the view that left-lateralized language processing mechanisms are in place at birth. Of course, further research is needed to clarify the somewhat mixed neurophysiological evidence for differences in sensitivity to native compared with non-native speech processing in this age group (for a theoretical review of the NIRS left-lateralized speech literature see Minagawa-Kawai et al., 2011b). In particular, and in light of the comparison between speech and non-speech auditory stimuli with comparable complexity, the increased hemodynamic activity demonstrated to be specific to speech is compelling. Nonetheless, these data have not simplified theoretical debates about the degree to which nature and nurture come into play differentially in early language development. Rather, they have served to push the focal age for this debate ever earlier in development. Because of the paucity

of neonatal (much less prenatal) data, we instead will focus on a well established postnatal phenomenon – categorical perception of speech sounds – to highlight how behavioral evidence has demonstrated that environment shapes biology strongly and quickly in the first year of life.

BEHAVIORAL AND NEUROPHYSIOLOGICAL EVIDENCE OF NATIVE LANGUAGE SENSITIVITY IN NEONATES

Given short exposure to the unfiltered speech signal, newborns demonstrate an impressive ability to process their native language. In spite of this limited experience, neonates can distinguish native speech from other complex, non-speech auditory stimuli when these are controlled for spectral and temporal factors (Vouloumanos and Werker, 2007). Recent research using NIRS has uncovered evidence of the neural mechanisms underlying these behavioral findings. Specifically, NIRS data demonstrate that the cortical areas utilized during speech processing are distinct compared with other auditory stimuli (for a review see Lloyd-Fox et al., 2010). For example, Kotilahti et al. (2010) compared music and speech processing in neonates. Overall, although no significant activation over baseline was observed in the right hemisphere for music or for speech, the auditory stimuli elicited bilateral activation, with increased responses in left relative to right temporal cortex during speech relative to music.

Apart from speech/non-speech distinctions, *in utero* experience with speech appears to be sufficient for newborns to differentiate between familiar speakers and stories in the context of native speech. For example, behavioral evidence clearly demonstrates that, within the first day of life, infants prefer their mother's voice over that of another woman (Decasper and Fifer, 1980). In addition, newborns can discriminate between a familiar (i.e., familiarized given prenatal exposure) and a novel story in their native language (Garnicka, 1977; Stern et al., 1983; Albin and Echols, 1996). They also discriminate between their own and another language from a different language family at birth. This recognition of familiar, native speech also extends to neurophysiological responses elicited from a single participant using electroencephalography (EEG). Radicevic et al. (2008) tested a single infant at 24 and 75 days of age. The child was read a story by the mother in her native tongue before and after birth (from 27 gestational weeks to 1 week before birth, then after birth for 7 days). Following birth, the infant was tested on several conditions: familiar (mother's) voice/familiar content, unfamiliar voice/familiar content, unfamiliar voice/familiar content in a non-native language. At 24 days, the familiar voice and content elicited a different delta rhythm from the non-native content in an unfamiliar voice, which resembled the resting state. In addition, at both ages, the unfamiliar language and voice elicited a response similar to rest. Moreover, 75-day measurements revealed similar delta and theta rhythms for both familiar and unfamiliar content regardless of language type or speaker familiarity.

BEHAVIORAL EVIDENCE OF PERCEPTUAL TUNING AND IMPLICATIONS FOR NEURAL CORRELATES

Perceptual tuning is a complex developmental phenomenon that has been the sole topic of substantial review papers (Sebastian-Galles, 2002; Werker and Tees, 2005). Here we focus specifically on the interplay between language environment and the neural

mechanisms that support this tuning process. We will briefly discuss behavioral evidence for language-specific perceptual tuning, and follow up with a more detailed discussion of those neurophysiological data that can inform us about the neural sensitivities present during infancy that may underlie this.

Language tuning is the narrowing of perception of speech sounds over the first year of life, from an initially broad ability to distinguish many minimally contrastive phonemes to an increasingly specialized capacity to distinguish (for the most part) only those phonemes relevant to one's ambient language (Eimas et al., 1971; Jusczyk et al., 1977; Werker and Tees, 1983, 1999; Werker and Lalonde, 1988; Polka and Werker, 1994; Jusczyk, 1997). For the purpose of this discussion, the term "perceptual tuning" will be used synonymously with "perceptual narrowing" and "perceptual reorganization" (Best, 1994). As we have already seen, infants demonstrate tremendous skill in processing the speech stream from an early age, and a substantial behavioral literature has demonstrated that perceptual tuning to native speech occurs relatively early in development. Specifically, between 10 and 12 months of age infants with monolingual language profiles become more adept at discriminating native compared with non-native phonemic contrasts (Werker and Tees, 1984). In addition, Kuhl et al. (2006) demonstrated that monolingual English-exposed and monolingual Japanese-exposed 6- to 8-month-old infants were able to make the English/r/and/l/distinction, but by 10–12 months of age, only the monolingual English-exposed infants continued to differentiate these phonemes.

These data are further supported by event-related potential (ERP) data from 4- to 5-month-olds demonstrating different mismatch negativity signatures during exposure to pseudowords with a stress pattern common to the native language versus one uncommon to that language (Friederici et al., 2007). Moreover, another ERP study found infant responses indicative of both native and non-native consonant contrast discrimination in 7-month-olds, whereas 11-month-olds did not demonstrate such sensitivity contrasts (Rivera-Gaxiola et al., 2005). However, when analyses were based on a different parsing of ERP components, 11-month-old infants appeared to still be sensitive to both types of contrast. These conflicting results (though given differing methods of data analysis) suggest that differences in native compared with non-native consonant contrasts may not be as robust as once thought, thus requiring further investigation, particularly using ERP methods. In addition to conflicting results on consonant contrasts, behavioral findings have suggested a somewhat earlier timeline (e.g., 6–8 months) for vowel discrimination (Kuhl et al., 1992; Polka and Werker, 1994). Thus, one way to conceptualize perceptual narrowing is to view experience with a native language as sharpening the boundaries between native contrasts (Aslin and Pisoni, 1980; Kuhl et al., 2001; Polka et al., 2001).

Furthermore, recent work by Narayan et al. (2010) has uncovered evidence that some phoneme contrasts are not available "pre-tuning,"¹ while others (even non-native) remain available after tuning should (arguably) have ended (i.e., after 8 months). Specifically, these researchers found that the non-native/na/-ŋa/

¹The pre-tuning label is reflective of the traditional behavioral literature's timeline for language tuning in monolingual infants (i.e., before 6 months of age).

contrast was *not* perceived by English 4- to 12-month-olds, nor was it perceived by Filipino 6- to 8-month-olds, for whom it was a native contrast. However, older Filipino 10- to 12-month-olds were able to make the distinction. This perceptual pattern is unusual, as non-native contrasts are typically perceived by pre-tuned infants of any language background (Trehub, 1976; Werker and Tees, 1984; Polka et al., 2001). This account highlights the intuitively appealing point that the perceptual tuning process is mediated by the *relative* difficulty or ease of speech signal processing (termed “acoustic salience” by Narayan et al. (2010)). Given the behavioral evidence for perceptual tuning to speech over the first year of life, it is reasonable to assume that the neural mechanism(s) responsible for this process may be undergoing a concurrent and pronounced period of development. This could be considered the key critical, or sensitive, period.

POSSIBLE NEURAL MECHANISMS FOR LANGUAGE TUNING IN INFANCY

Needless to say, language development researchers have long debated the timing and degree of any so-called critical period (Johnson, 2001, 2005; Werker and Tees, 2005; Armstrong et al., 2006; Thomas and Johnson, 2008). The following summary of the critical/sensitive/optimal outcome period literature highlights how different perspectives on the issue may inform our understanding of the possible neural mechanisms underlying language tuning.

A *critical period* has been described as a fixed time range during which an organism’s neural processing and behavior can be influenced by external environmental input (e.g., Werker and Tees, 2005). In contrast, a sensitive or optimal period (we will use the term *sensitive period*) is conceptualized as having a variable offset that depends on the organism’s experience and learning (Knudsen, 1999), making the system adaptable to changing environmental inputs for a more flexible length of time. Importantly, Werker and Tees (2005) have reviewed evidence suggesting that different aspects of language (e.g., syntax, phonology, morphology, semantics) may each have their own critical period, or at least develop in a(n) “inter-related” or “nested” set of critical periods. Specifically, as infants gain more experience with language, they become (incrementally) more aware of progressively complex components and of variability in the speech signal.

Behavioral evidence of sequential mastery of different aspects of language has generated several theories about the plasticity of neural mechanisms in early development. Neural plasticity is a state of functional changes within the dynamic, iterative process of neural development. Although the current literature has yet to tease apart the functional significance of the increase in synaptogenesis and subsequent pruning of those synapses during infancy, some have suggested a link between peaks and subsequent elimination of synapses co-occurring with more mature function (Webb et al., 2001). While there are certainly dynamic and plastic aspects to the normally developing brain, the definition of plasticity we will use here focuses on the connections that are selectively retained post-pruning because they are used to process input (Stiles, 2000).

Theories applicable to language learning must therefore address plasticity and at least somewhat account for the influence of early language experience on the process. Trainor (2005) has proposed two possible explanations for the particularly strong influence of

environment on neural development during infancy. The first, the *genetically mediated account*, posits that neural substrates’ capacity to change and adapt in response to environmental inputs (such as those that drive language tuning), is solely dependent upon maturation. Thus, infants would be expected to lose neural plasticity by a certain age regardless of environment, a position congruent with a “critical period” account of language tuning. Trainor’s (2005) second explanation, the *experientially mediated account*, holds that experience facilitates brain organization by reducing neural plasticity as connections become more functionally specified. This theory is very similar to the *neural commitment hypothesis* (Kuhl et al., 2005b; Kuhl and Rivera-Gaxiola, 2008), which asserts that language environment shapes the creation of infants’ neural connections (i.e., neural commitment). In particular, a language profile is thought to define the parameters of a “mental filter.” Input that does not get through this filter is less effectively processed. Both of these theories are congruent with a sensitive, rather than critical, period perspective on language tuning, because experience (rather than time-based maturation) controls the reduction in neural plasticity.

Thomas and Johnson (2008) have reviewed several theories of the mechanisms underlying such period, and at least two of these may have direct relevance to understanding the neural basis for language tuning. First, their mechanism of *termination of plasticity* regulates the rate of synaptic pruning, such that pruning increases with the biological maturation of brain regions and as they become specialized to perform specific functions, such as sensory, motor, or cognitive processing. As with the genetically mediated account discussed by Trainor (2005), this perspective implies a biologically programmed (rather than experience-based) process, and therefore would support a critical period view of language tuning. One could argue that the initially robust ability of neonates to discriminate speech-sound contrasts from a variety of different specific languages, followed by the marked reduction in this ability months later, could be explained by a termination of plasticity or genetically mediated account. However, as some non-native contrasts remain available after the language tuning period ends (e.g., Narayan et al., 2010), and given that bilingually exposed infants have shown a relatively extended period of sensitivity to contrasts (Werker and Byers-Heinlein, 2008), it is difficult to completely reconcile the language tuning process with these strictly maturation-based accounts.

Thomas and Johnson (2008) *self-termination of learning account* relies instead on a mechanism in which learning itself is thought to create neural changes that reduce plasticity. Similar to the previously discussed experientially mediated account (see Trainor, 2005) and the neural commitment hypothesis (Kuhl et al., 2005b), this view holds that responses of particular brain regions emerge largely from their functional and anatomical connections with other areas; in other words, their specialization is ultimately activity-dependent. As such, these experience-based explanations are consistent with a sensitive period view, in which the narrowing of categorical perception of phonemes that occurs during language tuning could be explained as a function of the infant’s exposure (or lack thereof) to those phonemes.

Experience-based explanations also complement an “acoustic salience” account of language tuning (Narayan et al., 2010), in which differential mastery for phoneme contrasts (with more salient ones being mastered first) begins to equalize as an infant

gains experience with the less salient ones over time. These could also explain why sensitivity to some rarely encountered non-native phonemes, or native language non-speech sounds used to communicate expressions (e.g., clicks used to convey exasperation or pity by an English-speaker), remain when others are lost (Best et al., 1988). Finally, experience-based accounts could reveal why neonates with bilingual language profiles *in utero* show equal preference for both languages (Byers-Heinlein et al., 2010), whereas their monolingually exposed counterparts show preference only for their single native language (Mehler et al., 1988; Moon et al., 1993; Nazzi et al., 1998; Ramus et al., 2000).

Although the exact mechanisms underlying language tuning – whether critical or sensitive in nature – are still debated, the general concept of such a period has important practical implications, as studies have time and again demonstrated that the timing of infants' language exposure (and thus tuning) may predict future language abilities. For example, early perceptual sensitivity to simple non-native syllable contrasts at ages 6, 14, 15, and 24 months has been shown to be correlated with subsequent language achievement, as established by the MacArthur Communicative Developmental Inventory performance on items such as phrases understood, words understood, and words produced (Tsao et al., 2004).

Kuhl et al. (2005b) have also reported that skilled early perception of native language phonemes can be a reliable predictor of later success in monolingual language development; in contrast, relatively sustained sensitivity to non-native language phoneme discrimination predicted a slower rate of language development. Kuhl et al. (2005b) argued that many markers of language development, including number of words produced, sentence complexity, and mean length of utterance, may be predicted based on an infant's early skill – or lack thereof – in discriminating native language phonemes. However, it is not clear from these behavioral data what might be driving the differences in the timing of language tuning and language development and its relation to subsequent language learning. We turn now to outcomes from neuroimaging and ERP studies that have attempted to identify possible neural mechanisms for this process.

NEUROPHYSIOLOGICAL EVIDENCE OF LANGUAGE TUNING IN INFANCY

Thus far, neurophysiological research on language learning in infants has revealed evidence of functional processing differences that may speak to the behavioral differences outlined in the previous section. In general, neurophysiological data demonstrate left-lateralized processing for language stimuli across the first 18 months of life. Notably, we will use the term *lateralization* to denote bilateral activity with significantly greater activity in the left compared with the right hemisphere. Most data on language processing fall into the category of lateralized, with bilateral activation being observed, though with one hemisphere (usually the left in the case of speech processing) showing significantly more activity than the other.

For example, fMRI data from typically developing, sleeping 2- to 3-month-olds demonstrated a significant left asymmetry (lateralization) for forward (but not backward) native speech (the infants were monolingually exposed), particularly in posterior

superior temporal regions, whereas the same study showed bilateral responses in other regions (e.g., right dorsolateral prefrontal cortex; left angular gyrus) in otherwise matched 2- to 3-month-olds (Dehaene-Lambertz et al., 2002). This lateralized response to forward, but not backward, speech was also observed in a NIRS study with sleeping neonates (Taga et al., 2007).

More recently, several NIRS studies have focused on testing awake and attentive infants to address the role of attentional state in infant language processing (Lloyd-Fox et al., 2010). For example, Taga and Asakawa (2007) exposed monolingual Japanese 2- to 4-month-olds to unfamiliar native words and found bilateral temporal activation. Notably, the majority of stimulus presentation in this case was audiovisual, as the three-word stimuli were presented at different alternating and overlapping intervals with flashing (4 Hz) checkerboard. In contrast, Minagawa-Kawai et al. (2011a) exposed monolingual Japanese 4-month-olds to short sentences from auditory-only film dialogs or a speech database delivered by a male or female speaker. Analyses revealed left-lateralized hemodynamic activity for native compared with non-native speech conditions. Further evidence in support of the left-lateralization of native speech processing during infancy includes a series of NIRS studies from our own lab, demonstrating increased activation in the left relative to the right temporal regions in response to native language stimuli in older monolingual 6- to 9-month-old infants (Bortfeld et al., 2007, 2009).

Interestingly, a cross-sectional NIRS study using 3- to 28-month-olds reported outcomes consistent with the idea of a reduction in plasticity for language tuning (Minagawa-Kawai et al., 2007). Awake monolingual infants were exposed to a set of four pseudowords that varied in terms of the duration of the last vowel. Specifically, two of the stimuli matched native language characteristics while the remaining two matched non-native characteristics. Total (i.e., a sum of oxygenated and deoxygenated) hemoglobin² results showed significant left-lateralization in response to the native as compared to the non-native contrast in most of the older infants, whereas all of the youngest infants showed bilateral activation. Moreover, only older infants showed significant left-lateralization for the native contrasts. These results suggest that the emergence of lateralization to the left corresponds to the neural tuning to a specific language and corresponding reduction in plasticity, at least for monolingually exposed infants.

Discrepancies in the current literature may be at least partially explained by differences across experiments in methodologies, specific cortical measurement locations, details of the speech stimuli used (e.g., words, syllables, or sentences), attentional states (i.e., awake versus asleep), and whether the stimuli are auditory or audiovisual in nature (or mixed). For example, studies in which infants cycle in and out of sleep (e.g., Dehaene-Lambertz et al., 2002) appear to demonstrate different patterns of hemodynamic activity given these two different states of alertness. It would make sense that an infant's attentional state would impact their processing of speech and understanding the role that attention plays in neurophysiological measures will require addi-

²Total hemoglobin is an accepted but less utilized measure of hemodynamic activity because oxygenated hemoglobin generally provides the strongest signal to noise ratio (see Lloyd-Fox et al., 2010).

tional research. It is also possible that different aspects of dynamic stimuli become more or less salient when presentation changes from uni- to multi-modal (Taga and Asakawa, 2007) and thus what is being measured may be different from studies in which stimulus presentation is consistently auditory or consistently audiovisual (Bortfeld et al., 2007, 2009; Minagawa-Kawai et al., 2011a). By the same token, studies that use auditory-only stimuli may differ from those that use audiovisual stimuli. Differences such as these across studies could be the source of variability in the degree of laterality observed. Nonetheless, at least for monolingually exposed infants, there seems to be converging evidence for relatively more involvement of the left than the right hemisphere during speech processing. However, this left-lateralized neurophysiological pattern of language processing may be influenced by experience with only language. Thus, we will next review the bilingual literature to examine the effect of additional language experience on the behavioral and neurophysiological aspects of language tuning in infancy.

LANGUAGE TUNING IN BILINGUALLY EXPOSED INFANTS

Theoretical accounts of the neural mechanisms involved in language tuning necessarily grapple with the interactive nature of biological and experiential factors. This dynamic relationship lies at the heart of data showing the fundamental impact of language environment on the language tuning process. Specifically studying the process of language tuning in bilingual infants, as compared to that in monolingual infants, is an important way to understand language tuning in light of a more diverse language profile (i.e., more complex experience). This holds the biological aspect of language constant between populations – in particular, when socioeconomic and other factors are likewise controlled, while contrasting the influence of language experience.

Of course, the degree to which bilingual language tuning differs from monolingual language tuning may depend upon the languages involved. The extant data suggest that comparison of bilingual and monolingual tuning timelines should be considered in the context of several factors, including the total number of contrastive phonemes within each language, the relative frequency of contrasts, the level of overlap between different categories across the languages, and, finally, the amount of exposure to each language profile (Burns et al., 2007; Sebastian-Galles and Bosch, 2009). These factors, in addition to choice of experimental paradigm, may account for the varied results currently available on the timing of language tuning in studies of monolingual versus bilingual infants (Curtin et al., 2011).

Current evidence suggests that *early* bilingual infants' language processing is similar to that of monolingual infants. Specifically in terms of differentiating native languages from one another, newborns with prenatal bilingual experience can discriminate (i.e., demonstrate equal preference for) their two *rhythmically distinct* languages at birth (Byers-Heinlein et al., 2010). Moreover, and relatively on par with the monolingual processing timeline, 4-month-old bilingual infants can discriminate either of two *rhythmically similar* languages from an unfamiliar language (with a different orientation response than monolinguals that is perhaps due to identifying the languages spoken before orienting; Bosch and Sebastian-Galles, 1997).

Thus, at least early in development, the pace of language tuning appears similar between infants with monolingual and bilingual language profiles.

Historically, results on bilingual language processing in infancy have appeared conflicted because bilingual infants have demonstrated language tuning at times on pace with, and at others lagging behind, their monolingual peers. Three studies in particular illustrate how monolingual and bilingual infants can retain similar language tuning timelines. First, using a habituation procedure, Burns et al. (2007) demonstrated that monolingual English and bilingual French–English 6- to 8-month-olds discriminated both English (i.e., /ba/and/pa/) and French voice onset time (VOT) contrasts (i.e., /ba/and/p^ha/) equally well. After 10 months, however, monolinguals only discriminated their native language contrast, while bilinguals remained able to make both distinctions. The authors note that these phonemic contrasts were high in frequency in both languages, and were likely unambiguous across languages, making this task relatively easy.

These results are congruent with findings using the head-turn preference paradigm to test monolingual English infants, monolingual Welsh infants, and English–Welsh bilingual infants, who demonstrated significantly longer looking times for familiar compared with unfamiliar VOTs (i.e., /ball/and/tall/) in their native language(s) at approximately 11 months of age (Vihman et al., 2007). Furthermore, Sundara et al. (2008) utilized an infant-controlled visual habituation procedure and exposed English monolinguals, French monolinguals, and French–English bilinguals to a contrast of the syllable/dae/, as the initial/d/phoneme differed in place of articulation for English and French (dental versus alveolar) and thus contrasted allophonically. Six- to 8-month-old infants from all three groups distinguished between these contrasts, while only the monolingual English and bilingual 10- to 12-month-olds were able to distinguish this contrast. These data suggest that, given highly frequent, similar phonemes with overlapping distributions across languages, bilingual infants remain on par with their monolingual counterparts. Thus, when overlap is coupled with high frequency in similar phonemes, overlap is unlikely to be a source of confusion or to cause delay in language tuning in bilingual infants.

However, other studies suggest language tuning may be delayed in bilingual infants compared with their monolingual counterparts. For example, Bosch and Sebastian-Galles (2003) used a familiarization/preference testing procedure to highlight differences in language tuning between monolingual and bilingual infants in the second half of the first year. Infants listened to disyllabic pseudowords with a stress pattern common to both languages, where the first vowel contrast was phonemic in Catalan, but not in Spanish (i.e., /e/versus/ɛ/). Following exposure to the variable tokens of one pseudoword, they were then tested on their ability to discriminate between either two new tokens of the familiarized pseudoword or two new tokens of the alternate (novel) pseudoword. Analyses revealed that monolingual infants responded to the familiarized stimuli by 8 months of age, whereas bilingual infants demonstrated reliable familiarization–preference at 4 months of age and again at 12 months, but not at 8 months. The authors attributed the transitory failure in phonemic discrimination at 8 months to bilinguals' denser distribution of phonetic space than monolinguals, which

could render recognition of salient distinctions while ignoring unimportant ones more difficult. Indeed, comparable behavioral delays in language tuning were observed in *monolingual* 8-month-old infants exposed to a crowded distribution of vowels (Sabourin et al., 2003) and stops (Conboy and Mills, 2006; Sundara et al., 2006, 2008; Fennell et al., 2007). As previously noted (Burns et al., 2007; Sundara et al., 2008), in addition to dense phonemic space, the frequency of the language contrasts used in the ambient language may also contribute to the relative timing of bilingual language tuning. Thus, it appears that both monolingual and bilingual infant populations cope with dense phonetic space by extending flexibility about this aspect of their language in the tuning process.

In order to further examine the roles of frequency and distribution of sounds on language tuning, Sebastian-Galles and Bosch (2009) tested 4-, 8-, and 12-month-old Spanish/Catalan bilingual infants with two vowel contrasts (i.e., /o-u/ and /e-u/). As noted by the authors, the vowels chosen here were contrastive in both languages and were found in sparse phonemic space. Furthermore, the Spanish /u/ is more infrequent than /o/, with the opposite distribution pattern found in Catalan. Using the /o-u/ contrast, a U-shaped pattern of discrimination was observed, where 4- and 12-month-olds were able to distinguish between the contrasts, while 8-month-olds were unable to perform this discrimination task, which was consistent with past data (Bosch and Sebastian-Galles, 2003). To further eliminate sources of ambiguity that may impact performance, Sebastian-Galles and Bosch (2009) tested 8-month-old bilingual infants using tokens uttered by fewer speakers, as well as tokens from a single speaker (i.e., reducing speaker-induced variability in the stimulus set). The data demonstrated that infants remained unable to perceive the /o-u/ contrast, suggesting that variability introduced by multiple speakers does not affect performance. Given the more acoustically distinct /e-u/ contrast, bilingual 8-month-olds (in addition to their monolingual peers) were able to discriminate the contrast. Thus, based on the properties of the /o-u/ contrast, it seems unlikely that the statistical properties of phonemic contrasts were the only factors influencing language tuning. The authors concluded that other factors, such as the degree of lexical similarity of the languages within the profile, the type of contrast studied (i.e., vowel or consonant), the density of the phonemic environment, and “socio-indexical” factors (e.g., language-switching) are more likely to explain differences in performance for the /o-u/ compared with the /e-u/ contrasts. Clearly, the issue will require additional data before any concrete conclusions can be drawn. Taken together, results from these behavioral studies tell us that bilingual language tuning is a complex process likely impacted by many factors, some of which also affect monolingual infants, and others that are unique to this population (Curtin et al., 2011).

In addition to all this, the choice of experimental paradigm may also affect the outcomes of bilingual tuning studies. For example, a categorization task administered through a visual choice paradigm revealed a somewhat different developmental timeline for the tuning process than more traditional testing methods. The visual choice method is an adaptation of the anticipatory eye movement paradigm previously used as a change detection measure with speech stimuli (McMurray and Aslin,

2004). Albareda-Castellot et al. (2011) employed this approach in a categorization study that began with an attention-getting, visual reinforce. When the infant attended to the reinforcer, it moved behind a T-shaped occluder and a disyllabic word from one of two vowel categories was played three times (Albareda-Castellot et al., 2011). Following initial occlusion, the reinforcer reemerged from the left or the right of the occluder, with the location predicted by the auditory stimulus. The reinforcer then reoriented the infants in preparation for the next trial. Based on the assumption that if they have already tuned to the auditory stimulus, infants will look in the direction predicted by that stimulus, the researchers analyzed the proportion of correctly anticipated place-of-emergence trials. They found that Catalan–Spanish bilingual infants were appropriately tuned to a contrast common to both languages by the age of 8 months. They further observed that the bilingual infants kept pace with their monolingual peers for contrasts found in only one of the bilingual infants’ languages (i.e., by 8 months). Thus, these results contrast with those from the habituation study by Bosch and Sebastian-Galles (2003). Given that the same contrasts were utilized in both studies, the conflict suggests that experimental design and choice of paradigm influenced an infants’ demonstration of tuning. The possibility that the type of performance required in an experimental paradigm (rather than competence or ability) may account for the differential timing of monolingual and bilingual language tuning observed in previous studies highlights the need for replication with and across paradigms.

Furthermore, although Albareda-Castellot et al. (2011) results suggest that monolingually and bilingually exposed infants may experience a similar onset of native language tuning, it is not clear whether the degree of *attainment* (that is, competence) of language tuning is similar. It is also unclear precisely which differences in task demand may be driving the different outcomes between the familiarization–preference procedure and the visual choice method. One possibility lies in the “testing” portion of these two paradigms. For example, the familiarization–preference procedure can demonstrate discrimination between a *single* familiarized token (F) and a novel token (N) at test (F versus N; Quinn, 2002). Thus, using the familiarization–preference procedure, only one of the two stimuli at test (F and not N) are presented in a context or enriched manner (i.e., as a single repetition); this enrichment facilitates processing and has been termed intersensory facilitation, (Bahrick and Lickliter, 2000, 2002). In contrast, the visual choice method (or anticipatory eye movement paradigm) pairs each auditory stimulus with an associated spatial location. Given that intersensory facilitation is more apparent in tasks with redundant multi-modal information (i.e., information highlighting differences between stimuli), which could include perception of speech stimuli (Bahrick et al., 2010), it is thus possible that the visual choice paradigm allows infants to demonstrate more mature language tuning behavior due to its use of intersensory facilitation. This is in contrast to the familiarization–preference procedure, which uses an enriched presentation for only one of the two stimuli being tested, thereby limiting the observable results. Regardless, the discrepancies between these paradigms make it clear that future studies are needed to examine whether the timeline of bilingual language tuning may actually be shifted,

or whether infant cognition may capitalize on different measurement procedures differently (Yoshida et al., 2009; Mattock et al., 2010; Albareda-Castellot et al., 2011).

In addition to the behavioral studies of bilingual language tuning in infancy, a handful of studies utilizing the ERP technique have examined neurophysiological correlates of language processing in bilingual infant populations. While these studies do not speak directly to differences in phonetic perception between monolinguals and bilinguals, they do provide insight into potential neurophysiological indicators of bilingualism in infancy. For example, a seminal study by Conboy and Mills (2006) examined 19- to 22-month-old English–Spanish bilinguals' neurophysiological responses to familiar and unfamiliar words in the children's dominant and non-dominant languages in the context of vocabulary size. These researchers found that the organization of the responses to words varied according to both individual and total vocabulary size, with higher producers producing a significant left-lateralized (P100) response in their dominant language. No such lateralization effect was observed in low producers. Furthermore, the researchers observed significantly different latencies (N200 versus N400 responses) for known compared with unknown words. Specifically, unknown words elicited a right lateralized response in the dominant language, while the non-dominant language demonstrated no such lateralization. Thus, these results demonstrate that language ability, as measured by vocabulary size, influences speed of processing and lateralization for processing the dominant and non-dominant languages.

Finally, Vihman et al. (2007) tested English monolingual and Welsh monolingual infants, as well as English–Welsh bilingual infants, all between 9 and 12 months of age, with ERP. Using familiar and unfamiliar word stimuli, these researchers observed the first neurophysiological indicator (i.e., in the form of an N2 response) of word form recognition at 10 months of age in English monolinguals, where behavioral data has shown such an effect only at 11 months of age. Furthermore, the ERP familiarity effect for English monolinguals vanished at 12 months in both paradigms, possibly indicating that the neurophysiological indicator of word form recognition is fundamentally different from the behavioral measure. Importantly, these authors noted that the Welsh monolingual infants showed no neurophysiological differences for the stimulus types, which may be due to unique qualities of the Welsh language (including mutation, reliance on later parts of words, and sociolinguistic factors). In contrast, 11-month-old English–Welsh bilingual infants showed the effect of familiarity in both languages and in both procedures, which may be due to the influence of English language learning on the Welsh language profile or to some of the other factors already mentioned. As with previously reviewed studies, elucidation of these findings will require additional research.

Ultimately, these data demonstrate that a bilingual language profile in infancy may result in different behavioral and neural consequences for language tuning in these (bilingual) infants compared with their monolingual peers. Numerous factors, including the languages involved in the bilingual language profile, the sociolinguistic environment, the level of overlap between the two languages, the total number of contrastive elements, and the frequency

distribution of the input, all likely play a role in bilingual language tuning. Therefore, the take-home message about the time course of bilingual language tuning is, at present, unclear. However, the extant data certainly indicate that there are differences between bilingual and monolingual language profiles in the respective rates of behavioral evidence of tuning, patterns of related hemodynamic activity, and ultimate outcomes in language proficiency.

LANGUAGE TUNING IN BILINGUALLY EXPOSED ADULTS

Although data from bilingually exposed infants are sparse and not always consistent, substantial evidence from bilingual adults has shown that age of second language (L2) acquisition and L2 proficiency may dramatically influence language organization in the brain as well (Vingerhoets et al., 2003; Wartenbuser et al., 2003; Briellmann et al., 2004; Indefrey, 2006; Hull and Vaid, 2007). A structural MRI study involving monolinguals and early and late bilinguals found relatively increased gray matter density in the bilinguals in the inferior parietal lobe near the temporo-parietal junction, and significantly more so for early than late bilinguals (Mechelli et al., 2004). This suggests that the onset age of bilingualism and/or the length of L2 experience may alter the brain's actual structure.

In addition to anatomical differences in structure, numerous behavioral, and functional imaging studies have indicated differential language organization for early bilinguals relative to late bilinguals and monolinguals. For example, an fMRI study of multilingual adults with varying ages of L2 acquisition and different levels of L2 proficiency has demonstrated that late multilingual speakers consistently show left-lateralization for processing all languages, regardless of proficiency (Briellmann et al., 2004). However, the lone early multilingual in this study appeared to show bilateral activation patterns for all languages, including the (non-proficient) language the participant had acquired in adulthood (i.e., well outside the language tuning period in infancy). These outcomes would suggest that the neural substrates for language are different for bilinguals who acquired one (late bilinguals) versus multiple (early bilinguals) languages during early development.

Findings for differential neural bases for language processing in early and late bilinguals have also been demonstrated in languages that are not audiovisual in nature. For example, an fMRI study involving deaf and hearing sign language users showed that both hearing and deaf early American Sign Language (ASL)–English bilinguals demonstrated bilateral activation during sign language processing (Neville et al., 1998). In contrast, late ASL–English bilinguals and ASL or English monolinguals displayed primarily left-lateralization of processing. These studies suggest that early exposure to multiple languages may result in recruitment of bilateral neural substrates – regardless of whether one of the languages is audiovisual and the other is not. Conversely, early exposure to a single language, again regardless of whether it is audiovisual or signed, may result in a left-lateralization for language. Finally, these outcomes suggest that the end result of language tuning may not only affect the languages present during that tuning process, but also any subsequently acquired languages.

In terms of behavioral evidence for differences in brain organization for language between monolinguals and early and late bilinguals, two meta-analyses of the language laterality literature have provided support for this view. One meta-analysis specifically compared behavioral outcomes from monolinguals and early and late bilinguals; bilinguals were individually coded for age of L2 acquisition (early, late), and level of L2 proficiency (proficient, non-proficient), and only the first language of bilinguals that matched the language of the monolinguals was assessed (Hull and Vaid, 2006). The authors observed that late bilinguals and monolinguals showed left hemisphere language dominance, regardless of proficiency level, whereas early bilinguals demonstrated bilateral involvement for language. Hull and Vaid suggested that increased involvement of the right hemisphere in early bilinguals could be a consequence of a relatively early need to recruit right hemisphere pragmatic strategies, such as to facilitate understanding of when and with whom to use one language versus the other (Obler, 1981; Beeman and Chiarello, 1998; Boatman, 2004).

A second meta-analysis by Hull and Vaid (2007) focused specifically on disentangling laterality differences among early and late bilinguals for both their languages. The outcomes replicated Hull and Vaid (2007) by demonstrating reliable bilateral organization in early bilinguals and left-lateralization in late bilinguals for their first languages. In addition, lateralization within a particular bilingual subgroup overlapped across first and second languages; that is, both the L1 and L2 of early bilinguals were organized bilaterally, and both the L1 and L2 of late bilinguals were left hemisphere lateralized. Taken together, outcomes from these meta-analyses point to differences in brain organization associated with differences in the number of languages experienced during early development.

Based on these data, Hull and Vaid (2007) posited the *anchoring hypothesis*, which argues that early exposure to two (or more) languages necessitates recruitment of neural support bilaterally, whereas single language exposure requires only left hemisphere priority. Moreover, the early establishment of the functional language pattern “anchors” later-learned languages so that they display that same pattern (presumably because they rely on the same neural substrates that were specialized for early learned languages). As such, the anchoring hypothesis concerning bilingual language organization is consistent with experience-based accounts of perceptual tuning [e.g., the experientially mediated (Trainor, 2005) and neural commitment (Kuhl et al., 2005b; Kuhl and Rivera-Gaxiola, 2008) hypotheses].

Several PET and fMRI studies have demonstrated converging evidence for overlapping neural substrates supporting L1 and L2 in bilinguals, both at the single word (Chee et al., 1999; Illes et al., 1999; Klein et al., 1999; Hernandez et al., 2000) and continuous speech levels (Perani et al., 1996; Chee et al., 1999; Vingerhoets et al., 2003; Briellmann et al., 2004). Because the lateralization of languages has been shown to differ for early and late bilinguals, but the patterns of L1 and L2 processing are nonetheless overlapping within each bilingual subtype, these outcomes are consistent with the notion that the functional specificity of the neural bases of language remains static once they are established. Presumably, such functional specificity is set up during the sensitive (or critical) period for language tuning.

At present, the bulk of existing evidence on language organization in bilingual infants and adults suggests that the end result of language tuning not only affects the neural organization of languages acquired during tuning but also any that may be acquired later in life. Thus, it seems clear that language tuning may be different for bilinguals and monolinguals, and this warrants a revisiting of the theories that characterize the neural mechanisms thought to underlie plasticity and neural specificity for language function.

The *experientially mediated hypothesis* (Trainor, 2005) and the *neural commitment hypothesis* (Kuhl et al., 2005b; Kuhl and Rivera-Gaxiola, 2008) could explain differences in the timing of language tuning, as some behavioral evidence with bilingual infants suggests a delay in the completion of language tuning, presumably to accommodate additional language organization to ensure the integrity of processing for multiple language systems. This position would be consistent with behavioral and neurophysiological evidence for reliable differences in adult bilingual laterality depending on age of L2 acquisition. However, if bilingual infants complete tuning for both languages at the same pace as their monolingual peers complete tuning for a single language (as some behavioral studies have suggested), the diversity of bilingual input may not complicate organization after all. Although it remains for future studies with bilingual infants to resolve this debate, it is clearly difficult to reconcile the latter position with existing adult bilingual evidence or with the experientially mediated and neural commitment accounts of language tuning.

The *self-termination of learning* account would also suggest that different neural connections are being forged for bilingual compared to monolingual infants because of the differences in their language experiences. This theory would account for the differential language laterality observed in adult early bilinguals compared with monolingual late bilinguals. In addition, this explanation could account for differences in connection formation as a function of differences in age of L2 acquisition. However, the self-termination of learning could also be consistent with the idea of a similar tuning timeline for bilingual and monolingual populations, in which case the generation of neural connections for language tuning would be similar for single and multiple languages (although this would be at odds with the adult neurophysiological data). Therefore, in the context of the evidence presented here, a sensitive period account of language tuning appears to better explain the current data in the infant and adult bilingual literatures, with the caveat that a reexamination of the sensitive period explanation would be warranted if future research demonstrates that bilinguals and monolinguals have comparably timed schedules of language tuning.

SENSORY DEPRIVATION (DEAFNESS) IN INFANTS AND CHILDREN

While the infant bilingual literature can be useful in addressing how multiple language inputs may affect language tuning and its neural bases, the language profile of deaf children who have received cochlear implants (CI) may offer useful information about how the *absence* of early auditory language input impacts language tuning in the developing brain. That is, congenital

deafness is a form of early sensory deprivation that can be later reversed by cochlear implantation surgery. What cannot be reversed, however, is the loss of early exposure, and this varies depending on the age of implantation. Given that tuning to native language phonemes in hearing infants requires auditory input and appears to develop during a period of high brain plasticity during infancy, the process of language tuning in children with CIs provides a unique opportunity to examine the boundary conditions for auditory plasticity and what, if any, language tuning takes place as a result.

Cochlear implants allow the recipient to experience auditory language for the first time, and this new sensory input ultimately results in neural reorganization to accommodate the perception of sound. Specifically, nearby brain areas are recruited to accommodate the new function of hearing (e.g., Lee et al., 2001; see also Desmond and Fiez, 1998; Recanzone, 2000, for related evidence from animal models). Importantly, CI research has demonstrated that cortical reorganization of the auditory cortex permits this region to adapt to the relatively limited frequency range produced by the CI, allowing for successful perception of even highly complex auditory stimuli (Shepherd et al., 1997). However, we are currently unaware whether the cortical reorganization during childhood cochlear implantation mimics that of infant language tuning.

The current literature focused on child CI users has identified several factors that influence the development of speech and language skills in a deaf child who is hearing through an implant. Unsurprisingly, evidence is accumulating that it is critical for the implantation to take place as early in development as possible (i.e., before 4 years of age). This allows the child to capitalize on the greater brain plasticity of that developmental period, thereby setting up the best conditions for normal auditory neural networks to emerge (Lee et al., 2001, 2007; Kang et al., 2004). Lee et al. (2001) used PET to show that by 7 years of age, cortical organization was significantly different from that of younger children, and this finding was corroborated by a converging paradigm (cortical auditory evoked potentials; Sharma et al., 2002). Later, Lee et al. (2007) and Kang et al. (2004) further narrowed the optimal window by establishing that post-implantation speech perception scores improved when implantation occurred before 4 years of age. More specifically, children who received a CI before the age of four performed best on sentence recognition tasks, whereas those implanted between 4 and 7 years of age showed a wide range of performance on this task; after age seven, implant recipients generally achieved low scores (Sharma et al., 2009).

However, as CI surgeries are performed on younger and younger patients we will be able to comment on how different durations of sensory deprivation affect language learning, and tuning. At least one recent behavioral study indicated that when testing infants with 1–2 months of initial CI use, infants that ranged between 4 and 10 months of age preferred their native language (Hebrew) compared with English (Kishon-Rabin et al., 2010). This timeline mirrors that seen for normal-hearing infants and corroborates the view that, with very early intervention, tuning may not be impacted (at least at the behavioral level).

Although the data from behavioral, PET, and EEG studies are helpful in providing evidence that the target period for normal language development via a CI following auditory deprivation ends by 4 years of age (Gordon et al., 2005; Dorman et al., 2007), it is generally quite difficult to obtain neurophysiological measurements from very young children using these techniques. Moreover, fMRI cannot be used in CI patients because the devices are incompatible with the scanner's magnetic field. However, the NIRS technique allows measurement of blood-oxygen level dependent changes in cortical activity, similar to fMRI, without interference with the implant. Indeed, NIRS evidence regarding language development has already been used to study the bilateral temporal responses of pediatric CI recipients to language.

Sevy et al. (2010) tested CI children (mean age 7.8 years, $n = 7$) on the same day they received their implants, providing a unique opportunity to observe how the untrained auditory cortex responds to its first exposure to sound. The results revealed significant hemodynamic responses to speech stimuli in the majority of deaf children, and the volume of hemodynamic response was not significantly different from that of hearing controls (mean age 10.2 years, $n = 9$). However, the majority of deaf children with a new CI showed a unilateral hemodynamic response to the auditory stimuli, and that response was most often ipsilateral to the location of the CI (typically in the right hemisphere). Once the children had at least 6 months of experience with their CI, responses to the auditory stimuli were predominantly bilateral and the unilateral responses that remained were most likely to be contralateral to the CI. These findings provide an initial glimpse into the boundaries for plasticity in congenitally deaf children whose deafness is subsequently ameliorated with a CI.

While this is the first study to use NIRS with this population, the Sevy et al. (2010) results demonstrate that NIRS can safely measure cortical responses in pediatric implant users, thereby establishing NIRS as a valuable tool for investigating the neural bases of language tuning and development in this special population. This application of NIRS promises to expand our understanding of the boundary conditions for development of normal auditory processing and of the likely neural mechanisms involved in perceptual tuning in general, and language tuning in particular. Thus, NIRS and other converging neurophysiological paradigms may further elucidate how sensory deprivation impacts auditory brain development in children (Neville and Bavelier, 2002; Kang et al., 2004; Giley et al., 2008).

CONCLUSION AND SYNTHESIS

Traditionally, researchers have discussed language tuning in terms of a critical, or sensitive, period (Werker and Tees, 2005). These terms identify language tuning as a phenomenon that is either constrained by biological maturation of systems important to language learning or as a period that capitalizes on the responsiveness of neural plasticity to language experience. In the context of language tuning, the present review highlights the importance of the differential role of experience (i.e., a child's language profile) in actively shaping the biological foundation for language best characterize the extant data. Three theories that characterize this aspect of the data particularly well are the *self-termination of*

learning account (Thomas and Johnson, 2008), the *experientially mediated account* (Trainor, 2005), and the *neural commitment hypothesis* (Kuhl et al., 2005a).

Although these theoretical constructs offer valuable explanations of the processes that drive perceptual tuning in general, they are somewhat limited in terms of their ability to explain the range of behavioral and neurophysiological data on language tuning. Based on the synthesis of data presented here, it appears that language tuning may be comprised of staggered, nested components that may tune at different times (Werker and Tees, 2005). Thus, a more specific means of reconciling these data can be encompassed with a supplementary explanation, namely, the *relative speed of tuning hypothesis*. Specifically, the relative speed of tuning hypothesis predicts that, within the maturational window that clearly exists for native-like language learning, the more transparent, obvious elements of language may tune faster than more opaque, ambiguous elements of language. We anticipate that this hypothesis will be especially useful in accounting for conflicting data within the literature, as it allows researchers to break the tuning process down based on the specific factors that influence tuning rate as measured using behavioral and neurophysiological paradigms. In particular, this approach has the potential to account for the impact of the *in utero* language profile, the characteristics of both heterogeneous and homogenous language profiles during infancy and childhood, and the roles of age and sensory deprivation on language tuning. Our hypothesis gives rise to several predictions, some of which are supported by the current data and others that will require further investigation.

First, the relative speed of tuning hypothesis predicts that behavioral testing with transparent, relatively easy tasks should demonstrate evidence of tuning in younger populations than those with more complicated, opaque tasks. For example, use of a categorization paradigm may demonstrate behavioral evidence of tuning earlier than a discrimination paradigm, due to the relative difficulties of these tasks (Bosch and Sebastian-Galles, 2003; Albareda-Castellot et al., 2011). And this should also be the case for neurophysiological paradigms, where evidence of language tuning should be observed earlier given tasks that tap into elements of language that tune early themselves. For example, an experimental design contrasting native and non-native languages of a different rhythmic class should demonstrate neural indicators of language tuning earlier than a task contrasting native and non-native languages within the same rhythmic class.

A further prediction is that “robust,” or cue-filled, sentence-level speech stimuli will elicit indicators of language tuning sooner than, for example, filtered speech (in which phonetic cues are removed) or contrast-level stimuli (in which only a fraction of the original speech stimulus is retained). This is borne out by recent neurophysiological data, in which infants showed initial neural evidence of tuning to word-level contrasts at 11 months (Minagawa-Kawai et al., 2007), while showing evidence of tuning to sentence-level stimuli at 4 months (Minagawa-Kawai et al., 2011a). Thus, evidence of language tuning – that is, a more left-lateralized hemodynamic response for native compared to non-native speech – was elicited in a younger group of infants given more robust stimuli. Our hypothesis also accounts for data from studies comparing infants’ ability to discriminate phonemes from more or less populated phonemic

space in which discrimination is better for phonemes from sparsely populated phonemic space relative to those from dense phonemic space (Burns et al., 2007; Sumner et al., 2008).

Some cues may be more robust than others given different language profiles. For example, ambiguity due to changes in speaker appears to be less influential to language tuning in bilingually exposed infant than do the relative frequency of phonetic contrasts across the ambient languages (Sebastian-Galles and Bosch, 2009). Since the relative speed of tuning hypothesis predicts that linguistic elements that are more robust and transparent will elicit earlier tuning than will weaker or more complex elements, different predictions emerge for bilingual relative to monolingual profiles. For example, where May et al. (under review) utilized speech stimuli that were essentially only prosodic in nature (i.e., low-pass filtered) and failed to obtain neural indicators of discrimination between native and non-native forms in monolingual neonates, bilingual neonates were able to distinguish between these same stimuli in a behavioral task (Byers-Heinlein et al., 2010). The relative speed of tuning hypothesis would predict such a difference given the different language profiles of the two groups of infants. Although unavailable from these studies, one can imagine that the full complement of data (e.g., monolingual and bilingual infants tested with the same stimuli using both behavioral and neurophysiological measures) would further support this outcome. The influence of these and other factors require further investigation to determine their relative impact on rate of language tuning.

Finally, the relative speed of tuning hypothesis can also account for data from populations deprived of sensory information, such as cochlear implant users. For these individuals, some of whom receive auditory input only after their system has undergone language tuning without it, we predict that tuning would occur more quickly for more robust aspects of the auditory signal that can be integrated with the greatest ease (e.g., phonemes that are clearly mapped to the visual signal, such as bilabial stops, or prosodic contours that easily map to, for example, vocal aperture or facial prosody). Again, such research will further inform the field.

In sum, given different language profiles, the relative speed of language tuning hypothesis provides a flexible way of framing the complex and sometimes contradictory behavioral and neurophysiological literature. Of course, this hypothesis is only one of several trying to account for the diverse data on language tuning. This review highlights the need for converging evidence from a variety of experimental designs, linguistic stimuli, and imaging modalities and neurophysiological methods to demonstrate reliable evidence of language tuning and to resolve some of the current inconsistencies across studies and language profiles. It is possible that evidence (both neural and behavioral) of language tuning will emerge from increased use of ecologically valid, sentence-level stimuli (as opposed to single syllable and pseudoword stimuli). Because the majority of neurophysiological data on early language development have been obtained from monolingually exposed infants, it will also be important for future studies to investigate the influences of different language profiles on the neural mechanisms that support language tuning. Nonetheless, the present review makes clear that the extant literature offers an important guide for future exploration of the basis of the language tuning process.

REFERENCES

- Albareda-Castellot, B., Pons, F., and Sebastián-Gallés, N. (2011). The acquisition of phonetic categories in bilingual infants: new data from an anticipatory eye movement paradigm. *Dev. Sci.* 14, 395–401.
- Albin, D. D., and Echols, C. H. (1996). Stressed and word-final syllables in infant-directed speech. *Infant Behav. Dev.* 19, 401–418.
- Armstrong, V. L., Brunet, P. M., He, C., Nishimura, M., Poole, H. L., and Spector, F. J. (2006). What is so critical?: a commentary on the reexamination of critical periods. *Dev. Psychobiol.* 48, 326–331.
- Aslin, R. N., and Pisoni, D. B. (1980). "Some developmental processes in speech perception," in *Child Phonology*, eds. G. H. Yeni-Komshian, J. F. Kavanagh, and C. A. Ferguson (New York, NY: Academic Press), 67–96.
- Bahrick, L. E., and Lickliter, R. (2000). Intersensory redundancy guides attentional selectivity and perceptual learning in infancy. *Dev. Psychol.* 36, 190–201.
- Bahrick, L. E., and Lickliter, R. (2002). "Intersensory redundancy guides early perceptual and cognitive development," in *Advances in Child Development and Behavior*, ed. R. Kail (New York, NY: Academic Press), 153–187.
- Bahrick, L. E., Lickliter, R., Castellanos, I., and Vaillant-Molina, M. (2010). Increasing task difficulty enhances effects of intersensory redundancy: testing a new prediction of the Intersensory Redundancy Hypothesis. *Dev. Sci.* 13, 731–737.
- Beeman, M., and Chiarello, C. (1998). Complementary left and right hemisphere language comprehension. *Curr. Dir. Psychol. Sci.* 7, 2–8.
- Best, C. (1994). "The emergence of native-language phonological influences in infants: a perceptual assimilation model," in *The Development of Speech Perception: The Transition from Speech Sounds to Spoken Words*, eds J. C. Goodman and H. C. Nusbaum (Cambridge, MA: The MIT Press), 167–224.
- Best, C., McRoberts, G. W., and Sithole, N. M. (1988). Examination of perceptual reorganization for nonnative speech contrasts: Zulu click discrimination by English-speaking adults and infants. *J. Exp. Psychol. Hum. Percept. Perform.* 14, 345–360.
- Boatman, D. (2004). Cortical bases of speech perception: evidence from functional lesion studies. *Cognition* 92, 47–65.
- Bortfeld, H., Fava, E., and Boas, D. A. (2009). Identifying cortical lateralization of speech processing in infants using near-infrared spectroscopy. *Dev. Neuropsychol.* 34, 52–65.
- Bortfeld, H., Wruck, E., and Boas, D. A. (2007). Assessing infants' cortical response to speech using near-infrared spectroscopy. *NeuroImage* 34, 407–415.
- Bosch, L., and Sebastián-Gallés, N. (1997). Native-language recognition abilities in 4-month-old infants from monolingual and bilingual environments. *Cognition* 65, 33–69.
- Bosch, L., and Sebastián-Gallés, N. (2003). Simultaneous bilingualism and the perception of a language specific vowel contrast in the first year of life. *Lang. Speech* 46, 217–244.
- Briellmann, R. S., Saling, M. M., Connell, A. B., Waite, A. B., Abbott, D. F., and Jackson, G. D. (2004). A high-field functional MRI study of quadrilingual subjects. *Brain Lang.* 89, 531–542.
- Burns, T. C., Yoshida, K. A., Hill, K., and Werker, J. F. (2007). The development of phonetic representation in bilingual and monolingual infants. *Appl. Psycholinguist.* 28, 455–474.
- Byers-Heinlein, K., Burns, T. C., and Werker, J. F. (2010). The roots of bilingualism in newborns. *Psychol. Sci.* 21, 343–348.
- Chee, M. W. L., Tan, E. W. L., and Thiel, T. (1999). Mandarin and English single word processing studied with functional magnetic resonance imaging. *J. Neurosci.* 19, 3050–3056.
- Conboy, B. T., and Mills, D. L. (2006). Two languages, one developing brain: effects of vocabulary size on bilingual toddlers' event-related potentials to auditory words. *Dev. Sci.* 9, F1–F11.
- Curtin, S., Byers-Heinlein, K., and Werker, J. F. (2011). Bilingual beginnings as a lens for theory development: PRIMIR in focus. *J. Phon.* doi: 10.1016/j.wocn.2010.12.002. [Epub ahead of print].
- Decasper, A. J., and Fifer, W. P. (1980). Of human bonding: newborns prefer their mothers' voices. *Nature* 208, 1174–1176.
- Decasper, A. J., Lecanuet, J. P., Busnel, M. C., Granier-Deferre, C., and Maugeais, R. (1994). Fetal reactions to recurrent maternal speech. *Infant Behav. Dev.* 17, 159–164.
- Dehaene-Lambertz, G., Dehaene, S., and Hertz-Pannier, L. (2002). Functional neuroimaging of speech perception in infants. *Science* 298, 2013–2015.
- Desmond, J. E., and Fiez, J. A. (1998). Neuroimaging studies of the cerebellum: language, learning and memory. *Trends. Cogn. Sci.* 2, 355–362.
- Dorman, M. F., Gifford, R. H., Spahr, A. J., and Mckarns, S. A. (2007). The benefits of combining acoustic and electric stimulation for the recognition of speech, voice and melodies. *Audiol. Neurotol.* 13, 105–112.
- Eimas, P. D., Siqueland, E. R., Jusczyk, P. W., and Vigorito, J. (1971). Speech perception in infants. *Science* 171, 303–306.
- Fennell, C. T., Byers-Heinlein, K., and Werker, J. F. (2007). Using speech sounds to guide word learning: the case of bilingual infants. *Child Dev.* 78, 1510–1525.
- Friederici, A., Friederich, M., and Christophe, A. (2007). Brain responses in 4-month-old infants are already language specific. *Curr. Biol.* 17, 1208–1211.
- Garnicka, O. K. (1977). "Some prosodic and paralinguistic features of speech to young children," in *Talking to Children: Language Input and Acquisition*, eds C. E. Snow and C. A. Ferguson (Cambridge: Cambridge University Press), 63–88.
- Gervain, J., Macagno, E., Cogoi, S., Pena, M., and Mehler, J. (2008). The neonate brain detects speech structure. *Proc. Natl. Acad. Sci. U.S.A.* 105, 14222–14227.
- Giley, P. M., Sharma, A., and Dorman, M. F. (2008). Cortical organization in children with cochlear implants. *Brain Res.* 1239, 56–65.
- Gordon, K. A., Tanaka, S., and Papsin, B. C. (2005). Atypical cortical responses underlie poor speech perception in children using cochlear implants. *NeuroReport* 16, 2041–2046.
- Hepper, P., and Shabidullah, B. (1994). Development of fetal hearing. *Arch. Dis. Child Fetal Neonatal Ed.* 71, F81–F87.
- Hernandez, A. E., Martinez, A., and Kohnert, K. (2000). In search of the language switch: an fMRI study of picture naming in Spanish-English bilinguals. *Brain Lang.* 73, 421–431.
- Hull, R., and Vaid, J. (2006). Laterality and language experience. *Laterality* 11, 436–464.
- Hull, R., and Vaid, J. (2007). Bilingual language lateralization: a meta-analytic tale of two hemispheres. *Neuropsychologia* 45, 1987–2008.
- Illes, J., Francis, W. S., Desmond, J. E., Gabrieli, J. D. E., Glover, G. H., Poldrack, R., Lee, C. J., and Wagner, A. D. (1999). Convergent cortical representation of semantic processing in bilinguals. *Brain Lang.* 70, 347–363.
- Indefrey, P. (2006). "A meta-analysis of hemodynamic studies on first and second language processing: which suggested differences can we trust and what do they mean?" in *The Cognitive Neuroscience of Second Language Acquisition*, eds M. Gullberg and P. Indefrey (Malden, MA: Blackwell), 279–304.
- Jardri, R., Pins, D., Houfflin-Debarge, V., Chaffiotte, C., Rocourt, N., Pruvo, J. P., Steinling, M., Delion, P., and Thomas, P. (2008). Fetal cortical activation to sound at 33 weeks of gestation: a functional MRI study. *NeuroImage* 42, 10–18.
- Johnson, M. H. (2001). Functional brain development in humans. *Nat. Rev. Neurosci.* 2, 475–483.
- Johnson, M. H. (2005). Sensitive periods in functional brain development: problems and prospects. *Dev. Psychobiol.* 46, 287–292.
- Jusczyk, P. W. (1997). *The Discovery of Spoken Language*. Cambridge, MA: MIT Press.
- Jusczyk, P. W., Rosner, B. S., Cutting, J. E., Foard, C. F., and Smith, L. B. (1977). Categorical perception of nonspeech sounds by 2-month-old infants. *Percept. Psychophys.* 21, 50–54.
- Kang, E., Lee, D. S., Kang, H., Lee, J. S., Oh, S., Lee, M. C., and Kim, C. S. (2004). Neural changes associated with speech learning in deaf children following cochlear implantation. *NeuroImage* 22, 1173–1181.
- Kishon-Rabin, L., Harel, M., Hildesheimer, M., and Segal, O. (2010). Listening preference for the native language compared to an unfamiliar language in hearing and hearing-impaired infants after cochlear implantation. *Otol. Neurotol.* 31, 1275–1280.
- Kisilevsky, B. S., Hains, S. M. J., Jacquet, A. Y., Grainer-Deferre, C., and Lecanuet, J. P. (2004). Maturation of fetal responses to music. *Dev. Sci.* 7, 550–559.
- Klein, D., Milner, B., Zatorre, R., Zhao, V., and Nikelski, J. (1999). Cerebral organization in bilinguals: a PET study of Chinese-English verb generation. *NeuroReport* 10, 4841–4846.
- Knudsen, E. I. (1999). Mechanisms of experience-dependent plasticity in the auditory localization pathway of the barn owl. *J. Comput. Physiol. A* 185, 305–321.
- Kotilahti, K., Nissila, I., Nissila, T., Lipiainen, L., Noponen, T., Merilainen, P., Huottilainen, M., and Fellman, V. (2010). Hemodynamic responses to speech and music in newborn infants. *Hum. Brain Mapp.* 31, 595–603.
- Kuhl, P. K., Conboy, B. T., Coffey-Corina, S., Padden, D., and Nelson, T. (2005a). Early speech perception and later language development: implications for the "critical period." *Lang. Learn. Dev.* 1, 237–264.
- Kuhl, P. K., Conboy, B. T., Padden, D., Nelson, T., and Pruitt, J. (2005b). Early speech perception and later language development: implications for the

- "critical period." *Lang. Learn. Dev.* 1, 237–264.
- Kuhl, P. K., and Rivera-Gaxiola, M. (2008). Neural substrates of language acquisition. *Annu. Rev. Neurosci.* 31, 511–534.
- Kuhl, P. K., Stevens, E. B., Hayashi, A., Deguchi, T., Kiritani, S., and Iverson, P. (2006). Infants show a facilitation effect for native language phonetic perception between 6 and 12 months. *Dev. Sci.* 9, F13–F21.
- Kuhl, P. K., Tsao, F. M., Liu, H. M., Zhang, Y., and Deboer, B. (2001). Language/culture/mind/brain: progress at the margins between disciplines. *Annu. N. Y. Acad. Sci.* 935, 136–174.
- Kuhl, P. K., Williams, K. A., Lacerda, F., Stevens, K. N., and Lindblom, B. (1992). Linguistic experience alters phonetic perception in infants by 6 months of age. *Science* 255, 606–608.
- Lecanuet, J. P., Fifer, W. P., Krasnegor, N., and Smotherman, W. P. (1995). *Fetal development: a psychobiological perspective*. Hillsdale, NJ: Lawrence Erlbaum Association.
- Lee, H. J., Truay, E., Mamou, G., Sappey-Marini, D., and Giraud, A.-L. (2007). Visual speech circuits in profound acquired deafness: a possible role for latent multimodal connectivity. *Brain* 130, 2929–2941.
- Lee, S. D., Lee, J. S., Oh, S. H., Kim, S. K., Kim, J. W., Chung, J. K., Lee, M. C., and Kim, C. S. (2001). Deafness: cross-modal plasticity and cochlear implants. *Nature* 409, 149–150.
- Lloyd-Fox, S., Blasi, A., and Elwell, C. E. (2010). Illuminating the developing brain: the past, present and future of functional near infrared spectroscopy. *Neurosci. Behav. Rev.* 34, 269–284.
- Mattock, K., Polka, L., Rvachew, S., and Krehm, M. (2010). The first steps in word learning are easier when the shoes fit: comparing monolingual and bilingual infants. *Dev. Sci.* 13, 229–243.
- McMurray, B., and Aslin, R. N. (2004). Anticipatory eye movements reveal infants' auditory and visual categories. *Infancy* 6, 203–229.
- Mechelli, A., Crinion, J. T., Noppeney, U., O'doherty, J., Ashburner, J., Frackowiak, R. S., and Price, C. J. (2004). Neurolinguistics: structural plasticity in the bilingual brain. *Nature* 431, 757.
- Mehler, J., Bertoncini, J., and Barriere, M. (1978). Infant recognition of mother's voice. *Perception* 7, 491–497.
- Mehler, J., Jusczyk, P. W., Lambertz, G., Halsted, G., Bertoncini, J., and Amiel-Tison, C. (1988). A precursor of language acquisition in young infants. *Cognition* 29, 143–178.
- Minagawa-Kawai, Y., Van Der Lely, H., Ramus, F., Sato, Y., Mazuka, R., and Dupoux, E. (2011a). Optical brain imaging reveals general auditory and language-specific processing in early infant development. *Cereb. Cortex* 21, 254–261.
- Minagawa-Kawai, Y., Cristia, A., and Dupoux, E. (2011b). Cerebral lateralization and early speech acquisition: a developmental scenario. *Dev. Cogn. Neurosci.* 1, 217–232.
- Minagawa-Kawai, Y., Mori, K., Naoi, N., and Kojima, S. (2007). Neural attunement processes in infants during the acquisition of a language-specific phonemic contrast. *J. Neurosci.* 27, 315–321.
- Moon, C., Panneton-Cooper, R., and Fifer, W. P. (1993). Two-day olds prefer their native language. *Infant Behav. Dev.* 16, 495–500.
- Morokuma, S., Fukushima, K., Kawai, N., Tomonaga, M., Satoh, S., and Nakano, H. (2004). Fetal habituation correlates with functional brain development. *Behav. Brain Res.* 153, 459–463.
- Narayan, C. R., Werker, J. F., and Beddor, P. S. (2010). The interaction between acoustic salience and language experience in developmental speech perception: evidence from nasal place discrimination. *Dev. Sci.* 13, 407–420.
- Nazzi, T., Bertoncini, J., and Mehler, J. (1998). Language discrimination by newborns: toward an understanding of the role of rhythm. *J. Exp. Psychol. Hum. Percept. Perform.* 24, 756–766.
- Neville, H., and Bavelier, D. (2002). Human brain plasticity: evidence from sensory deprivation and altered language experience. *Prog. Brain Res.* 138, 177–188.
- Neville, H. J., Bavelier, D., Corina, D., Rauschecker, J., Karni, A., Lalwani, A., Braun, A., Clark, V., Jezzard, P., and Turner, R. (1998). "Cerebral organization for language in deaf and hearing subjects: biological constraints and effects of experience," in *Proceedings of the National Academy of Sciences* (Washington, DC: The National Academy of Sciences), 95, 922–929.
- Obler, L. K. (1981). "Right hemisphere participation in second language acquisition," in *Individual Differences and Universals in Language Learning Aptitude*, ed. K. Diller (Rowley, MA: Newbury Press), 53–64.
- Peña, M., Maki, A., Kovacic, D., Dehaene-Lambertz, G., Koizumi, H., Bouquet, F., and Mehler, J. (2003). Sounds and silence: an optical topography study of language recognition at birth. *Proc. Natl. Acad. Sci. U.S.A.* 100, 11702–11705.
- Perani, D., Dehaene, S., Grassi, F., Cohen, L., Cappa, S. F., Dupoux, D., Fazio, F., and Mehler, J. (1996). Brain processing of native and foreign languages. *NeuroReport* 7, 2439–2444.
- Polka, L., Colantoni, C., and Sundara, M. (2001). A cross-language comparison of /d/-/th/ perception: evidence for a new developmental pattern. *J. Acoust. Soc. Am.* 109, 2190–2201.
- Polka, L., and Werker, J. F. (1994). Developmental changes in perception of non-native vowel contrasts. *J. Exp. Psychol. Hum. Percept. Perform.* 20, 421–435.
- Quinn, P. C. (2002). "Early categorization: a new synthesis," in *Blackwell Handbook of Childhood Cognitive Development*, ed. U. Goswami (Malden, MA: Blackwell), 84–101.
- Radicevic, Z., Vujovic, M., Jelacic, L., and Sovilj, M. (2008). Comparative findings of voice and speech: language processing at an early ontogenetic age in quantitative EEG mapping. *Exp. Brain Res.* 184, 529–532.
- Ramus, F., Hauser, M. D., Miller, C., Morris, D., and Mehler, J. (2000). Language discrimination by human newborns and by cotton-top tamarin monkeys. *Science* 288, 349–351.
- Recanzone, G. H. (2000). "Cerebral cortical plasticity: perception and skill acquisition," in *The New Cognitive Neurosciences*, 2nd ed., ed. M. S. Gazzaniga, (Cambridge, MA: MIT Press), 237–247.
- Rivera-Gaxiola, M., Silva-Pereyra, J., and Kuhl, P. K. (2005). Brain potentials to native and non-native contrast in 7- and 11-month-old american infants. *Dev. Sci.* 8, 162–172.
- Sabourin, L., Werker, J. F., Bosch, L., and Sebastián-Gallés, N. (2003). Perceiving vowels in a tight vowel space: evidence from monolingual infants. *J. Cogn. Dev.* Available at: <http://aix1.uottawa.ca/~lsabour3/publicat.htm>
- Saito, Y., Aoyama, S., Kondo, T., Fukumoto, R., Konishi, N., Nakamura, K., Kobayashi, M., and Toshima, T. (2007). Frontal cerebral blood flow change associated with infant-directed speech. *Arch. Dis. Child Fetal Neonatal Ed.* 92, F113–F116.
- Sebastian-Galles, N. (2002). Comment on Werker and Tees (1984). Cross-language speech perception: evidence for perceptual reorganisation during the first year of life. *Infant Behav. Dev.* 25, 144–146.
- Sebastian-Galles, N., and Bosch, L. (2009). Developmental shift in the discrimination of vowel contrasts in bilingual infants: is the distributional account all there is to it? *Dev. Sci.* 12, 874–887.
- Sevy, A., Bortfeld, H., Huppert, T., Beauchamp, M. S., Tonini, R., and Oghalai, J. (2010). Neuroimaging with near-infrared spectroscopy demonstrates speech-evoked activity in the auditory cortex of deaf children following cochlear implantation. *Hear. Res.* 270, 39–47.
- Sharma, A., Dorman, N. F., and Spahr, A. J. (2002). A sensitive period for the development of the central auditory system in children with cochlear implants: implication for age of implantation. *Ear Hear.* 23, 532–539.
- Sharma, A., Nash, A. A., and Dorman, M. (2009). Cortical development, plasticity and re-organization in children with cochlear implants. *J. Commun. Disord.* 42, 272–279.
- Shepherd, R. K., Hartmann, R., Heid, S., Hardie, N., and Klinke, R. (1997). The central auditory system and auditory deprivation: experience with cochlear implants in the congenitally deaf. *Acta Otolaryngol.* 117, 28–33.
- Stern, D. N., Spieker, S., Barnett, R. K., and Mackain, K. (1983). The prosody of maternal speech: infant age and context related changes. *J. Child Lang.* 10, 1–15.
- Stiles, J. (2000). Neural plasticity and cognitive development. *Dev. Neuropsychol.* 18, 237–272.
- Sumner, P., Anderson, E. J., Sylvester, R., Haynes, J. D., and Rees, G. (2008). Combined orientation and colour information in human V1 for both L-M and S-cone chromatic axes. *NeuroImage* 39, 814–824.
- Sundara, M., Polka, L., and Genesee, F. (2006). Language-experience facilitates discrimination of /d/-in monolingual and bilingual acquisition of English. *Cognition* 100, 369–388.
- Sundara, M., Polka, L., and Molnar, M. (2008). Development of coronal stop perception: bilingual infants keep pace with their monolingual peers. *Cognition* 108, 232–242.
- Taga, G., and Asakawa, K. (2007). Selectivity and localization of cortical response to auditory and visual stimulation in awake infants aged 2 to 4 months. *NeuroImage* 36, 1246–1252.
- Taga, G., Homae, F., and Watanabe, H. (2007). Effects of source-detector distance of near infrared spectroscopy on the measurement of the cortical hemodynamic response in infants. *NeuroImage* 38, 452–460.
- Thomas, M. S. C., and Johnson, M. H. (2008). New advances in understanding sensitive periods in brain development. *Curr. Dir. Psychol. Sci.* 17, 1–5.
- Trainor, L. J. (2005). Are there critical periods for musical development? *Dev. Psychol.* 46, 262–278.
- Trehub, S. E. (1976). The discrimination of foreign speech contrasts by infants and adults. *Child Dev.* 47, 466–472.
- Tsao, F. M., Liu, H. M., and Kuhl, P. K. (2004). Speech perception in infancy predicts language development in the second year of life: a longitudinal study. *Child Dev.* 75, 1067–1084.

- Vihman, M. M., Thierry, G., Lum, J. A., Keren-Portnoy, T., and Martin, P. (2007). Onset of word form recognition in English, Welsh, and English–Welsh bilingual infants. *Appl. Psycholinguist.* 28, 475–493.
- Vingerhoets, G., Van Borsel, J., Tesink, C., Van Den Noort, M., Deblaere, K., Seurinck, R., Vandemaele, P., and Achten, E. (2003). Multilingualism: an fMRI study. *NeuroImage* 20, 2181–2196.
- Vouloumanos, A., and Werker, J. F. (2007). Listening to language at birth: evidence for a bias for speech in neonates. *Dev. Sci.* 10, 159–164.
- Wartenbuger, I., Heekeren, H. R., Abutalebi, J., Cappa, S. F., Villringer, A., and Perani, D. (2003). Early setting of grammatical processing in the bilingual brain. *Neuron* 37, 159–170.
- Webb, S. J., Monk, C. S., and Nelson, C. A. (2001). Mechanisms of postnatal neurobiological development: implications for human development. *Dev. Neuropsychol.* 19, 147–171.
- Werker, J. F., and Byers-Heinlein, K. (2008). Bilingualism in infancy: first steps in perception and comprehension. *Trends Cogn. Sci.* 12, 144–151.
- Werker, J. F., and Lalonde, C. E. (1988). Cross-language speech perception: initial capabilities and developmental change. *Dev. Psychobiol.* 24, 672–683.
- Werker, J. F., and Tees, R. C. (1983). Developmental changes across childhood in the perception of non-native speech sounds. *Can. J. Psychol.* 37, 278–286.
- Werker, J. F., and Tees, R. C. (1984). Cross-language speech perception: evidence for perceptual reorganization during the first year of life. *Infant Behav. Dev.* 7, 49–63.
- Werker, J. F., and Tees, R. C. (1999). “Experimental influences on infant speech processing: toward a new synthesis,” in *Annual Review of Psychology*, ed. J. T. Spence (Palo Alto, CA: Annual Reviews), 509–535.
- Werker, J. F., and Tees, R. C. (2005). Speech perception as a window for understanding plasticity and commitment in language systems of the brain. *Dev. Psychobiol.* 46, 233–251.
- Yoshida, K. A., Fennell, C. T., Swingle, D., and Werker, J. F. (2009). Fourteen-month-old infants learn similar-sounding words. *Dev. Sci.* 12, 412–418.
- Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 15 January 2011; accepted: 12 July 2011; published online: 01 August 2011.

Citation: Fava E, Hull R and Bortfeld H (2011) Linking behavioral and neurophysiological indicators of perceptual tuning to language. *Front. Psychology* 2:174. doi: 10.3389/fpsyg.2011.00174

This article was submitted to *Frontiers in Language Sciences*, a specialty of *Frontiers in Psychology*.

Copyright © 2011 Fava, Hull and Bortfeld. This is an open-access article subject to a non-exclusive license between the authors and Frontiers Media SA, which permits use, distribution and reproduction in other forums, provided the original authors and source are credited and other Frontiers conditions are complied with.



Studying neonates' language and memory capacities with functional near-infrared spectroscopy

Silvia Benavides-Varela¹, David M. Gómez^{1,2} and Jacques Mehler^{1*}

¹ Cognitive Neuroscience Sector, International School for Advanced Studies, Trieste, Italy

² Center for Advanced Research in Education, University of Chile, Santiago, Chile

Edited by:

Judit Gervain, Université Paris Descartes, France

Reviewed by:

Isabell Wartenburger, University of Potsdam, Germany

Ramon Mariano Guevara, Basque Center for Brain, Language and Cognition, Spain

*Correspondence:

Jacques Mehler, Cognitive Neuroscience Sector, International School for Advanced Studies, Via Bonomea 265, 34136 Trieste, Italy.
e-mail: mehler@sissa.it

The measurement of newborns' brain hemodynamic activity has improved our understanding of early cognitive processes, in particular of language acquisition. In this paper, we describe two experimental protocols adapted to study neonates' speech-processing capacities using functional near-infrared spectroscopy (fNIRS): the block design and the familiarization-recognition design. We review some of their benefits and disadvantages, and refer to research issues that can be explored by means of these protocols. We also illustrate the use of the two experimental designs through representative fNIRS studies that reveal specific patterns of activation of the newborn brain during speech perception, learning of repetition structures, and word recognition.

Keywords: cognitive development, language acquisition, newborns' memory, fNIRS

INTRODUCTION

Neonates represent an ideal population for the investigation of the biological dispositions that guide humans to acquire their native language. Neonate studies reveal how we begin processing speech, and inform us about the cognitive faculties that we possess before the accumulation of significant linguistic experience.

However, working with this population is a great challenge. Newborns stay awake only for short periods of time; and when they are awake they eat, interact with caretakers, cry, or are in a quiet state of drowsiness. These changes of state make it difficult to obtain behavioral measures useful for research. In addition, most of the current neuroimaging techniques are impractical for testing healthy young participants because these techniques a) use machines that produce high levels of acoustic noise, b) require the application of liquids or gel in the newborn's head, or c) have very low tolerance to facial and body movement. Functional near-infrared spectroscopy (fNIRS) is an imaging technique that has been employed for clinical purposes (Wolf et al., 2007) and more recently in cognitive research as well. It is regarded as one of the most appropriate to study language faculties and cognitive capacities in the newborn's brain (see Aslin and Mehler, 2005; Minagawa-Kawai et al., 2008; Lloyd-Fox et al., 2010; Obrig et al., 2010, for reviews on fNIRS developmental studies, but see also Dehaene-Lambertz and Peña, 2001; Kushnerenko et al., 2001; Cheour et al., 2002, 2004; Imada et al., 2006, for examples of language studies in newborns using magnetoencephalography and electroencephalography). fNIRS is non-invasive, and there is no need to use any substance, not even to keep the device in place on the infant's head. It is ideal to study how neonates process auditory stimuli because the device makes hardly any noise. Moreover, we have observed that the number of participants providing useful fNIRS data is very high compared to many behavioral methods used to investigate neonatal cognition: in typical behavioral paradigms used to study auditory capacities, rejection

rate may reach beyond 50%. In our fNIRS studies, the number of infants that are excluded from the analysis ranges between 20 and 25%. One reason for this is that, in fNIRS studies regarding auditory perception, participants need not be excluded because of behaviors like falling asleep¹ or refusing the pacifier, very common in neonates. However, infants must be relatively quiet because head movements can easily displace the fNIRS probes.

In what follows, we describe some studies carried out with newborns in our laboratory using fNIRS, with particular focus on the technical and methodological aspects.

EQUIPMENT AND EXPERIMENTAL SETUP

Functional near-infrared spectroscopy works by measuring changes in cerebral blood flow volume and oxygen saturation using optical means. It is based on the emission of near-infrared laser light on the subject's scalp. Lloyd-Fox et al. (2010) and Gervain et al. (2011) provide excellent reviews about the workings of fNIRS and its applications to infant research.

Our laboratory is located in the Hospital Azienda Ospedaliera Santa Maria della Misericordia in Udine, in the same building as the Obstetrics and Neonatology Department. We use the ETG-4000 machine (Hitachi Medical Corporation, Tokyo, Japan), which emits continuous near-infrared light at two wavelengths—695 nm and 830 nm—through optical fibers. The sampling rate is 10 Hz, and the total laser power output per fiber we use is 0.75 mW.

A silicon holder usually called “probe” keeps the optical fibers at a fixed distance from one another. In each probe, five fibers act as emitters and four as detectors, allowing simultaneous recording

¹For studies with auditory stimulation, it is not considered as an issue that infants sleep during the experimental session. Indeed, sleeping newborns can discriminate and learn speech sounds (Cheour et al., 2002), and sleeping 3-month-olds show brain responses associated with memory for previously heard sentences (Dehaene-Lambertz et al., 2002).

from 12 points per probe. The separation between each emitter–detector couple is 3 cm. Two probes are used, one for each of the cerebral hemispheres (**Figure 1**), providing a total of 24 recording sites.

Because we are interested in normal development and language acquisition, we focus our studies on healthy hearing newborns. Infants are considered eligible to participate in our studies if their gestational age is between 38 and 42 weeks, their Apgar score is at least eight in the first minute of life, no problems were observed in the hearing test, and do not present hematomas. Moreover, to maximize the likelihood of monitoring the same areas of the brain across different infants, we recruit only those whose head diameter ranges between 33.5 and 36.0 cm. Personnel of the Hospital carry out the whole recruitment process.

Each infant is tested individually inside a dimly lit sound-attenuated booth (**Figure 2**). The placement of the probes on the infant's head takes a few minutes. The use of pacifier is usually not necessary in this stage. Newborns come to the test session when they are in a quiet state of rest, either awake or sleeping. They remain in their own nursery crib and are tested lying in it, without receiving any reinforcement. A nurse or a pediatrician assists the neonates.

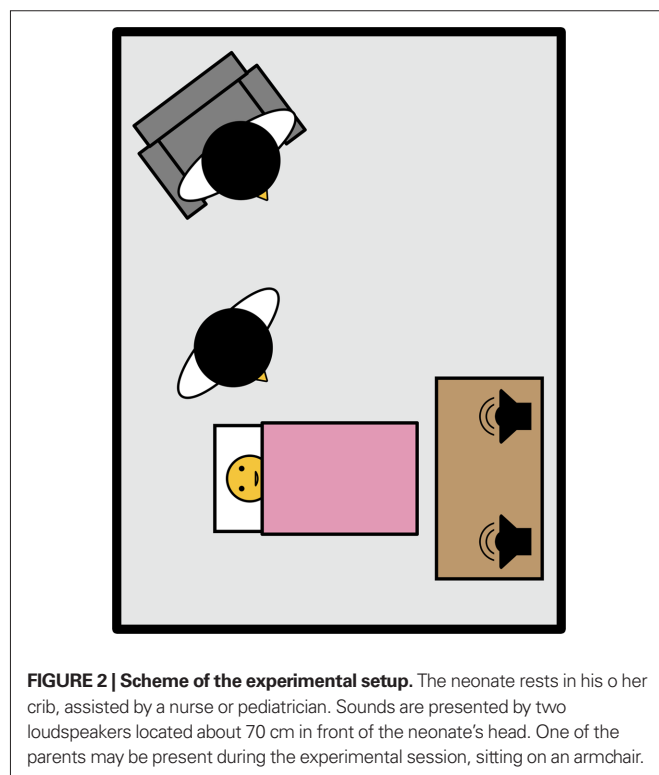
Because of our focus on language-related processes, we have mostly worked with auditory stimuli. These are presented at an appropriate intensity via two loudspeakers inside the experimental booth. The fNIRS machine and the control computer are placed outside the booth, to avoid undesired noise and heating in

the testing room. An infrared videocamera is used to monitor the infant's behavior. It is common that one or both parents attend the testing session. They choose whether to be inside the testing booth (provided that they will remain still and silent throughout the session) or outside, observing the infant's behavior through the online video recording. Parents are informed of how fNIRS functions, and they sign a consent form after they have understood how the experiment works and all their questions have been answered.

RELIABILITY OF THE METHOD

A long tradition in developmental research has used behavioral methods such as the high-amplitude sucking paradigm for tackling many infant speech perception questions. Thus, during our first experiences with fNIRS, we aimed to determine whether previous results obtained with these behavioral methods could be reproduced. While replicating previous findings, the fNIRS studies aimed to provide a link between behavioral observations and their underlying brain mechanisms.

Our very first study assessed the specialization of the brain to process speech stimuli at birth. In healthy adult participants, the left-hemisphere dominance for speech is a well-known fact supported by a wealth of clinical, behavioral, and brain imaging studies. The question that arose among developmental cognitive scientists was whether the left and right hemispheres are equal in function at birth and then specialized through experience, or whether at birth both hemispheres are already predisposed to process distinct types of stimuli. Subsequent research with young infants suggested that the second alternative was the most probable (Segalowitz and Chapman, 1980; Best, 1988;



Bertoncini et al., 1989; Holowka and Petitto, 2002). Many researchers used indirect behavioral measures to assess a diversity of questions: for instance, Bertoncini et al. (1989) tested 2-weeks-old infants with a dichotic listening technique during a high-amplitude sucking experimental session. Two stimuli, a syllable and a musical stimulus, were presented simultaneously to the infant, one in each ear. When infants listened to repeated syllables with their right ear, they displayed a change in behavior when a different syllable was presented, which was interpreted as a discrimination response. Instead, no discrimination occurred if infants listened to the linguistic material with their left ear. The converse was observed when the musical stimulus was changed.

Peña et al. (2003) used fNIRS to investigate the patterns of brain activation in response to auditory stimulation in full-term neonates. Their study thus provided a direct measure to assess hemispheric specialization while participants listened to normal speech, backward speech, or silence. Peña et al. (2003) found that the newborn left-hemisphere shows greater hemodynamic activity in response to normal speech than to backward speech² or silence. In addition, no area of the right hemisphere showed differential activation when contrasting forward and backward speech.

Peña et al.'s (2003) findings have yielded important theoretical and methodological implications. First, their study supported the notion of an early left-hemisphere specialization for speech, a fact intrinsically related to the emergence of the language faculty. Second, it validated fNIRS as a technique capable of both replicating previous behavioral results and enriching them by providing direct measurements of the activity in left and right perisylvian areas when processing speech.

Being a relatively young technique (see Wolf et al., 2007, for a history of near-infrared spectroscopy techniques), fNIRS must be evaluated also in terms of the reliability of the provided measures. In this respect, researchers have showed promising results for certain procedures based on fNIRS. For instance, Plichta et al. (2006) reported a high reproducibility at the group level, but not when re-testing single participants. Our studies do not address the issue of re-testing yet. However, the results of Peña et al. (2003) have been successfully reproduced in our laboratory in several occasions after the original work.

EXPERIMENTAL PROTOCOLS

Below we present in detail the two main testing protocols that we use in our research with neonates.

BLOCK DESIGN

We started our exploration of cognitive core language acquisition mechanisms by borrowing block designs from the fMRI tradition. We have used them in the aforementioned study by Peña et al. (2003), and also in Gervain et al. (2008).

Newborns are presented with sets of stimuli (blocks) corresponding to the different experimental conditions. The duration of the blocks and their content vary according to the purposes of

each study. For example, Peña et al. (2003) presented neonates with continuous stimuli: backward or forward infant-directed utterances in blocks of 15 s of duration. Instead, in the Gervain et al. (2008) studies, blocks were composed of 10 discrete items separated by pauses of varying length (0.5–1.5 s), yielding blocks of about 18 s. There were two kinds of blocks, which were presented in an interleaved fashion, avoiding the presentation of more than two consecutive blocks of the same condition (**Figure 3A**). From the first published study, our group used and continued using a variable separation of blocks (25–35 s) to avoid the effects of spontaneous oscillations frequently detected in fNIRS recordings (Diehl et al., 1998). For the statistical analyses we usually consider oxyhemoglobin or total hemoglobin concentration changes during the time window spanning from 10 to 20 s after the onset of stimulation, which roughly coincides with the plateau of the hemodynamic response. In each experimental block a given channel is rejected either because of movement artifact—that is, if the hemodynamic signal shows a variation per unit of time above a given threshold level—or saturation of the optical channel due to displacement of the probes. Only infants with a minimum amount of accepted block-channel pairs are further considered.

Block designs are useful to contrast brain responses to two or more experimental conditions (e.g., backward speech, forward speech, silence). One advantage is that one can compare conditions without the need of additional *ad hoc* control groups. However, a trade-off between number of contrasts and length of the experimental session should be considered, because the longer the session, the higher is the probability of the infant becoming fussy, and therefore of rejection.

It is important to highlight that experiments using this design have proven sensible not only for the study of prosodic or acoustic properties, but also for assessing learnability of specific structures in the first days of life: Gervain et al. (2008) examined newborns' ability to learn and detect repetition patterns (e.g., words like “mubaba,” “penana”), showing that these sequences evoke greater activation in specific areas of the newborn brain as compared to non-repeating sequences (e.g., words like “mubage,” “penaku”).

FAMILIARIZATION-RECOGNITION DESIGN

Behavioral methods have delivered very useful data that increased our understanding of cognitive development and language acquisition. The habituation-dishabituation method was the most frequently used to test neonates, yielding important findings that advanced our knowledge of how infants acquire language³. Among many other colleagues that applied this paradigm, we highlight the work by Eimas et al. (1971), the cardinal study who first discovered that 1- and 4-month-olds distinguish phonemes that differ by one feature, for example [b] from [p] or [b] from [d]. Furthermore, some years later Kuhl (1983) explored how infants categorize syllables despite the variability of the speech stimuli, and Mehler et al. (1978) discovered that newborns recognize

²Backward speech is widely considered a very good control to compare whether infants specifically react to language stimuli and not to some physical features such as intensity, duration, or pitch. Indeed, a snippet of language played backward has the same energy of the corresponding stimuli played forward, but all the phonetic, phonological, syntactical, and prosodic transitions are lost.

³We refer to the vast majority of studies concerning language acquisition and infants' general auditory capacities. In this field, the most common behavioral procedure was high-amplitude sucking.

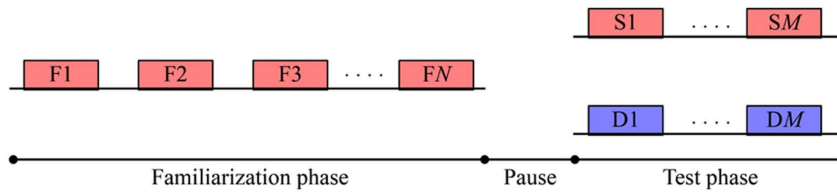
A Blocked design**B Familiarization-recognition design**

FIGURE 3 | Scheme depicting the experimental protocols used in our laboratory. Each rectangle represents a block, consisting in either continuous auditory stimulation or a series of discrete sounds separated by short pauses of about 1 s. Consecutive blocks are separated by silent intervals of varying duration. **(A)** The block design consists in the intermixed presentation of N blocks belonging to

each of the conditions X and Y. **(B)** The familiarization-recognition design consists in a Familiarization phase (F) in which all neonates listen to N blocks of stimulation. After a silent retention interval—Pause—that can last for some minutes, a Test phase follows. Here, half of the group of infants listens to M stimuli blocks of the same kind (S) of the familiarization, whereas the other half listens to a different kind of stimuli (D).

their own mother's voice (see also Jusczyk, 1997, for an extended review on infant's speech-processing studies based on behavioral methods).

Notwithstanding the important findings obtained with the habituation-dishabituation paradigm, the studies using it suffered from the aforementioned problems of behavioral methods such as high rejection rate. Moreover, using only behavioral paradigms restricted investigations to discrimination capacities mostly, leaving practically unattended other cognitive functions equally important for language acquisition. For instance, very little is known about the memory capacities of neonates, in spite of the great progress that cognitive scientists and neuroscientists have made in identifying brain regions underlying memory processes in adults (e.g., Nyberg and Cabeza, 2000; Baddeley et al., 2009).

One of the most important challenges of fNIRS is to address questions that are difficult to assess behaviorally. Benavides-Varela, Gómez, Bion, Macagno, Peretz and Mehler (in preparation) adapted a habituation-deshabituation paradigm for testing newborn's memory for speech sounds with fNIRS, which could also provide indications of neural activity associated with encoding and recognition at birth.

Benavides-Varela and Colleagues tested neonates between 2 and 5 days old in a recognition memory task. Newborns listened to a single CVCV word repeated for 10 blocks (each block contained six identical words) for a total of 6 min of exposure⁴ (see **Figure 3B**). A 2-min-long silent pause separated the familiarization and test phases, providing a means of tapping infants' memory. After the pause, the same word of the familiarization phase was presented to half of the infants, whereas the other half listened to a novel word. The test consisted of five blocks (3 min). Differential responses in hemoglobin concentration between the two groups of newborns would indicate that the familiar stimulus is remembered. Results were auspicious: participants hearing a novel word showed higher

relative concentrations of oxyhemoglobin in the first block of the test phase, as compared to neonates who heard again the familiar word. fNIRS revealed these differences in a bilaterally distributed network, involving temporo-parietal and anterior areas. These results open a field of possibilities for studying and better understanding the development of early memory capacities, and represent a promising step forward with respect to behavioral approaches.

The use of fNIRS has the advantage that the experimental session can be shortened or lengthened to study important variables such as amount of exposure or long-lasting retention. This represented a problem in behavioral studies, because infants often fall asleep during long silent pauses, or because the behavioral responses of the control groups may become noisy (see Jusczyk et al., 1995). Furthermore, the findings of Benavides-Varela and Colleagues suggest that 6 min of familiarization (see footnote 4) are enough for newborns to form a lasting representation of the presented word. This exposure time is considerably shorter than the ones used in comparable behavioral studies.

The usual implementation of the habituation-dishabituation procedure requires the adoption of a criterion for shifting from habituation phase to test phase, depending on each infant's behavior. The amount of exposure required to achieve this criterion differs considerably across infants. We point out that the fNIRS version of this method allowed us to equate the amount of exposure each infant received in the familiarization phase. This is particularly important in memory studies, in which the amount of exposure partly determines the robustness of the memory trace, and therefore of the recognition response. Finally, we stress the fact that the fNIRS technique not only provides a yes/no answer to the experimental question at hand, but also informs us of the activation of several cortical areas.

We believe that this type of design will prove useful for studies aiming beyond infants' discrimination abilities and preferences, representing an opportunity to track the time course of learning, and the encoding and recognition processes in newborns and infants at different stages of development.

⁴Of this total duration, however, the fraction corresponding to effective acoustic stimulation was about 1 min.

CONCLUDING REMARKS

Developmental scientists face the challenge of devising reliable methodologies to assess cognitive capacities in neonates. The incorporation of the fNIRS imaging technique promotes our understanding of early language acquisition and memory capacities. We hope that the successful experiences with fNIRS in several laboratories (e.g., Peña et al., 2003; Bortfeld et al., 2007; Gervain et al., 2008; Nakano et al., 2009; Lloyd-Fox et al., 2010; Obrig et al., 2010) will encourage further exploitation of this valuable tool for studying early human development.

An important line of methodological progress is focusing on designs that combine different paradigms (behavioral and neuroimaging) or techniques (e.g., fNIRS and electroencephalography), either across different experimental sessions or in simultaneous recordings. The use of two or more techniques has the potential to provide complementary information about the functioning of the newborn's brain. A pioneer study by Telkemeyer et al. (2009) has

already used NIRS and EEG concurrently to test auditory processing in newborns. We are confident that such adaptations will play a central role in improving our knowledge about human innate cognitive capacities.

ACKNOWLEDGMENTS

We thank Laurence White for helpful comments on earlier versions of this manuscript; Marijana Sjekloča, Francesca Gandolfo, and Alessio Isaja for their permanent administrative and technical support; Dr. Francesco Macagno and the personnel of Udine's Hospital Neonatology and Obstetrics Departments for their assistance in the recruitment of neonates. We also thank the parents of the young participants for their collaboration. This work was supported by McDonnell Foundation Grant 21002089, and a grant from Ministerio de Ciencia y Tecnología (MICIT) and Consejo Nacional de Investigaciones Científicas y Tecnológicas (CONICIT) of Costa Rica to the first author.

REFERENCES

- Aslin, R. N., and Mehler, J. (2005). Near-infrared spectroscopy for functional studies of brain activity in human infants: promise, prospects, and challenges. *J. Biomed. Opt.* 10, 0001–0003.
- Baddeley, A., Eysenck, M. W., and Anderson, M. C. (2009). *Memory*. New York: Psychology Press.
- Bertoncini, J., Morais, J., Bijeljac-Babic, R., McAdams, S., Peretz, I., and Mehler, J. (1989). Dichotic perception and laterality in neonates. *Brain Lang.* 37, 591–605.
- Best, C. T. (1988). "The emergence of cerebral asymmetries in early human development: a literature review and a neuroembryological model," in *Brain Lateralization in Children*, eds D. L. Molfese and S. J. Segalowitz (New York: Guilford), 3–5.
- Bortfeld, H., and Wruck, E., and Boas, D. A. (2007). Assessing infants' cortical response to speech using near-infrared spectroscopy. *Neuroimage* 34, 407–415.
- Cheour, M., Imada, T., Taulu, S., Ahonen, A., Salonen, J., and Kuhl, P. (2004). Magnetoencephalography is feasible for infant assessment of auditory discrimination. *Exp. Neurol.* 190, S44–S51.
- Cheour, M., Martynova, O., Näätänen, R., Erkkola, R., Sillanpää, M., Kero, P., Raz, A., Kaipio, M.-L., Hiltunen, J., Aaltonen, O., Savela, J., and Hämäläinen, H. (2002). Speech sounds learned by sleeping newborns. *Nature* 415, 599–600.
- Dehaene-Lambertz, G., Dehaene, S., and Hertz-Pannier, L. (2002). Functional neuroimaging of speech perception in infants. *Science* 298, 2013–2015.
- Dehaene-Lambertz, G., and Peña, M. (2001). Electrophysiological evidence for automatic phonetic processing in neonates. *Neuroreport* 12, 3155–3158.
- Diehl, R. R., Linden, D., Lücke, D., and Berlitz, P. (1998). Spontaneous blood pressure oscillations and cerebral autoregulation. *Clin. Auton. Res.* 8, 7–12.
- Eimas, P. D., Siqueland, E. R., Jusczyk, P. W., and Vigorito, J. (1971). Speech perception in infants. *Science* 171, 303–306.
- Gervain, J., Macagno, F., Cogoi, S., Peña, M., and Mehler, J. (2008). The neonate brain detects speech structure. *Proc. Natl. Acad. Sci. U.S.A.* 105, 14222–14227.
- Gervain, J., Mehler, J., Werker, J. F., Nelson, C. A., Csibra, G., Lloyd-Fox, S., Shukla, M., and Aslin, R. N. (2011). Near-infrared spectroscopy: a report from the McDonnell infant methodology consortium. *Dev. Cogn. Neurosci.* 1, 22–46.
- Holowka, S., and Petitto, L. A. (2002). Left hemisphere cerebral specialization for babies while babbling. *Science* 297, 1515.
- Imada, T., Zhang, Y., Cheour, M., Taulu, S., Ahonen, A., and Kuhl, P. (2006). Infant speech perception activates Broca's area: a developmental magnetoencephalography study. *Neuroreport* 17, 957–962.
- Jusczyk, P. W. (1997). *The Discovery of Spoken Language*. Cambridge, MA: MIT Press.
- Jusczyk, P. W., Kennedy, L. J., and Jusczyk, A. M. (1995). Young infants' retention of information about syllables. *Infant. Behav. Dev.* 18, 27–41.
- Kuhl, P. K. (1983). Perception of auditory equivalence classes for speech in early infancy. *Infant. Behav. Dev.* 6, 263–285.
- Kushnerenko, E., Cheour, M., Ceponiene, R., Fellman, V., Renlund, M., Soininen, K., Alku, P., Koskinen, M., Sainio, K., and Näätänen, R. (2001). Central auditory processing of durational changes in complex speech patterns by newborns: an event-related brain potential study. *Dev. Neuropsychol.* 19, 83–97.
- Lloyd-Fox, S., Blasi, A., and Elwell, C. E. (2010). Illuminating the developing brain: the past, present and future of functional near infrared spectroscopy. *Neurosci. Biobehav. Rev.* 34, 269–284.
- Mehler, J., Bertoncini, J., Barriere, M., and Jassik-Gerschenfeld, D. (1978). Infant recognition of mother's voice. *Perception* 7, 491–497.
- Minagawa-Kawai, Y., Mori, K., Hebden, J. C., and Dupoux, E. (2008). Optical imaging of infants' neurocognitive development: recent advances and perspectives. *Dev. Neurobiol.* 68, 712–728.
- Nakano, T., Watanabe, H., Homae, F., and Taga, G. (2009). Prefrontal cortical involvement in young infants' analysis of novelty. *Cereb. Cortex* 19, 455–463.
- Nyberg, L., and Cabeza, R. (2000). "Brain imaging of memory," in *The Oxford Handbook of Memory*, eds E. Tulving and F. I. M. Craik (Oxford University Press, New York), 501–519.
- Obrig, H., Rossi, S., Telkemeyer, S., and Wartenburger, I. (2010). From acoustic segmentation to language processing: evidence from optical imaging. *Front. Neuroenergetics* 2:13. doi: 10.3389/fnene.2010.00013
- Peña, M., Maki, A., Kovačič, D., Dehaene-Lambertz, G., Koizumi, H., Bouquet, F., and Mehler, J. (2003). Sounds and silence: an optical topography study of language recognition at birth. *Proc. Natl. Acad. Sci. U.S.A.* 100, 11702–11705.
- Plichta, M. M., Herrmann, M. J., Baehne, C. G., Ehli, A.-C., Richter, M. M., Pauli, P., and Fallgater, A. J. (2006). Event-related functional near-infrared spectroscopy (fNIRS): are the measurements reliable? *Neuroimage* 31, 116–124.
- Segalowitz, S. J., and Chapman, J. S. (1980). Cerebral asymmetry for speech in neonates: a behavioral measure. *Brain Lang.* 9, 281–288.
- Telkemeyer, S., Rossi, S., Koch, S. P., Nierhaus, T., Steinbrink, J., Poeppel, D., Obrig, H., and Wartenburger, I. (2009). Sensitivity of newborn auditory cortex to the temporal structure of sounds. *J. Neurosci.* 29, 14726–14733.
- Wolf, M., Ferrari, M., and Quaresima, V. (2007). Progress of near-infrared spectroscopy and topography for brain and muscle clinical applications. *J. Biomed. Opt.* 12, 062104.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 19 January 2011; accepted: 28 March 2011; published online: 18 April 2011.

Citation: Benavides-Varela S, Gómez DM and Mehler J (2011) Studying neonates' language and memory capacities with functional near-infrared spectroscopy. *Front. Psychology* 2:64. doi: 10.3389/fpsyg.2011.00064

This article was submitted to *Frontiers in Language Sciences*, a specialty of *Frontiers in Psychology*.

Copyright © 2011 Benavides-Varela, Gómez and Mehler. This is an open-access article subject to a non-exclusive license between the authors and Frontiers Media SA, which permits use, distribution and reproduction in other forums, provided the original authors and source are credited and other Frontiers conditions are complied with.



Language and the newborn brain: does prenatal language experience shape the neonate neural response to speech?

Lillian May^{1*}, Krista Byers-Heinlein², Judit Gervain³ and Janet F. Werker¹

¹ Department of Psychology, University of British Columbia, Vancouver, BC, Canada

² Department of Psychology, Concordia University, Montreal, QC, Canada

³ Department of Psychology, CNRS, Université Paris Descartes, Paris, France

Edited by:

Heather Bortfeld, University of Connecticut, USA

Reviewed by:

Nadege Roche-Labarbe, Massachusetts General Hospital, USA
Ioulia Kovelman, University of Michigan, USA

*Correspondence:

Lillian May, Department of Psychology, University of British Columbia, 2136 West Mall, Vancouver, BC, Canada V6T 1Z4.
e-mail: lamay@psych.ubc.ca

Previous research has shown that by the time of birth, the neonate brain responds specially to the native language when compared to acoustically similar non-language stimuli. In the current study, we use near-infrared spectroscopy to ask how prenatal language experience might shape the brain response to language in newborn infants. To do so, we examine the neural response of neonates when listening to familiar versus unfamiliar language, as well as to non language stimuli. Twenty monolingual English-exposed neonates aged 0–3 days were tested. Each infant heard low-pass filtered sentences of forward English (familiar language), forward Tagalog (unfamiliar language), and backward English and Tagalog (non-language). During exposure, neural activation was measured across 12 channels on each hemisphere. Our results indicate a bilateral effect of language familiarity on neonates' brain response to language. Differential brain activation was seen when neonates listened to forward Tagalog (unfamiliar language) as compared to other types of language stimuli. We interpret these results as evidence that the prenatal experience with the native language gained *in utero* influences how the newborn brain responds to language across brain regions sensitive to speech processing.

Keywords: language, near-infrared spectroscopy, neonates

INTRODUCTION

It is well known that the adult brain is specialized in its response to native language (Perani et al., 1996; Dehaene et al., 1997). Recent evidence has suggested that the human brain is tuned to language from the earliest stages of development. Only a few days after birth, neonates respond differently to language than to non-linguistic sounds. Very young infants demonstrate a preference for listening to speech over non-speech (Vouloumanos and Werker, 2007), and are capable of discriminating languages from different rhythmical classes (Mehler et al., 1988; Nazzi et al., 1998; Ramus et al., 2000). However, what is unknown from past research is the extent to which early prenatal experience with language may play a role in determining the organization of neonates' neural tuning for language. In particular, no one has yet investigated whether the experience that neonates have with the native language while *in utero* influences the pattern and location of brain activity to familiar versus unfamiliar language. In the current study, we use near-infrared spectroscopy (NIRS) to take the first steps in exploring this question.

Research to date examining the neonate brain response to language versus non-language has shown that brain responses to familiar language are both stronger and more specialized when compared to the response to non-language (Dehaene-Lambertz et al., 2002, 2010; Pena et al., 2003). Using behavioral methods, a left hemisphere advantage for language has been inferred through dichotic listening to individual syllables as measured by high-amplitude sucking in newborns (Bertoncini et al., 1989), as well as through mouth asymmetries during babbling in 5 to

12-month-olds (Holowka and Petitto, 2002). In neuroimaging research, optical imaging studies with newborns have shown a greater left hemisphere response to audio recordings of forward versus backward speech (Pena et al., 2003), as well as evidence that the left hemisphere plays an important role in processing repetition structures in language (e.g., ABB versus ABC syllable sequences; Gervain et al., 2008). Similarly, fMRI studies with infants 2–3 months of age indicate differential responses in the left hemisphere to continuous forward versus backward speech (Dehaene-Lambertz et al., 2002), and to speech versus music (Dehaene-Lambertz et al., 2010). These functional studies are supported by structural MRI analyses indicating asymmetries at birth in the left hemisphere language areas of the brain (Dubois et al., 2010). All of the above studies, however, have focused on young infants' neural response to familiar language, leaving open the question of how much responses may have been driven by language experience.

At birth, neonates are experiencing extra-uterine language for the first time. However, *in utero* they have had the opportunity to learn about at least some of the properties of language. The peripheral auditory system is mature by 26 weeks gestation (Eisenberg, 1976), and the properties of the womb are such that the majority of low-frequency sounds (less than 300 Hz) are transmitted to the fetal inner ear (Gerhardt et al., 1992). The low-frequency components of language that are transmitted through the uterus include pitch, some aspects of rhythm, and some phonetic information (Querleu et al., 1988; Lecanuet and Granier-Deferre, 1993). Moreover, the fetus has access to the mother's speech via bone

conduction (Petitjean, 1989). There is evidence that the fetus can hear and remember language sounds even before birth. Fetuses respond to and discriminate speech sounds (Lecanuet et al., 1987; Zimmer et al., 1993; Kisilevsky et al., 2003). Moreover, newborn infants show a preference for their mother's voice at birth (DeCasper and Fifer, 1980) and show behavioral recognition of language samples of children's stories heard only during the pregnancy (DeCasper and Spence, 1986). Finally, and of particular interest to our work, newborn infants born to monolingual mothers prefer to listen to their native language over an unfamiliar language from a different rhythmical class (Mehler et al., 1988; Moon et al., 1993). These studies suggest that infants may have learned about the properties of the native language while still in the womb.

In a recent extension of the work showing a preference for the native language at birth, Byers-Heinlein et al. (2010) investigated how prenatal bilingual experience influences language preference at birth. Infants from 0 to 5 days of age born to either monolingual English or bilingual English–Tagalog mothers were tested in a high-amplitude sucking procedure. Infants were played sentences in both English (a stress-timed language) and Tagalog (a Filipino language that is syllable-timed). Sentences from both languages were low-pass filtered (to a 400-Hz cut-off), to maintain the rhythmical information of each language while eliminating most surface segmental cues that may be different across languages. Byers-Heinlein et al. (2010) found that while all infants could discriminate English and Tagalog, the monolingual-exposed infants showed a preference for only English and the bilingual-exposed infants had a similar preference for both English and Tagalog. These results provide strong evidence that language preference at birth is influenced by the language heard *in utero*, even when infants have had prenatal experience with multiple languages. The neural correlates of this behavioral preference for familiar language(s) at birth are however, unknown.

A recent neuroimaging study of infants' processing of speech and non-speech has provided some support for the hypothesis that language experience may impact early neural specialization for processing some aspects of language (Minagawa-Kawai et al., 2011). Using NIRS, 4-month-old Japanese infants' brain response was assessed while listening to a familiar language (Japanese), to an unfamiliar language (English), and to different non-speech sounds (emotional voices, monkey calls, and scrambled speech). Greater left hemisphere activation was reported for both familiar and unfamiliar language when compared to the non-speech conditions. Critically, activation was also significantly greater to the familiar language when compared to the unfamiliar language. This latter finding implies that by 4 months of age, the young brain responds differently to familiar versus unfamiliar language and is thus influenced by language experience. However, the infants studied by Minagawa-Kawai et al. (2011) were 4 months of age – meaning that these infants have dramatically more experience with their native language than newborn infants. It is unknown whether infants with only a few hours of post-natal experience will show a similar difference in neural activation to a familiar versus an unfamiliar language.

In contrast to the above studies demonstrating the impact of language experience on newborn infants' language processing,

other areas of research have uncovered aspects of language perception that appear unaffected by specific language experience early in development. For example, neonates' rhythm-based language discrimination has been shown to be based on language-general abilities. Phonologists have traditionally classified the world's languages into three main rhythmic categories: stress-timed (e.g., English, Dutch), syllable-timed (e.g., Spanish, French), and mora-timed (e.g., Japanese). This distinction is critically important to language learning as rhythmicity is associated with word order in a language (Nespor et al., 2008), rendering it one of the most potentially informative perceptual cues for bootstrapping language acquisition. Recent cross-linguistic investigations have more finely quantified the distinction between rhythmical classes, finding that languages fall into rhythmical class on the basis of two parameters: percent vowel duration within a sequence and the standard deviation of the duration of consonantal intervals (Ramus et al., 1999; see also Grabe and Low, 2002 for a different measurement scheme).

In a long series of studies, it has been demonstrated that young infants are able to discriminate languages from different rhythmical classes (Mehler et al., 1988; Nazzi et al., 1998; Ramus et al., 2000). This ability does not depend on familiarity with one or both of the languages being tested. Infants with prenatal experience with a single language can discriminate the native language from a rhythmically dissimilar unfamiliar language (Mehler et al., 1988), as well as discriminate two unfamiliar rhythmically different languages (Nazzi et al., 1998). Further, infants with prenatal bilingual exposure are able to discriminate their two native languages when those languages are from different rhythmical classes, even though both languages are familiar (Byers-Heinlein et al., 2010). These findings show that rhythm-based language discrimination in newborns is not based on experience with the native language, but instead on initial universal biases. It therefore may be the case that the early neural response to language in neonates also reflects similar language-universal processing.

CURRENT STUDY

The goal of the present study was to test whether neonates' early brain specialization for language is driven exclusively by a universal preparation for language, or whether there is influence from prenatal language experience. To test these competing hypotheses, we measured the patterns and location of neonates' brain response to a familiar (the primary language heard *in utero*) versus an unfamiliar language. Building on previous research, we compared the pattern and location of neural responses to forward speech versus backward speech in both a familiar and an unfamiliar language. We tested newborn infants, an age group that has not previously been tested for the influence of listening experience on neural organization.

We employed NIRS to measure neural activity in neonates when listening to familiar and unfamiliar language. Participants in the current study were born to monolingual English-speaking mothers, and each infant was tested in four language conditions. In two forward-language conditions, neonates were played sentences of adult-directed English and Tagalog. In two backward-language control conditions, infants were played the same English and Tagalog sentences reversed. Backward speech has often been used

in neuroimaging studies exploring the brain response to language versus no-language, including both studies with adults (Perani et al., 1996; Carreiras et al., 2005) and infants (Dehaene-Lambertz et al., 2002; Pena et al., 2003). Backward-language is believed to be a useful non-linguistic control because it matches the forward-language in both intensity and pitch, but is distinctly non-linguistic, as humans are unable to produce some of the sound sequences created (such as backward aspirated stops), because words in backward speech do not have a proper syllable form, and because the prosodic structure of sentences is disturbed. As a consequence, infants fail to discriminate a pair of reversed languages despite succeeding in differentiating the same pair of forward-played languages (Mehler et al., 1988). This suggests that backward spoken language does not carry the same linguistic relevance as forward speech (Mehler et al., 1988; Ramus et al., 2000).

The language stimuli used in the current study were the identical low-pass filtered sentences used in the previously described behavioral study investigating language preference and discrimination in neonates conducted by Byers-Heinlein et al. (2010). Low-pass filtered speech has been used by many cross-linguistic preference and discrimination studies (Mehler et al., 1988; Nazzi et al., 1998; Byers-Heinlein et al., 2010), as the filtering is believed to eliminate surface acoustic and phonetic cue differences between languages that may lead to irrelevant preferences. The filtering allows much of the rhythmical structure of the language to remain intact, and infants' ability to use rhythmical information to discriminate between languages is thought to remain the same across unfiltered and filtered language (Mehler et al., 1988). Filtered speech was used in the present study as it likely mimics the properties of language perceived *in utero* (Lecanuet and Granier-Deferre, 1993). As such, we expected that the neonate response to filtered speech should reflect how initial prenatal experience might shape the neural response to speech. It should be noted, however, that while there is considerable behavioral work on language preference and discrimination using filtered speech, ours is the first study to use filtered speech in a neural imaging study addressing these questions. Given that some of the features of speech important to neural localization may be removed in filtering the stimulus (i.e., the fast phonetic change dynamics; Zatorre and Belin, 2001; Poeppel, 2003; Zatorre and Gandour, 2007), we anticipated that the neural response to filtered speech might differ from patterns of results found previously with unfiltered speech.

MATERIALS AND METHODS

PARTICIPANTS

Twenty full-term, healthy neonates (ranging in age from 0 to 3 days, mean age = 1.6 days) born to English-speaking mothers were included in the analyses. All mothers reported speaking at least 90% English during their pregnancy, and no Tagalog (19 mothers reported speaking 100% English, and one mother 90% English and 10% Romanian). An additional 10 infants were tested, but were excluded due to the infant becoming awake or fussy and failing to complete the procedure (5), equipment failure (2), insufficient analyzable data (2), or parental interference (1). All infants were tested while asleep or during a quiet state of wakefulness. All

parents of infants gave informed consent prior to beginning the experiment.

STIMULUS MATERIALS

The language samples used in the current study were taken from those used by Byers-Heinlein et al. (2010). Stimuli consisted of six English sentences and six Tagalog sentences recorded by native speakers of each language and spoken in an adult-directed manner. All sentences were matched in pitch, duration, and number of syllables, and were produced by adult native language speakers. Sentences were low-pass filtered to 400 Hz, to remove surface segmental cues while maintaining rhythmical structure and prosody. Backward-language sentences were formed by reversing the English and Tagalog sentences using Praat (Boersma and Weenink, 2011). Sentence lengths in English ranged from 3.28 to 4.09 s, with a mean of 3.55 s. Sentence lengths in Tagalog ranged from 3.07 to 4.19 s, with a mean of 3.61 s.

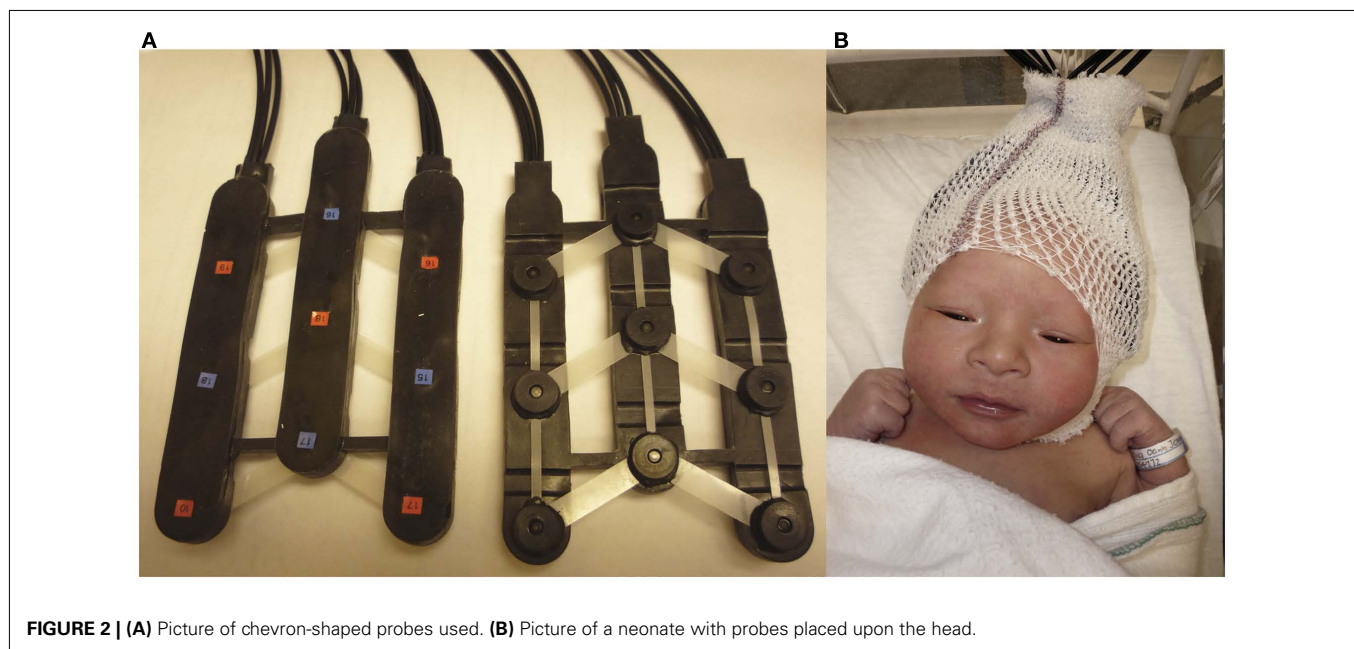
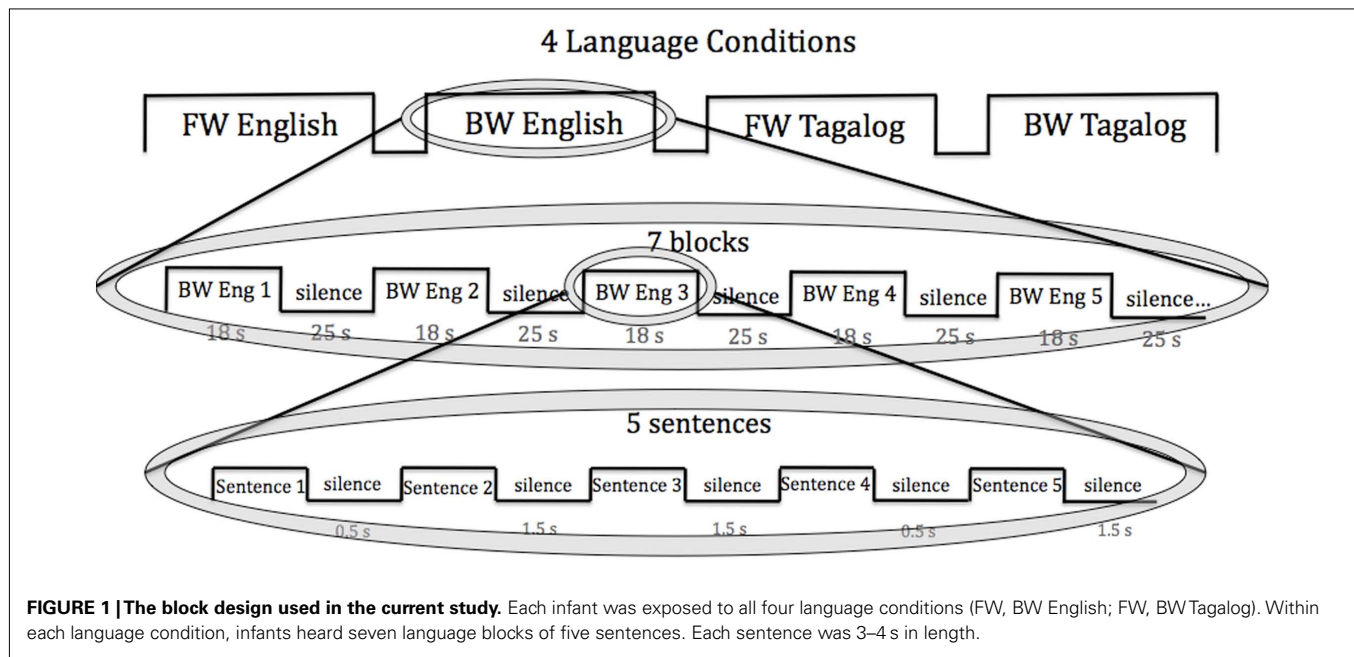
Each infant was tested in all four language conditions: forward English, forward Tagalog, backward English, and backward Tagalog. The conditions were randomly ordered across infants and presented consecutively. The blocked design was chosen as it has been used by many infant NIRS studies (e.g., Pena et al., 2003; for a review see Gervain et al., 2011). Each condition lasted 5.6 min. Within each condition, stimuli were organized within seven blocks that each lasted 18–20 s. Each block consisted of five sentences. There were six sentences in total for each language, and for each block five different sentences were randomly selected. Within a block, the five sentences were separated by brief pauses of variable length (0.5–1.5 s), following Gervain et al. (2008). Blocks were separated from each other by 25–35 s of silence. The total testing time for each infant was 22.4 min. The block design used is presented in Figure 1.

PROCEDURE

Neonates were tested in a local maternity hospital, while asleep or at rest in a bassinet. Testing occurred in a silent, private experimental room. A Hitachi ETG-4000 NIRS machine with a source detector separation of 3 cm and two continuous wavelengths of 695 and 830 nm was used to record the NIRS signal, using a sampling rate of 10 Hz. For further technical details regarding the machine, see Gervain et al. (2011).

Two chevron-shaped probes were used, each consisting of nine 1 mm optical fibers. Of these nine fibers, five were emitters and four detectors. As such, there were 12 recording channels in each probe. One probe set was placed over the perisylvian area of the neonate's scalp of the left hemisphere, with the second probe set over the symmetrical area of the right hemisphere. The chevron shape of the probes was situated to nestle above the infant's ears (see Figure 2 for image of the probes, and probes placed on infant; see Figure 3 for probe configuration). A stretchy cap was used to keep the probes in place. The NIRS machine used a laser power of 0.75 mW.

A MacBook laptop or a Mac Mini desktop computer running Psyscope × (Build 36) controlled the experiment, playing the language stimuli and sending markers to the NIRS machine. The language stimuli were played through two speakers approximately



1.5 m from the infants' head. The intensity of the stimuli was set to 70–75 dB.

DATA ANALYSIS

Analyses were initially conducted on oxyHb and deoxyHb in a time window between 0 and 35 s after stimulus onset to capture the full time course of the hemodynamic response in each block (Gervain et al., 2008). Data were averaged across blocks within the same condition. Data were band-pass filtered between 0.01 and 0.7 Hz, as to remove low-frequency noise (i.e., slow drifts in Hb concentrations) as well as high frequency noise (i.e., heartbeat). Movement

artifacts were removed by isolating blocks in which a change in concentration greater than $0.1 \text{ mmol} \times \text{mm}$ over a period of 0.2 s, i.e., two samples, occurred, and rejecting the block. On average, 3.69 blocks were retained for data analysis in the English FW, 3.17 in the English BW, 3.46 in the Tagalog FW, and 3.61 in the Tagalog BW condition. For all retained blocks, a baseline was established by linearly fitting the 5-s preceding the onset of the block and the 5-s beginning 15 s after the end of the block. This timeline is used to allow the hemodynamic response function that occurs in response to the experimental stimuli to return to the original steady state (Pena et al., 2003; Gervain et al., 2008).

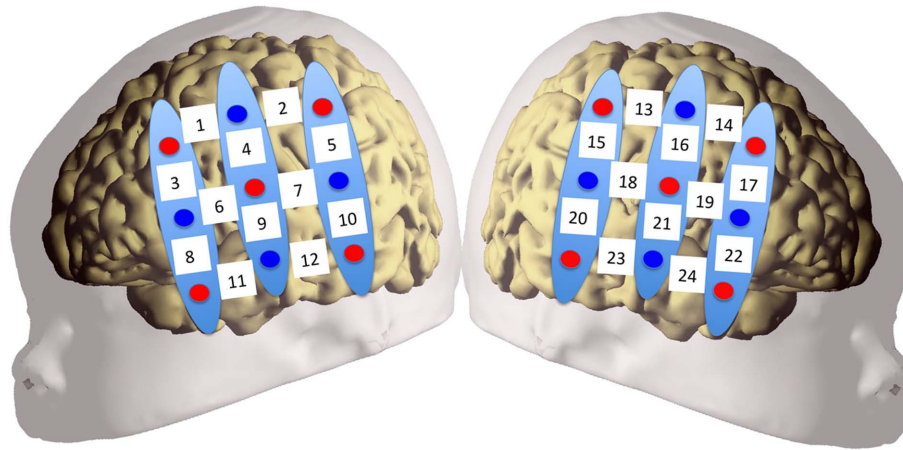


FIGURE 3 | Configuration of probe sets overlaid on a schematic infant head. Red circles indicate emitter fibers, while blue circles indicate detectors. Separation in all emitter–detector channels was 3 cm. The probes were placed on the participants' heads using surface landmarks, such as the ears or the

vertex. Channels 1–12 were placed over the left hemisphere, while channels 13–24 were placed over the right. Probes were placed so that the bottom-most channels (11, 12 in LH; 23, 24 in RH) ideally nestled above the infant's ear.

The region of interest (ROI) was defined following Pena et al. (2003). Channels 7–12 and 19–24 were chosen as the ROI in each hemisphere. These ROIs correspond to the lower ROIs in Pena et al. (2003), which is the area where significant activation was found in that study. These ROIs comprise the temporal (auditory processing) brain areas, where one can expect to find the strongest speech-related response.

RESULTS

The grand average results of the experiment are presented in Figures 4 and 5. The figures show the averages of oxyHb and deoxyHb concentration change in all blocks for each condition across all infants. A table of the oxyHb results is presented in Figure 6. We conducted a repeated measures analysis of variance (ANOVA) within the target ROI (lower channels, as used by Pena et al., 2003) with factors Language (English/Tagalog) \times Direction (BW/FW) \times Hemisphere (LH/RH) separately for oxyHb and deoxyHb, similar to Pena et al.'s (2003) analysis. The ANOVA for oxyHb yielded a significant main effect for Direction [$F(1,19) = 5.342$, $p = 0.032$], as BW speech gave rise to a larger response than FW speech. The interaction between Language \times Direction was marginally significant [$F(1,19) = 3.882$, $p = 0.064$], as FW Tagalog gave rise to a decrease in oxyHb (inverted response), whereas FW and BW English as well as BW Tagalog resulted in an increase in oxyHb (canonical response; significant and marginal Bonferroni *post hoc* tests: FW Tagalog versus BW Tagalog $p = 0.002$; FW Tagalog versus FW English $p = 0.054$; FW Tagalog versus BW English $p = 0.077$; Figure 7). A similar ANOVA with deoxyHb yielded no significant results.

DISCUSSION

Our findings demonstrate that the neural processing of language is influenced by language experience even by the first few days of life. When newborn infants listened to English (familiar) and Tagalog (unfamiliar) language stimuli, we observed a difference

in brain response. When processing forward-played sentences of English, neonates showed an increase in overall oxygenated hemoglobin across both hemispheres. In contrast, when infants listened to sentences of unfamiliar forward Tagalog, we observed a decrease in oxygenated hemoglobin. No language familiarity effects were found in the brain response to backward speech, as neonates had a similar neural response to backward English and backward Tagalog. While we observed different patterns of brain activation to forward English versus forward Tagalog, we did not find a consistent difference in the localization of brain activity between language conditions. For both English and Tagalog, similar patterns of activation were found in the temporal regions across the left and right hemispheres. Our results therefore suggest that prenatal language experience does shape how the brain responds to familiar and unfamiliar language. These results echo behavioral findings using the same stimuli (Byers-Heinlein et al., 2010), where neonates were shown to both prefer and discriminate a familiar language from an unfamiliar language. However, at least with the filtered language stimuli used in our study, we find no evidence that the neonate brain uses distinct brain regions to process different languages.

Our data also produced several unexpected findings. First, the lack of any observed hemisphere differences in neonates' response to familiar or unfamiliar language contrasts with previous studies showing left hemisphere dominance for language processing in young infants in the area of the planum temporale (Dehaene-Lambertz et al., 2002; Pena et al., 2003). This is surprising given the left lateralization of neonate brain response to language found by Pena et al. (2003) which used very similar methodology to our study. We propose two hypotheses to explain the difference in hemispheric findings between the current study and Pena et al. (2003). One possibility is that the subtle differences in procedure led to the differential findings. While we attempted to place the probes in similar temporal areas to Pena et al. (2003) it is impossible to know if the placement was completely comparable across

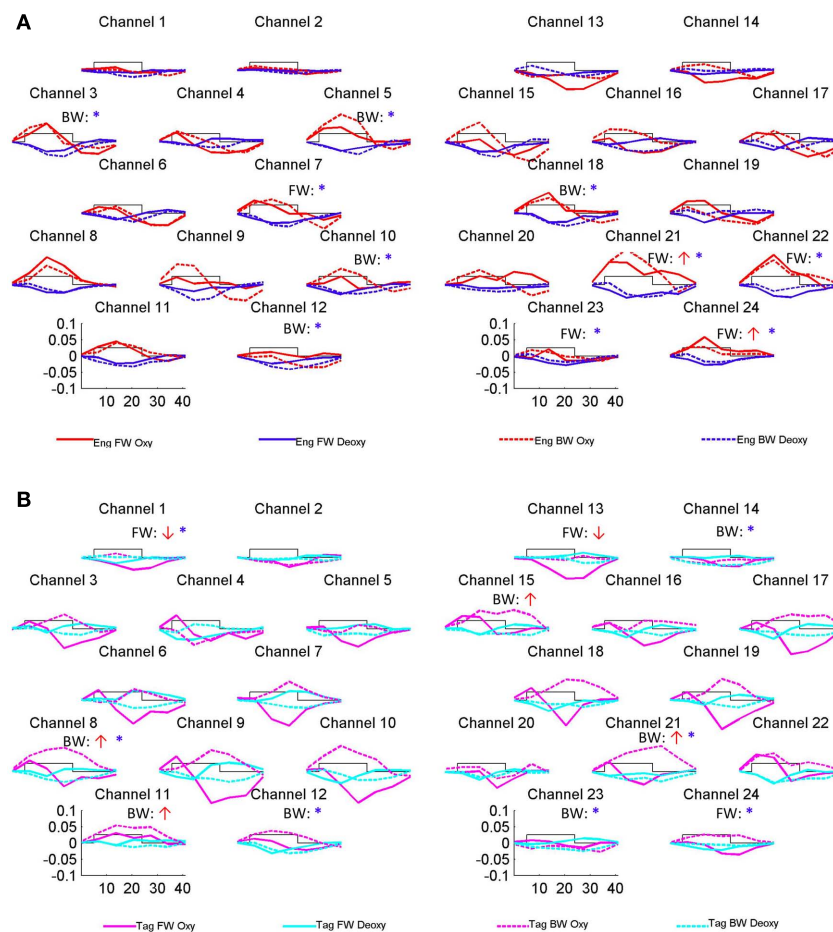


FIGURE 4 | Grand average results of the experiment. Numbers and location of channels correspond to the placement shown in **Figure 3**. **(A)** The oxyHb and deoxyHb response to the English FW and BW conditions. **(B)** The oxyHb and deoxyHb response to the Tagalog FW and BW conditions. The arrows indicate significant changes in concentrations in channel-by-channel *t*-tests ($p < 0.05$, uncorrected) as compared to a zero baseline: the red

upward arrow indicates significant increase in oxyHb (canonical response), the red downward significant decrease in oxyHb (inverted response), the blue asterisk a significant decrease in deoxyHb (canonical response). After correction for multiple comparisons, using the False Discovery Rate as defined by Benjamini and Hochberg (1995), none of these comparisons reached significance. The channel-wise results are therefore only suggestive.

studies. It may be that slightly different brain regions are being measured, and that our study did not pick up on areas that are lateralized to language at birth (such as the planum temporale).

However, we believe that the difference in lateralization in our study and by Pena et al. (2003) is more likely based on the stimuli used. While the current study used low-pass filtered samples of speech, Pena et al. (2003) used unfiltered speech. When low-pass filtering speech to 400 Hz, much of the segmental information, such as consonant formant transitions, is removed, while most of the prosodic information is retained. We believe that this alteration of the speech stimuli may cause neural activation that is more bilateral rather than left hemisphere dominant. Several lines of research have demonstrated that different aspects of speech are processed in different brain areas (Zatorre and Belin, 2001; Poeppel, 2003; Zatorre and Gandour, 2007). While rapid changes in speech (such as formant transitions in consonants) result in a left hemisphere bias in processing, slower changes in speech (such as prosody) result in a right hemisphere bias. This sensitivity in

brain processing has recently been evidenced in very young infants, including neonates (Homae et al., 2006; Telkemeyer et al., 2009; Minagawa-Kawai et al., 2011). We therefore suggest that filtered speech, as compared to unfiltered speech, would likely emphasize slower prosodic changes and de-emphasize faster consonant formant transitions in the speech, therefore resulting in bilateral activation. However, further research is needed to investigate this hypothesis, by directly comparing the neural response to filtered versus unfiltered speech.

A second unexpected finding in the current study was the lack of a differential brain response to forward and backward English. This finding also contrasts with the results from Pena et al. (2003) where greater left hemisphere activation was found for forward versus backward native language stimuli. Again, we propose that this result may be affected by the nature of the filtered speech used. As noted above, reversed speech is made non-linguistic in nature due to two factors: First, many of the consonants in backward speech cannot be produced by the human

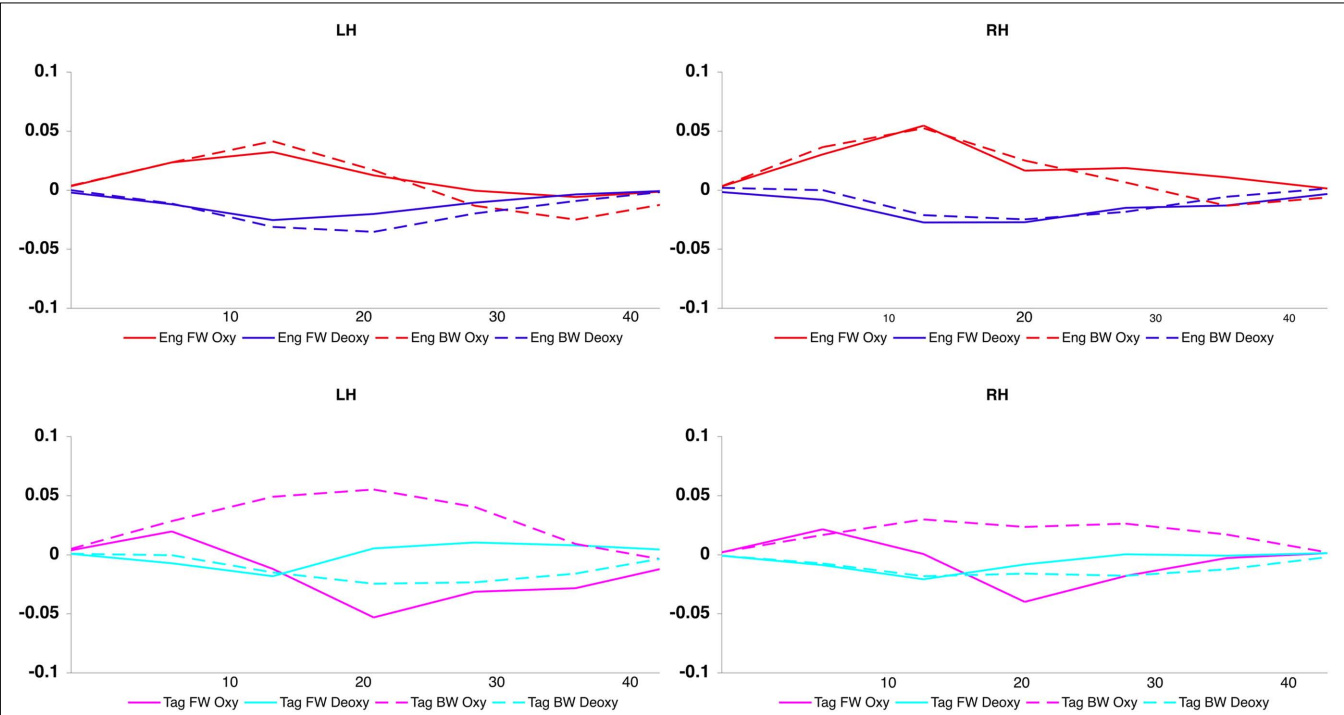
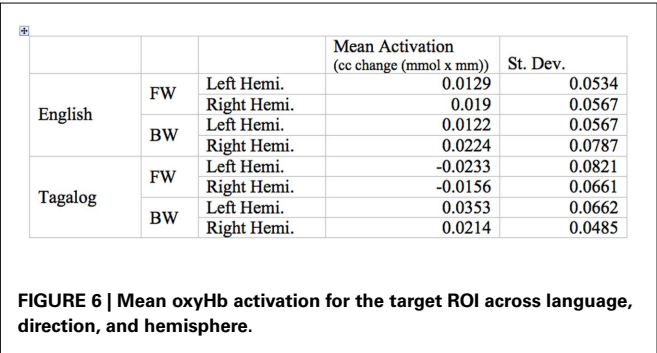


FIGURE 5 | Averaged oxy- and deoxyHb response for each condition, across hemisphere and channel.



vocal tract. Second, the prosodic structure of speech is disturbed when reversed. While the filtering likely reduces the first cue to unnaturalness, the second cue still remains in filtered backward speech. One possible *post hoc* explanation for our pattern of results is that as the filtering does maintain rhythm, neonates may be able to detect a familiar rhythmical structure in both the forward and backward English, leading to similar neural processing of both types of English language stimuli. In contrast, the rhythm of Tagalog is unfamiliar in both forward and backward forms, meaning that no familiarity might lead to similar processing of both Tagalog conditions. As infants might not be able to detect that FW and BW Tagalog are the normal and reversed versions of the same stimuli or come from the same language, there is no reason to expect a similar response to these two conditions. However, this hypothetical possibility requires further refinement and testing.

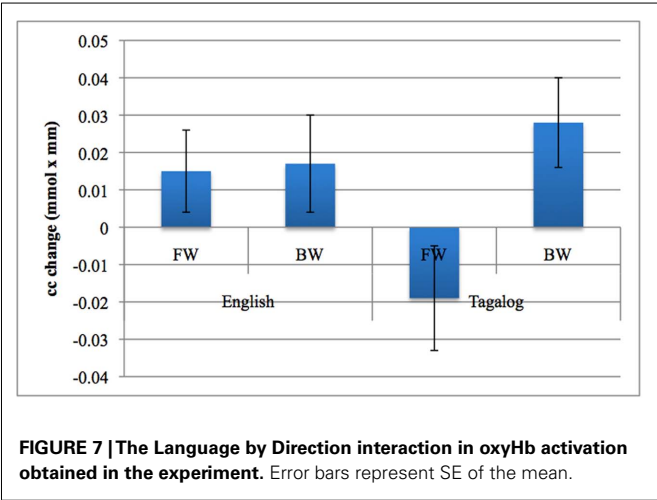


FIGURE 7 | The Language by Direction interaction in oxyHb activation obtained in the experiment. Error bars represent SE of the mean.

Thirdly, the statistical analyses revealed a negative oxyHb response to forward Tagalog. This hemodynamic response shape to forward Tagalog requires further investigation. However, what is important to note is that the size and shape of the brain response to forward Tagalog is clearly different from the shape of the response we obtained in the English conditions, further underscoring the difference between the processing of the native and a non-native language. Regardless of the basis of the differential brain response to English and Tagalog, our main finding remains that neonates showed a dissimilar pattern how the brain responds to familiar versus

familiar language. We cannot make definitive claims as to the exact nature of this processing difference on the basis of the results from the current study. Nonetheless, the results do highlight this as an area prime for future research.

Our results raise several additional questions for future study. How much prenatal language experience is sufficient to shape the neural response to language? Do premature infants show an equivalent different response to familiar and unfamiliar languages as found in the current study with full-term infants? Furthermore, if infants are raised post-natally in surroundings where prenatal and post-natal language experience differs, how might the initial brain response to language shift during development? How much post-natal experience with an unfamiliar language is needed to

alter neural activation? Finally, our evidence that prenatal language experience impacts neonates' initial neural response to language raises the question of whether and how this early neural activation might impact later language processing and learning of familiar versus unfamiliar language.

CONCLUSION

In the current study, we provide the first exploration of whether the newborn infant's neural processing of language is influenced by early language experience. We find a clear difference in how the neonate brain responds to familiar versus unfamiliar language. These results indicate that even prior to birth, the human brain is tuning to the language environment.

REFERENCES

- Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Series B Stat. Methodol.* 57, 289–300.
- Bertoncini, J., Morais, J., Bijeljac-Babic, R., McAdams, S., Peretz, I., and Mehler, J. (1989). Dichotic perception and laterality in neonates. *Brain Lang.* 37, 591–605.
- Boersma, P., and Weenink, D. (2011). *Praat: Doing Phonetics by Computer [Computer Program], Version 5.2.12*. Available at: <http://www.praat.org/> [Retrieved January 28, 2011].
- Byers-Heinlein, K., Burns, T. C., and Werker, J. F. (2010). The roots of bilingualism in newborns. *Psychol. Sci.* 21, 343–348.
- Carreiras, M., Lopez, J., Rivero, F., and Corina, D. (2005). Neural processing of a whistled language. *Nature* 433, 31–32.
- DeCasper, A. J., and Fifer, W. P. (1980). Of human bonding: newborns prefer their mother's voices. *Science* 208, 1174–1176.
- DeCasper, A. J., and Spence, M. J. (1986). Prenatal maternal speech influences newborns' perception of speech sounds. *Infant Behav. Dev.* 9, 133–150.
- Dehaene, S., Dupoux, E., Mehler, J., Cohen, L., Paulesu, E., Perani, D., van der Moortele, P., Lehericy, S., and LeBihan, D. (1997). Anatomical variability in the cortical representation of first and second language. *Neuroreport* 8, 3809–3815.
- Dehaene-Lambertz, G., Dehaene, S., and Hertz-Pannier, L. (2002). Functional neuroimaging of speech perception in infants. *Science* 298, 2013–2015.
- Dehaene-Lambertz, G., Montavont, A., Jobert, A., Alliol, L., Dubois, J., Hertz-Pannier, L., and Dehaene, S. (2010). Language or music, mother or Mozart? Structural and environmental influences on infants' language networks. *Brain Lang.* 114, 53–65.
- Dubois, J., Benders, M., Lazeyras, F., Borradori-Tolsa, C., Ha-Vinh Leuchter, R., Mangin, J. F., and Hüppi, P. S. (2010). Structural asymmetries of perisylvian regions in the preterm newborn. *Neuroimage* 52, 32–42.
- Eisenberg, R. B. (1976). *Auditory Competence in Early Life: The Roots of Communicative Behavior*. Baltimore: University Park Press.
- Gerhardt, K. J., Otto, R., Abrams, R. M., Colle, J. J., Burchfield, D. J., and Peters, A. J. M. (1992). Cochlear microphones recorded from fetal and newborn sheep. *Am. J. Otolaryngol.* 13, 226–233.
- Gervain, J., Macagno, F., Cogoi, S., Pena, M., and Mehler, J. (2008). The neonate brain detects speech structure. *Proc. Natl. Acad. Sci. U.S.A.* 105, 14222–14227.
- Gervain, J., Mehler, J., Werker, J. F., Nelson, C. A., Csibra, C., Lloyd-Fox, S., Shukla, M., and Aslin, R. N. (2011). Near-infrared spectroscopy: a report from the McDonnell infant methodology consortium. *Dev. Cogn. Neurosci.* 1, 22–46.
- Grabe, E., and Low, E. L. (2002). "Durational variability in speech and the rhythm class hypothesis," in *Papers in Laboratory Phonology*, eds C. Gussenhoven and N. Warner (Berlin: Mouton de Gruyter), 515–546.
- Holowka, S., and Petitto, L. A. (2002). Left hemisphere cerebral specialization for babies while babbling. *Science* 297, 1515.
- Homae, F., Watanabe, H., Nakano, T., and Taga, G. (2006). The right hemisphere of sleeping infant perceives sentential prosody. *Neurosci. Res.* 54, 276–280.
- Kisilevsky, B. S., Hains, S. M. J., Lee, K., Xie, X., Ye, H. H., Zhang, K., and Wang, Z. (2003). Effects of experience on fetal voice recognition. *Psychol. Sci.* 14, 220–224.
- Lecanuet, J. P., and Granier-Deferre, C. (1993). "Speech stimuli in the fetal environment," in *Developmental Neurocognition: Speech and Face Processing in the First Year of Life*, eds B. De Boysson-Bardies, S. de Schonen, P. Jusczyk, P. MacNeilage, and J. Morton (Norwell, MA: Kluwer Academic Publishing), 237–248.
- Lecanuet, J. P., Granier-Deferre, C., DeCasper, A. J., Maugeais, R., Andrieu, A. J., and Busnel, M. C. (1987). Perception and discrimination of language stimuli. Demonstration from the cardiac responsiveness. Preliminary results. *C. R. Acad. Sci. III* 305, 161–164.
- Mehler, J., Jusczyk, P. W., Lambertz, G., Halsted, N., Bertoncini, J., and Amiel-Tison, C. (1988). A precursor of language acquisition in young infants. *Cognition* 29, 143–178.
- Minagawa-Kawai, Y., van der Lely, H., Ramus, F., Sato, Y., Mazuka, R., and Dupoux, E. (2011). Optical brain imaging reveals general auditory and language-specific processing in early infant development. *Cereb. Cortex* 21, 254–261.
- Moon, C., Cooper, R. P., and Fifer, W. P. (1993). Two-day-olds prefer their native language. *Infant Behav. Dev.* 16, 495–500.
- Nazzi, T., Bertoncini, J., and Mehler, J. (1998). Language discrimination by newborns: toward an understanding of the role of rhythm. *J. Exp. Psychol. Hum. Percept. Perform.* 24, 756–766.
- Nespor, M., Shukla, M., van de Vijver, R., Avesani, C., Schraudolf, H., and Donati, C. (2008). Different phrasal prominence realization in VO and OV languages. *Lingue e Linguaggio* 7, 1–28.
- Pena, M., Maki, A., Kovacic, D., Dehaene-Lambertz, G., Koizumi, H., Bouquet, F., and Mehler, J. (2003). Sounds and silence: an optical topography study of language recognition at birth. *Proc. Natl. Acad. Sci. U.S.A.* 100, 11702–11705.
- Perani, D., Dehaene, S., Grassi, F., Cohen, L., Cappa, S. F., Dupoux, E., Fazio, F., and Mehler, J. (1996). Brain processing of native and foreign languages. *Neuroreport* 7, 2439–2444.
- Petitjean, C. (1989). *Une condition de l'audition foetale: la conduction sonore osseuse. Conséquences cliniques et applications pratiques*. MD dissertation, University of Franche-Comté, Besançon.
- Poeppel, D. (2003). The analysis of speech in different temporal integration windows: cerebral lateralization as "asymmetric sampling in time." *Speech Commun.* 41, 245–255.
- Querleu, D., Renard, X., Versyp, F., Paris-Delrue, L., and Crepin, G. (1988). Fetal hearing. *Eur. J. Obstet. Gynecol. Reprod. Biol.* 28, 191–212.
- Ramus, F., Hauser, M. D., Miller, C. T., Morris, D., and Mehler, J. (2000). Language discrimination by human newborns and cotton-top tamarin monkeys. *Science* 288, 349–351.
- Ramus, F., Nespor, M., and Mehler, J. (1999). Correlates of linguistic rhythm in the speech signal. *Cognition* 73, 265–292.
- Telkemeyer, S., Rossi, S., Koch, S. P., Nierhaus, T., Steinbrink, J., Poeppel, D., Obting, H., and Wartenburger, I. (2009). Sensitivity of newborn auditory cortex to the temporal structure of sounds. *J. Neurosci.* 29, 14726–14733.

- Vouloumanos, A., and Werker, J. F. (2007). Listening to language at birth: evidence for a bias for speech in neonates. *Dev. Sci.* 10, 159–164.
- Zatorre, R. J., and Belin, P. (2001). Spectral and temporal processing in human auditory cortex. *Cereb. Cortex* 11, 946–953.
- Zatorre, R. J., and Gandour, J. T. (2007). Neural specializations for speech and pitch: moving beyond the dichotomies. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 363, 1087–1104.
- Zimmer, E. Z., Fifer, W. P., Kim, Y. I., Rey, H. R., Chao, C. R., and Myers, M. M. (1993). Response of the premature fetus to stimulation by speech sounds. *Early Hum. Dev.* 33, 207–215.
- Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.
- Received: 19 February 2011; paper pending published: 15 April 2011; accepted: 22 August 2011; published online: 21 September 2011.
- Citation: May L, Byers-Heinlein K, Gervain J and Werker JF (2011) Language and the newborn brain: does prenatal language experience shape the neonate neural response to speech? *Front. Psychology* 2:222. doi: 10.3389/fpsyg.2011.00222
- This article was submitted to *Frontiers in Language Sciences*, a specialty of *Frontiers in Psychology*.
- Copyright ©2011 May, Byers-Heinlein, Gervain and Werker. This is an openaccess article subject to a non-exclusive license between the authors and Frontiers Media SA, which permits use, distribution and reproduction in other forums, provided the original authors and source are credited and other Frontiers conditions are complied with.



Acoustic processing of temporally modulated sounds in infants: evidence from a combined near-infrared spectroscopy and EEG study

Silke Telkemeyer^{1,2,3,4*}, Sonja Rossi^{3,5}, Till Nierhaus^{3,5}, Jens Steinbrink³, Hellmuth Obrig^{3,5†} and Isabell Wartenburger^{1,3,4†}

¹ Languages of Emotion Cluster of Excellence, Freie Universität Berlin, Berlin, Germany

² Department of Cognitive Psychology, Humboldt-Universität Berlin, Berlin, Germany

³ Berlin NeuroImaging Center, Department of Neurology, Charité University Medicine, Berlin, Germany

⁴ Department of Linguistics, University of Potsdam, Potsdam, Germany

⁵ Department of Cognitive Neurology, Max Planck Institute for Human Cognitive and Brain Sciences, University Hospital, Leipzig, Germany

Edited by:

Judit Gervain, University Paris Descartes, France

Reviewed by:

Marcela G. Pena, University of Chile, Chile

Martin Meyer, University of Zurich, Switzerland

*Correspondence:

Silke Telkemeyer, Languages of Emotion Cluster of Excellence, Freie Universität Berlin, Habelschwerdter Allee 45, 14195 Berlin, Germany.
e-mail: silke.telkemeyer@fu-berlin.de

[†]Hellmuth Obrig and Isabell Wartenburger have contributed equally to this work.

Speech perception requires rapid extraction of the linguistic content from the acoustic signal. The ability to efficiently process rapid changes in auditory information is important for decoding speech and thereby crucial during language acquisition. Investigating functional networks of speech perception in infancy might elucidate neuronal ensembles supporting perceptual abilities that gate language acquisition. Interhemispheric specializations for language have been demonstrated in infants. How these asymmetries are shaped by basic temporal acoustic properties is under debate. We recently provided evidence that newborns process non-linguistic sounds sharing temporal features with language in a differential and lateralized fashion. The present study used the same material while measuring brain responses of 6 and 3 month old infants using simultaneous recordings of electroencephalography (EEG) and near-infrared spectroscopy (NIRS). NIRS reveals that the lateralization observed in newborns remains constant over the first months of life. While fast acoustic modulations elicit bilateral neuronal activations, slow modulations lead to right-lateralized responses. Additionally, auditory-evoked potentials and oscillatory EEG responses show differential responses for fast and slow modulations indicating a sensitivity for temporal acoustic variations. Oscillatory responses reveal an effect of development, that is, 6 but not 3 month old infants show stronger theta-band desynchronization for slowly modulated sounds. Whether this developmental effect is due to increasing fine-grained perception for spectrotemporal sounds in general remains speculative. Our findings support the notion that a more general specialization for acoustic properties can be considered the basis for lateralization of speech perception. The results show that concurrent assessment of vascular based imaging and electrophysiological responses have great potential in the research on language acquisition.

Keywords: infants, speech perception, language acquisition, auditory processing, near-infrared spectroscopy, event related potentials, brain oscillations

INTRODUCTION

The analysis of acoustic features in the continuous auditory speech stream is a prerequisite for language acquisition in infancy. Among other functions it serves the segmentation of the speech stream into smaller units, like words and phrases (Mehler et al., 2004; Gervain and Mehler, 2010). This very early step of speech perception necessitates temporal and spectral differentiation of the acoustic input. In the context of speech the differentiation of the temporal structure of the acoustic input is critical, as illustrated by the clinical finding that infants with a deficit in differentiating rapidly varying auditory stimuli are more likely to develop a specific language impairment (Benasich and Tallal, 2002; Choudhury et al., 2007). The relevance of categorical acoustic feature analysis during early language acquisition is uncontroversial. However, knowledge on the underlying neuronal network and its maturation during early development is sparse. While in adults the brain clearly relies on functionally specialized areas to process speech (and language) little is known on

how this efficient network develops from birth and which “inborn” foundations endow the human brain with the unique ability to reach language competence. In newborns and 3 month old infants seminal work using functional magnetic resonance imaging (fMRI) and near-infrared spectroscopy (NIRS) demonstrated asymmetrical responses to forward compared to backward speech especially in the left angular gyrus (Dehaene-Lambertz et al., 2002, 2006; Pena et al., 2003). Additionally, a larger sensitivity of right hemispheric fronto-temporal regions in response to prosodic features has been demonstrated already in 3 month old infants (Homae et al., 2006). The dominance of the right hemispheric auditory cortex for music processing, which relies on melodic and concise pitch information (Zatorre and Belin, 2001), has been recently shown to be present from birth: A NIRS study in newborns revealed right-lateralized activation during the presentation of music excerpts (Perani et al., 2010). Taken together there is converging evidence that basic aspects of lateralization in the network in response to complex auditory

stimuli, necessary for speech and music comprehension, evolve very early. However, because very young infants clearly lack linguistic and musical knowledge, an intriguing question is how acoustic features may “guide” the lateralized processing. With regard to language, psychoacoustic models propose lateralized auditory processing as a more general basis of the lateralization in the language network. Supported by lesion and functional imaging data in adults, these models highlight that hemispheric specialization for different aspects of language processing may partially be driven by the auditory analysis. In this vein different psychoacoustic models proposed functional asymmetries based on spectral and/or temporal feature analysis (Zatorre and Belin, 2001; Poeppel, 2003; Zaehle et al., 2004; Schönwiesner et al., 2005). Though they stress different aspects of the functional anatomy of processing complex auditory stimuli they partly converge because they posit differential specializations for features of such stimuli. As an example, the multi-time resolution model (Hickok and Poeppel, 2007; Poeppel et al., 2008) postulates at least two temporal integration windows for the processing of the auditory speech input, which operate in parallel. According to the model, the integration of rapidly varying acoustic features (20–50 ms window) fundamental for the perception of phonetic contrasts, recruits areas in the left and right auditory cortices. Conversely, modulations of the acoustic signal at slower rates (150–300 ms) – more relevant for suprasegmental feature analysis (e.g., prosody) – are predominantly processed in the right hemisphere. Using noise stimuli that were modulated at different temporal rates within the predicted windows, Boemio et al. (2005) confirmed these predictions in an fMRI study in adults. In newborns a NIRS study used the same stimuli to explore whether similar lateralization can be found during a very early stage of brain development (Telkemeyer et al., 2009). The results indicate that already the newborn’s auditory cortex is sensitive to the different temporal features of the acoustic input. In particular we could show a right hemispheric lateralization for slow acoustic modulations to be present at birth. These results support the notion that basic acoustic features within the speech signal drive the hemispheric lateralization from the earliest stages of language acquisition.

Lateralization of speech perception based on its auditory features and the acquisition of linguistic competence may interact. During human development the lateralization of linguistic processing increases and consolidates (Holland et al., 2001). Experimental evidence for a successive lateralization of linguistic contrasts during early development has been supplied by longitudinal studies investigating how an initially bilateral processing of a phonemic contrast is progressively lateralized with increasing age (Minagawa-Kawai et al., 2007). Since the infant passes crucial milestones of native language acquisition in the first 6 months (Kuhl et al., 1992; Kuhl, 2004; Friederici, 2005) and changes in the underlying neuronal mechanisms have been demonstrated (Kuhl and Rivera-Gaxiola, 2008), we focused on the development of auditory feature analysis of non-linguistic stimuli in this age group.

The rationale was that processing of non-linguistic contrasts, potentially supporting the evolving lateralization of speech perception can be investigated during the maturation of the network by temporally modulated noise stimuli (Boemio et al., 2005; Telkemeyer et al., 2009). In 6 and 3 month old infants we measured the hemodynamic and electrophysiological brain responses

using simultaneous NIRS and EEG. With regard to lateralization we expected to find a pattern in both, the 3 and 6-months age-group, similar to that reported in our previous study in newborns (Telkemeyer et al., 2009). Oxygenation changes as measured by NIRS should lateralize to the right hemisphere for the slowly modulated stimuli in both age groups.

In our study in newborns we did not find differences in the electrophysiological signals simultaneously assessed (Telkemeyer et al., 2009). The transient event-related potentials (ERP) elicited by the onset of the auditory stimulus reliably showed that infants process the auditory input. It did, however, not show sensitivity for the different modulation frequencies. Also a time-frequency analysis (TFA) of the electrophysiological data did not yield a reliable effect in response to the auditory stimulus. In newborns this may be due to a discontinuous EEG (Lippe et al., 2009) which does not allow for analyses as used in data from adults (e.g., Hoogenboom et al., 2006; Koch et al., 2009). However, a recent study by Pena et al. (2010) showed that gamma-band response and ERPs evidence differential processing of linguistic input in infants. They used stimuli in the infants’ native language and stimuli in languages of rhythmically similar or grossly different classes and compared 3 and 6 month old full-term infants to 6 and 9 month old preterm infants. The results are that increased gamma-band power in response to the native language is present at 6 months in full-term but only at 9 months in preterm infants. These results elegantly corroborate the hypothesis that neuronal maturation plays a pivotal role in the earliest steps of language acquisition, most likely due to acoustic features such as rhythmic class. In the present study we report on the results of ERPs and oscillatory EEG response which nicely extend these findings to the processing of non-linguistic complex auditory material. This may be of great relevance to better understand the interaction between language acquisition and auditory feature analysis in the auditory cortex in the first 6 months of life.

With regard to the EEG data we therefore analyzed two parameters and their dependence on neuronal maturation:

- (i) We analyzed the AEP to find out how the latency and amplitude of two major components (N1, P2) change with age (Kushnerenko et al., 2002; Choudhury and Benasich, 2010). Here we expected differences in latency and amplitude of the early AEP components between 3 and 6 month old infants and were interested in whether AEPs would differ between the different modulation frequencies. Studies in adult subjects demonstrated that auditory non-speech stimuli with different grades of temporal variations modulate the evoked electrophysiological response (Zaehle et al., 2007, 2009). Unless in newborns we did not find effects of temporal modulation frequency on the AEP, the above findings (Pena et al., 2010) suggest that this may evolve between 3 and 6 months.
- (ii) We analyzed the oscillatory activity of neural assemblies, which are more sensitive to the sustained activity in response to the stimuli (Shahin et al., 2010). Brain oscillatory systems act as neural communication mechanisms, dynamically integrating signals from different brain regions and are a potential neuronal correlate of feature binding. They are also known to play a crucial role in attentional processes (Lopes da Silva, 1991; Singer, 1993, 1999). Information on the synchronization

and desynchronization between local and distant neuronal ensembles may hence inform our understanding of whether and how sustained auditory features are processed by the infant's brain. Comparing signal power of cortical oscillations in specific frequency bands after stimulus onset to a pre-stimulus (baseline) interval enables quantification of oscillatory EEG responses to the stimulus. Event-related de-/synchronization of neuronal ensembles results in decreased/increased oscillatory activity, respectively (Pfurtscheller and Lopes da Silva, 1999). This is the more relevant because the electrophysiological de/synchronization can be tonic, that is, it provides us with a measure of neuronal activity over the full duration of a stimulus. In the present study we therefore also ask whether the temporal acoustic modulations are reflected in the oscillatory electrophysiological response. In adults a simultaneous EEG and fMRI study showed that the spontaneous gamma power (28–40 Hz) correlates with activation in the left auditory cortex, while the fluctuations in the theta-range (3–6 Hz) correlate with BOLD-contrast changes in the right hemisphere (Giraud et al., 2007). These data nicely fit into the theory of parallel lateralized processing of slow and fast modulation frequencies (Poeppel et al., 2008). In infants the recent study on rhythmic classes in early language perception suggest that the analysis of oscillatory brain responses to auditory stimulation in young infants may be a promising tool to understand maturational processes (Pena et al., 2010). This approach has only rarely been used (e.g., Stefanics et al., 2007; Lippe et al., 2009) and we are aware of the explorative nature of such an analysis potentially complementing our ERP and NIRS results.

In sum, our study investigates the perception of basic auditory precursors relevant to the decoding of speech in infancy. The novel approach to combine EEG and NIRS measurements allows us to monitor the temporal and topographic aspects of neuronal processing. Using the same set-up and experimental procedure as our previous study in newborns and sharing the stimuli with an fMRI study in adults may be of relevance for a better longitudinal understanding of how auditory feature analysis of linguistic material is shaped by more basic principles of auditory perception.

MATERIALS AND METHODS

SUBJECTS

We examined two groups of infants: 6 and 3 month olds. Generally the protocol was similar to a previous study we conducted in newborns (Telkemeyer et al., 2009). We acquired informed consent from both parents. The study protocol was approved by the local ethics committee of the Charité University Medicine Berlin.

Six month old infants

In this group we measured 44 healthy infants (mean age = 185 days, ± 9.5 days; 17 boys). Their mean gestational age was 40 weeks (± 1.4 weeks), and their average birth weight was 3367 g (± 487 g). Information on familial language development and handedness was obtained from the parents. In 82% ($n = 36$) of the subjects both parents were right-handed, in 18% ($n = 8$) one parent was left-handed. In 9% ($n = 4$) one parent reported some kind of language

impairment (e.g., articulation/reading problems) during childhood. In 5% ($n = 2$) both parents reported some kind of language impairment during childhood.

Five subjects were excluded from further analysis because the experiment was ended when the infant showed signs of discomfort. Another three infants were excluded from the NIRS analysis as a result of technical problems during data acquisition. Thus, 36 infants entered the NIRS analysis.

For the EEG analysis we included all subjects in whom at least 50% of the segments survived the artifact correction procedure (see section Data Analysis). Twenty-three subjects fulfilled this criteria and were included in the final EEG analysis.

Three month old infants

In this group 40 healthy infants (mean age = 94 days, ± 10.3 days; 22 boys) were measured. They had a mean gestational age of 40 weeks (± 1.4 weeks), and their average birth weight was 3564 g (± 529 g). Information on familial language development and handedness revealed that in 80% ($n = 32$) of the subjects both parents were right-handed, and in 20% ($n = 8$) one parent was left-handed. In 7.5% ($n = 3$) one parent reported some kind of language impairment during childhood, in one case language impairments were reported by both parents.

Two infants showed signs of discomfort leading to a discontinuation of the experiment. Thus, data of 38 subjects entered the NIRS analysis. After artifact correction of the EEG data, the data of 22 infants survived our criteria (more than 50% segments after artifact correction) and entered the final EEG analysis.

STIMULI

In analogy to our previous study in newborns (Telkemeyer et al., 2009) we selected four different auditory stimuli from a larger stimulus-set published by Boemio et al. (2005). The tonal stimuli with a total duration of 9 s each are formed by concatenated noise-segments. Each noise-segment has a center frequency in the spectral range relevant for the discrimination of speech formants (1000–1500 Hz). While this spectral information is largely kept constant over the stimulus conditions, the temporal modulation of the segments was manipulated. This was achieved by modulating each noise-segment with a bandwidth of 125 Hz around its central frequency thereby yielding segments of varying length. In the present study and our previous study in newborns, segment lengths of 12, 25, 160, or 300 ms were assembled, thus forming four stimulus conditions differing in their temporal modulation. Because experimental time is limited in infants, we did not use the whole range of stimulus conditions used by Boemio et al. (2005), but selected two specific “time-windows” of temporal modulation frequencies: (1) fast (12 and 25 ms) acoustic modulations that correspond to fast such as phonetic modulations within the speech stream; (2) slow (160 and 300 ms) acoustic modulations which are associated with slow (e.g., syllabic) variations within the speech signal (Stevens, 1998). While the two fast acoustic modulations correspond to modulation frequencies of 83 and 40 Hz, the two slowly modulated stimulation patterns correspond to modulation frequencies of 6 and 3 Hz. We presented 23 stimuli per condition with variable inter-stimulus intervals (ISI) ranging from 1 to 12 s (mean 4.1 s) in a pseudo-randomized order. Please note, that the

two fast (12 and 25 ms) and the two slow (160 and 300 ms) acoustic modulations were pooled together for the data analysis (NIRS as well as EEG) to achieve a larger number of trials per condition (i.e., 46 fast modulated stimuli and 46 slowly modulated stimuli).

For audio-examples of the stimuli, please refer to Boemio et al. (2005) and Telkemeyer et al. (2009).

PROCEDURE

Throughout the experiment the infants sat on their parent's lap. To keep the experiment as transparent as possible parents were not acoustically masked. We considered undesirable influences by the parents' behavior in response to the stimuli relatively unlikely, because the stimulus material consisted of artificial, noise-like stimuli. To sustain the infants' attention a silent video of moving objects was shown temporally unrelated to the acoustic presentation. The auditory stimuli were presented via two stereo speakers (sound level of 70 dB). Stimulus presentation was controlled by Presentation software (V0.7.1, Neurobehavioral Systems). The experiment consisted of two blocks of 10 min each, separated by a variable break that could be used to interact with the infant or parent if necessary. The total duration of the experiment was approximately 20 min. The experiment was interrupted whenever the infant showed any sign of discomfort, and continued only if infant and parent were willing to further participate.

DATA ACQUISITION

Near-infrared spectroscopy

Cortical oxygenation changes in response to the auditory stimulation were assessed by NIRS. Near-infrared light ($\lambda = \sim 600\text{--}900$ nm) penetrates biological tissue up to several centimeters depth, reaching the cerebral cortex, when applied on the head. Models of the neuro-vascular coupling (Fox and Raichle, 1986) predict that increases in neuronal activation lead to an increase in regional cerebral blood flow overcompensating the local demand in oxygen. This results in a focal cortical hyperoxygenation which translates into an increase in oxygenated hemoglobin (oxy-Hb) and a decrease in deoxygenated hemoglobin (deoxy-Hb) concentrations. It should be noted that an increase in regional cerebral blood volume and an increase in blood flow velocity is expected in an activated cortical area (Fox and Raichle, 1986). With regard to the NIRS parameters that translates into an increase in oxy-Hb and a decrease in deoxy-Hb (Obrig and Villringer, 2003). The debate on which parameter is more "powerful" is on sensitivity and specificity. Sensitivity is larger for oxy-Hb, partially due to the larger amplitude. Specificity, however, is larger for deoxy-Hb, because an increased washout of deoxy-Hb in an activated area can be considered a specific feature of the cerebral hemodynamic response as opposed to changes in hemodynamics in the extracerebral tissue (Boden et al., 2007). NIRS unfortunately is extremely sensitive to changes in extracerebral hemodynamic changes. Therefore we advocate to always report deoxy-Hb changes also, because deoxy-Hb decreases are the major source of the BOLD-contrast (Steinbrink et al., 2006). The matter is complicated by the debate on the "typical" response pattern in infants, for a more detailed discussion see Obrig et al. (2010). We here report both increases in oxy-Hb and decreases in deoxy-Hb.

Technically light of two different wavelengths is guided to and from the subject's head by fiber-optic bundles. Detector probes are placed pairwise some 2–3 cm from the emitting probes to collect the reflected light. Each source-detector pair defines a sampling volume. Focal changes in oxy- and deoxy-Hb are derived from the changes in attenuation measured at two wavelengths, based on the modified Beer–Lambert law (Cope and Delpy, 1988). Event-related decreases in deoxy-Hb correlate well with BOLD-contrast increases, termed "activation" in the fMRI literature (Kleinschmidt et al., 1996; Obrig and Villringer, 2003).

We used a NIRS system (Omniat Tissue Oxymeter, ISS, USA) consisting of four light detectors and eight light emitters. The instrument works with modulated light sources at 690 and 830 nm. Raw data were sampled at a rate of 10 Hz. All optical probes and the EEG electrodes were integrated into an EEG cap (EASYCAP, Germany). Emitter and detector probes for the NIRS measurement were separated by an interprobe distance of 2.5 cm. The NIRS array resulted in 6 measurement volumes over each hemisphere (see **Figure 1**): (1) inferior frontal, (2) superior frontal, (3) inferior temporal, (4) superior temporal, (5) posterior temporal, and (6) temporo-parietal. The probe placement paralleled the EEG electrode placement and partially corresponded to positions of the 10–20 system (Sharbrough et al., 1991).

Electroencephalography

Electroencephalography was recorded with 17 Ag/AgCl electrodes (Brainproducts, Germany) also mounted with the elastic EEG cap (EASYCAP, Germany). Electrodes were located according to the 10–20 system (Sharbrough et al., 1991) at the following positions: F3, F4, C3, C4, P3, P4, F7, F8, T7, T8, F9, F10, Fp1, Fp2, Fz, Cz, Pz, online-referenced against the left mastoid, with the AFz as ground electrode (see **Figure 1**). The EEG signal was recorded with a sampling rate of 1000 Hz and digitized online from 0.53 to 120 Hz.

DATA ANALYSIS

Near-infrared spectroscopy

Attenuation changes at 690 and 830 nm were converted into concentration changes of oxy- and deoxy-Hb using the modified Lambert–Beer law (Cope and Delpy, 1988). Data were low-pass filtered at 0.3 Hz (Butterworth, third order) and additionally high-pass filtered at 0.03 Hz to correct for high-frequency noise and slow drifts and fluctuations. Attenuation of movement artifacts is of special relevance in data recorded in infants. In line with previous infant studies using the same methodology (Taga et al., 2003; Minagawa-Kawai et al., 2011) we detected motion-induced artifacts characterized by sudden and sharp signal changes through visual inspection of the data. Artifacts were digitally marked and replaced by linear interpolation of uncontaminated data-points (10 data-points before and after the artifact), thus avoiding exclusion of whole segments or even whole data-sets. Next, the concentration changes of oxy- and deoxy-Hb were analyzed using a general linear model (GLM) approach. To increase the number of trials per condition the two fast modulated stimulus conditions were pooled (12 and 25 ms) as well as the two slowly modulated stimulus conditions (160 and 300 ms). Thus, the design matrix included two boxcar functions with the stimulus duration of 9 s relative to the onset of each stimulus modeling the pseudo-randomized succession of the fast and slowly modulated stimuli. These predictors were convolved

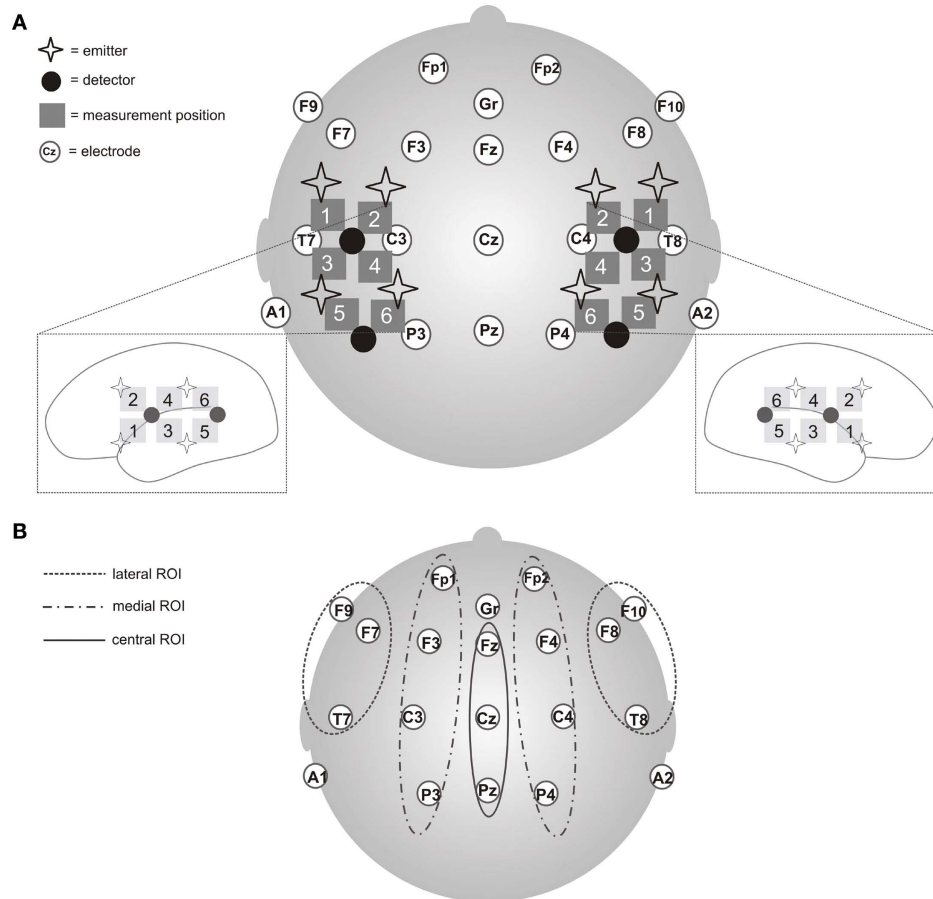


FIGURE 1 | (A) Details of the combined EEG and near-infrared spectroscopy setup. EEG was recorded from 17 scalp positions according to the international 10–20 system (Gr = ground, A1/A2 = reference). The six measurement positions per hemisphere for assessing the vascular response by near-infrared spectroscopy are represented by all available emitter-detector pairs

(= measurement position): (1) inferior frontal; (2) superior frontal; (3) inferior temporal; (4) superior temporal; (5) posterior temporal; (6) temporo-parietal.

(B) Regions of Interest (ROIs) for the EEG analyses: left-medial: Fp1/F3/C3/P3, right-medial: Fp2/F4/C4/P4, left-lateral: F7/T7/F9, right-lateral: F8/T8/F10, central: Fz/Cz/Pz.

with the canonical hemodynamic response function (Boynton et al., 1996). The GLM analysis yields beta-values for oxy- and deoxy-Hb for the two stimulus conditions. The contrast between conditions and the *post hoc* statistical analyses (resulting in *t*-values) were performed in analogy to “Statistical Parametric Mapping,” as used for fMRI data. Paired *t*-tests were performed between left and right channels for fast and slow modulations, for each age group and for oxy- and deoxy-Hb separately.

Electroencephalography

Off-line analyses were performed using Brain Vision Analyzer 2.0. Data were filtered off-line at 0.53 Hz low cutoff, 70 Hz high cutoff, and a 50 Hz notch filter (bandwidth 5 Hz, 24 dB/octave) was applied to attenuate line-voltage artifacts. We re-referenced the data to the averaged left and right mastoids. After the filtering procedure very noisy electrode channels were rejected. These were channels that showed either a flat line or signals stemming from predominantly technical artifacts throughout the whole experiment. Data were segmented into units of 10 s (1 s pre-stimulus onset, 9 s post-stimulus onset). We then applied a semi-automated artifact cor-

rection procedure (Brain Vision Analyzer 2.0). First, each data-set is automatically scanned for segments with maximal voltage step of 50 μ V/ms, and maximal absolute differences of 200 μ V. To ensure the quality of this automated procedure, each segment was again checked, and excluded manually if necessary.

Only participants for whom a minimum of 50% of the trials survived the artifact correction were included in the further EEG analyses. In the 23 subjects included in the analysis of the 6 month olds an average of $30.2 \pm 12.7\%$ of the trials were removed by the artifact correction procedure (fast condition: mean = 32.2%, SD = 14.8%; slow condition: mean = 29.2%, SD = 13.8%). In 3 month olds 22 subjects were included in whom an average of 23.3% (SD = 15.8%) of the trials were removed (fast condition mean = 25.0%, SD = 16.0%; slow condition: mean = 25.3%, SD = 18.6%). A repeated measures analyses of variance (ANOVA) with the factor *condition* (fast versus slow acoustic variations) and *age group* as between-subject factor was performed to assess whether the amount of excluded segments differed across conditions and age groups. The ANOVA did not reveal a significant effect of condition [$F(1,43) = 1.85$; $p = 0.18$], and of condition \times age group [$F(1,43) = 2.84$; $p = 0.10$].

Analysis of auditory-evoked potentials. Auditory-evoked potentials (AEP) upon stimulus onset were computed for each participant and each experimental condition by averaging 1000 ms after stimulus onset referenced to a 100 ms pre-stimulus baseline. We were interested in developmental effects on the general features of the AEPs but also on specific effects of fast and slow acoustic modulations on the AEPs. Therefore we conducted three different AEP analyses:

(1) Analysis of general features of the AEPs in both age groups

To investigate how AEPs change with age we computed averaged AEPs across all stimulus conditions for the 6 and 3 month olds separately and performed peak-latency and -amplitude analyses. Previous studies suggest that in infants the most prominent components of the AEP occur within a time window of approximately 500 ms after stimulus onset (e.g., Kushnerenko et al., 2002; Wunderlich et al., 2006). To identify the average peak amplitudes and latencies of the AEP components we therefore ran a peak analysis for the first 500 ms after stimulus onset (Brain Vision Analyzer 2.0). Two peaks were identified within the first 500 ms after stimulus onset, and peak amplitude as well as peak latency for each electrode and each participant of the two age groups were assessed. To evaluate group differences in amplitude and latency we conducted univariate ANOVA separately for latency and amplitude of each peak, using *age group* as between-subject factor. The following electrodes entered statistical analysis subdivided into five regions of interest (ROIs): left-medial: Fp1/F3/C3/P3, right-medial: Fp2/F4/C4/P4, left-lateral: F7/T7/F9, right-lateral: F8/T8/F10, central: Fz/Cz/Pz.

(2) Analyses on mean amplitudes of fast and slow acoustic modulations

Next we analyzed differences of the AEP components with regard to the different stimulus conditions (fast and slow acoustic modulations). To increase the signal to noise ratio the two fast modulations (12 and 25 ms) were pooled and compared to the pooled slowly modulated conditions (160 and 300 ms). To identify peaks of the components the general peak-latency analysis yielded values which are in line with the literature (Kushnerenko et al., 2002; Picton and Taylor, 2007; Lippe et al., 2009). The following time windows were analyzed: for the 6 month olds: 0–100 ms (N1), and 100–225 ms (P2). Due to longer latencies in the younger age group, in the 3 month olds different windows were used: 0–200 ms (N1), and 200–500 ms (P2). Because the time windows differ between 6 and 3 month olds, separate analyses on mean amplitudes were performed for the two age groups. The analysis was performed in the previously specified ROIs (*left-medial, right-medial, left-lateral, right-lateral and central*). The repeated measures ANOVAs tested the within-subject factors *condition* (fast versus slow), and *hemisphere* (left versus right) the latter including *left-medial* versus *right-medial*, and *left-lateral* versus *right-lateral* ROIs. For the central ROI we calculated repeated measures ANOVA with the factor *condition*. When an ANOVA revealed a significant ($p \leq 0.05$) main effect or interactions between either *condition* and/or *hemisphere*, *post hoc* paired *t*-tests were calculated between the next levels of the respective factor. Greenhouse and Geisser (1959) corrected significances are reported.

(3) Analyses on peak amplitudes of fast and slow acoustic modulations

To test for significant differences between fast and slow modulation frequencies we additionally performed *peak* amplitude analyses (Rossi et al., 2010) of the AEP-components. Analysis of *peak* amplitude was performed because: (i) the analysis on *mean* amplitudes did not reveal a significant effect for the P2; (ii) *mean* amplitudes cannot be compared between age groups since the lengths of the time-windows differed between age groups. Statistical analyses of the *peak* amplitudes were performed on the same ROIs following the same schema reported for *mean* amplitude analyses above. However, we now extended the ANOVA with the between-subject factor *age group*.

Time-frequency analysis. To investigate tonic differences in the electrophysiological response between conditions and age groups we performed TFA to reveal stimulus-induced changes in oscillatory brain activity in the frequency range from 4 to 70 Hz. Artifact corrected EEG data (see above) were downsampled to 500 Hz. Further analysis was performed using custom-built Matlab scripts (version R2007a, Mathworks, Natick, MA, USA). For calculating the time-frequency representations from 4 to 70 Hz we used segments from –900 to 9000 ms relative to stimulus onset, and performed wavelet analyses (Morlet wavelet) on each trial (Tallon-Baudry and Bertrand, 1999; Jensen et al., 2002). Baseline power was calculated in a 850 ms pre-stimulus interval (–900 to –50 ms prior to stimulus onset to avoid stimulus related contamination of the baseline by smearing effects). One possibility to quantify oscillatory EEG responses is to assess the relative increase or decrease in signal power of cortical oscillations in specific frequency bands in an interval after stimulus onset compared to a pre-stimulus interval. Thereby a resulting event-related synchronization or desynchronization quantifies changes in signal power relative to the event (Pfurtscheller and Lopes da Silva, 1999). Therefore we averaged the time-frequency representations across the trials of the two fast modulated stimuli (12 and 25 ms), and the two slow acoustic modulations (160 and 300 ms), and displayed relative changes to the baseline. Fast and slow conditions comprised up to a maximum of 46 trials for both conditions in each infant. Relative changes were averaged over frequency and time. The frequency windows for the computation of the different frequency bands were chosen according to literature (Pfurtscheller and Lopes da Silva, 1999; Nierhaus et al., 2009). **Figure 5** shows de-/synchronization in the frequency bands from 4–8, and 10–15 Hz from 500 to 9000 ms after stimulus onset. For statistical analysis we computed mean values across the time-frequency windows: 4–8 Hz; 500–8900 ms and 10–15 Hz 500–8900 ms. The mean values entered a repeated measures ANOVA with the within-subject factors *condition* and *hemisphere* and the between-subject factor *age group*, performed in the above defined ROIs.

RESULTS

NEAR-INFRARED SPECTROSCOPY

The GLM based on the oxygenation changes yielded β -values of changes in oxy- and deoxy-Hb for fast (12 and 25 ms) and slow (160 and 300 ms) acoustic modulations. To assess lateralization they

were compared by paired *t*-tests between hemispheres. **Figure 2** illustrates in which areas oxy- and/or deoxy-Hb responses showed significant lateralization ($p \leq 0.05$). The upper panel illustrates the results in the 6 month olds the lower those in 3 month olds.

In 6 month olds ($n = 36$) fast acoustic modulations lead to larger hemodynamic responses (oxy-Hb \uparrow and deoxy-Hb \downarrow) over the left compared to right inferior temporal position [position (3): deoxy-Hb: $t_{(35)} = -2.37$, $p = 0.012$; oxy-Hb: $t_{(35)} = 2.26$, $p = 0.015$]. Fast acoustic modulations additionally elicited a larger hemodynamic response (deoxy-Hb \downarrow) in the right compared to left temporo-parietal region [position (6): deoxy-Hb: $t_{(35)} = 1.83$, $p = 0.038$].

For the slow acoustic modulations statistics confirmed a larger hemodynamic response (deoxy-Hb \downarrow) in two right hemispheric positions [position (1): deoxy-Hb: $t_{(35)} = 1.88$, $p = 0.034$; position (6): deoxy-Hb: $t_{(35)} = 1.89$, $p = 0.033$] in inferior frontal and temporo-parietal regions.

The NIRS results for the 3 month olds ($n = 38$) are illustrated in the lower panel of **Figure 2**. In this age group we found a larger increase in oxy-Hb for left compared to right hemispheric brain regions: Left superior frontal and posterior temporal regions

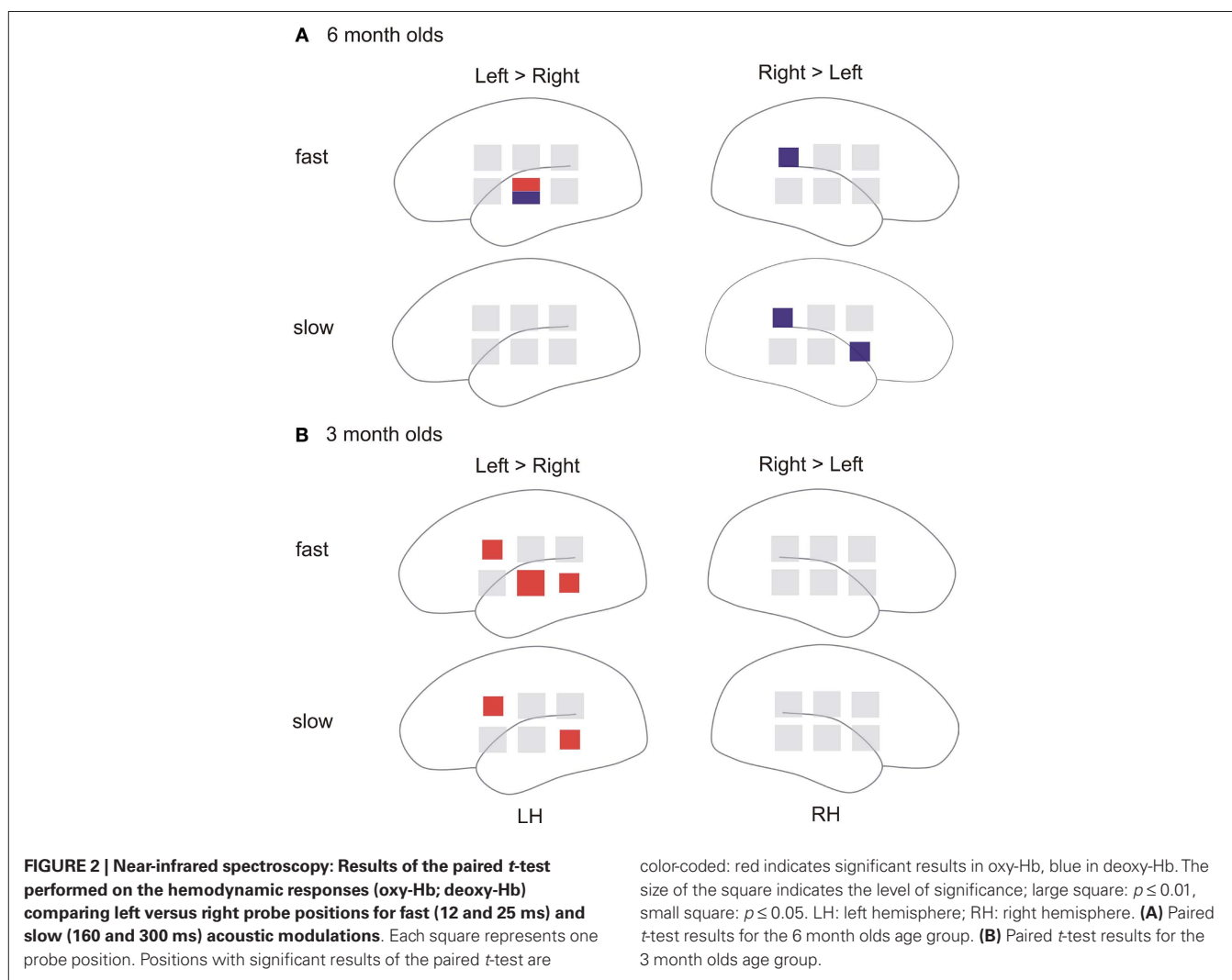
showed increased responses for both, fast [position (2); oxy-Hb: $t_{(37)} = 1.83$, $p = 0.038$; position (5); oxy-Hb: $t_{(37)} = 1.91$, $p = 0.032$], and slow [position (2); oxy-Hb: $t_{(37)} = 1.72$, $p = 0.047$; position (5); oxy-Hb: $t_{(37)} = 1.96$, $p = 0.029$] acoustic modulations. For the fast acoustic modulations we additionally found a stronger response in the left inferior temporal position [position (3); oxy-Hb: $t_{(37)} = 2.47$, $p = 0.009$].

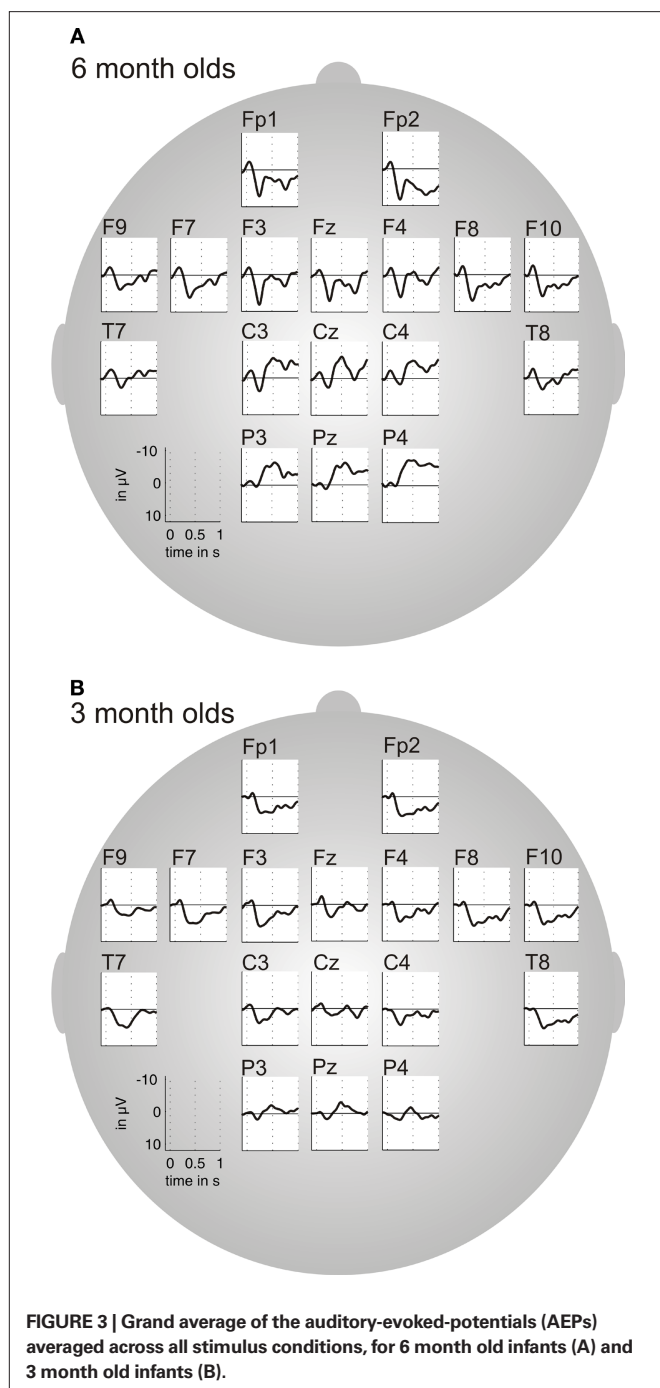
ELECTROENCEPHALOGRAPHY

The analysis of the EEG data focused on two properties. First we report the results concerning the evoked potentials upon onset of the stimulus periods (AEPs representing the phasic response). To assess the response over the full length of the stimulation period we next report the results of the TFA for two frequency bands at 4–8 and 10–15 Hz (tonic response).

Auditory-evoked potentials

General features of the AEPs in both age groups. To reveal a general effect of maturation of the AEPs, we first calculated the averaged AEPs across all stimulus conditions. **Figure 3** illustrates the





results for 6 month old infants ($n = 23$) and 3 month old infants ($n = 22$) separately. We performed peak-latency analyses on the AEPs between 0 and 500 ms, which revealed a first peak with a negative polarity (N1) followed by a second component with a positive polarity (P2) in the AEPs of both age groups.

In 6 month olds, the N1 peaked at 58 ms on average (range 25–93 ms, $SD = 21$ ms). This component showed the same latency in the 3 month olds (mean 59 ms, range 27–98 ms, $SD = 20$ ms). The univariate ANOVA confirmed that there was no statistically significant difference for *peak latency* in any of the ROIs. On the contrary the *amplitude* of the N1 was larger in 6 compared to 3 month old

infants, which was confirmed by the univariate ANOVA for *peak amplitude* over the *left-lateral ROI*: $F(1,42) = 7.32$, $p < 0.01$, and *right-lateral ROI*: $F(1,43) = 6.95$, $p < 0.01$. The mean amplitude in the 6 month olds group was $-4.1 \mu V$ (range -13.6 to $1.74 \mu V$, $SD = 3.2 \mu V$). In 3 month olds the mean amplitude of the N1 was $-2.4 \mu V$ (range -7.3 to $2.75 \mu V$, $SD = 2.3 \mu V$).

The P2 is clearly visible in the grand averages of both age groups (Figure 3). In 6 month old infants the P2 peaks at 226 ms on average (range 153–277 ms, $SD = 35$ ms), while in 3 month olds the peak occurs later at around 315 ms (range 154–453 ms, $SD = 80$ ms). The univariate ANOVA on *peak latency* of the P2 revealed significant differences between age groups for all ROIs (*left-medial*: $F(1,43) = 19.97$, $p < 0.001$; *right-medial*: $F(1,43) = 14.27$, $p < 0.001$; *left-lateral*: $F(1,42) = 22.61$, $p < 0.001$; *right-lateral*: $F(1,43) = 18.09$, $p < 0.001$; *central*: $F(1,43) = 17.02$, $p < 0.001$). With regard to the *peak amplitude* of the P2 there was no difference between age groups.

In sum, N1 and P2 components were seen in the AEPs of both age groups. The N1 peaks around 60 ms in both age groups and increases in amplitude with age over bilateral fronto-temporal regions. The P2, on the contrary, decreases in latency with age over all regions but does not change in amplitude.

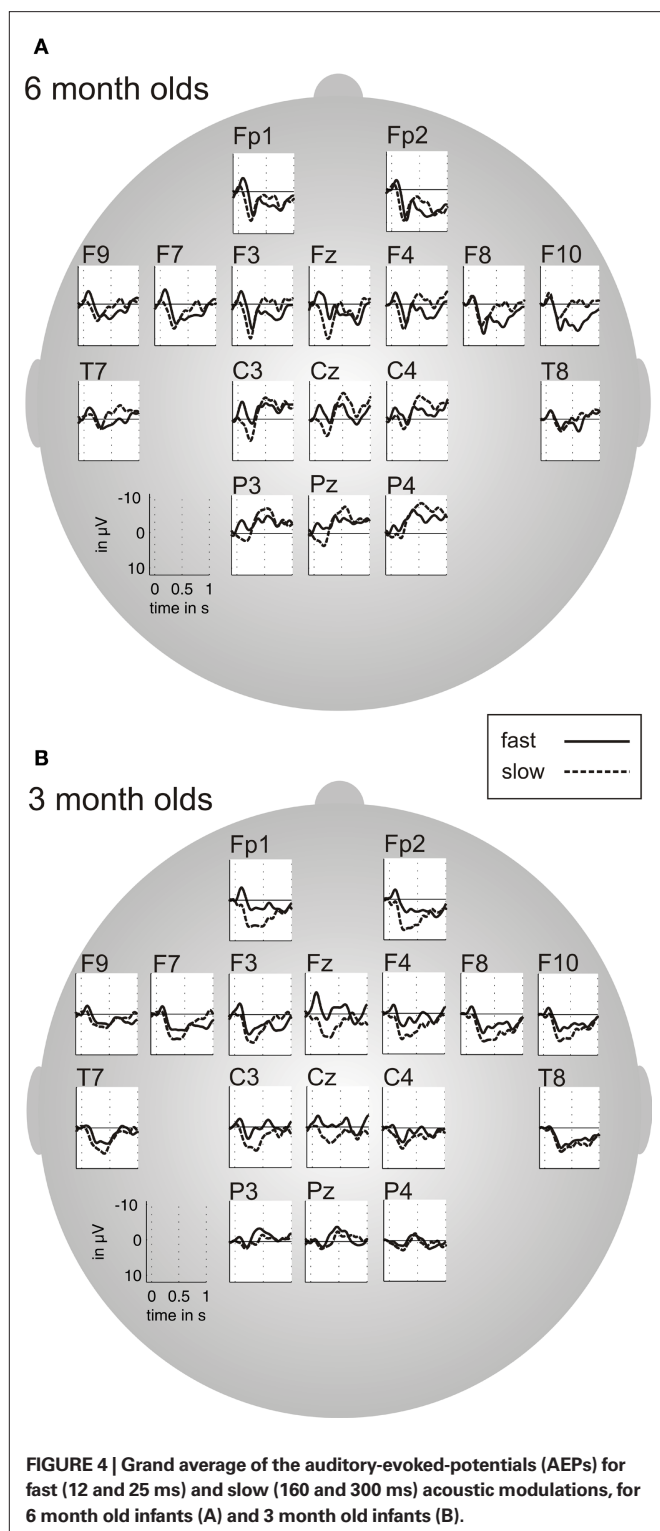
Analyses on mean amplitudes of fast and slow acoustic modulations. To test whether fast and slow acoustic modulations elicit differential phasic electrophysiological responses we computed AEPs separately for fast and slow acoustic modulations. Figure 4 shows the results separately for the two different age groups. N1- and P2-component are clearly seen in all conditions.

In 6 month olds the ANOVA for the N1-window (0–100 ms) reveal a significant effect of the factor *condition* only. Therefore we averaged the respective ROI pairs for the paired *t*-tests to compare fast and slow modulations. In the *medial ROI* we found a larger mean amplitude of the N1 for fast compared to slow acoustic modulations ($F(1,22) = 4.67$, $p < 0.04$; $t_{(22)} = -2.16$, $p = 0.04$). In the 3 month olds the ANOVA for the N1-window (0–200 ms) revealed a trend for the main effect *condition* over the *central ROI* ($F(1,21) = 4.04$, $p < 0.057$). Here N1 was larger in amplitude for fast in contrast to slow modulations. The effect was most pronounced over *Fz* (see Figure 4). Separate paired *t*-tests for each of the three midline electrodes confirmed a significantly larger N1 for fast compared to slow acoustic modulations ($t_{(18)} = -3.65$, $p = 0.002$) over *Fz*.

The analysis on the mean amplitude of the P2 (100–225 ms in the 6 months olds and 200–500 ms in the 3 months olds) for fast versus slow stimuli did not yield any statistically significant effects.

In sum, the mean amplitude analysis yielded significant differences between the two conditions only for the N1. In both age groups the N1 was larger for the onset of fast compared to slowly modulated stimuli over bilateral fronto-central ROIs (please also refer to Table 1 for an overview of the results).

Analyses on peak amplitudes of fast and slow acoustic modulations. We additionally performed statistical analyses on *peak amplitudes* for N1 and P2. Both peaks (N1 and P2) were identified by an automatic peak detection (see section General Features of the AEPs in Both Age Groups). The within-subject factors *condition*, *hemisphere*, and the between-subject factor *age group* were tested by



repeated measures ANOVAs for the medial and lateral ROIs. For the analysis of the *central* ROI an ANOVA with the within-group factor *condition* and the between-subject factor *age group* was computed. Neither the ANOVA for the peak amplitude of the N1, nor for the P2 did reveal any effect of the between-subject factor *age group*. Therefore we averaged across the two age groups for *post hoc* paired *t*-tests.

For the N1, the ANOVA revealed a significant main effect of *condition* in the *medial*: $F(1,43) = 6.01$, $p < 0.02$, and *central* ROI: $F(1,43) = 4.95$, $p < 0.03$. The *post hoc* paired *t*-test for the averaged left and right *medial* ROI revealed a significantly larger N1 for fast compared to slow acoustic modulations ($t_{(44)} = -2.48$, $p = 0.02$). The same effect was seen for the *central* ROI ($t_{(44)} = -2.24$, $p = 0.03$).

With regard to the P2 the ANOVA also revealed significant main effects of *condition* for the *medial* ($F(1,43) = 4.35$, $p < 0.04$), and *central* ROI: $F(1,43) = 5.88$, $p < 0.02$. The *post hoc* paired *t*-test for the *medial* ROI revealed a significantly larger P2 for slow acoustic modulations ($t_{(44)} = -2.09$, $p = 0.04$), which also held true for the *central* ROI ($t_{(44)} = -2.45$, $p = 0.02$).

In summary, the differential peak analyses of the AEPs for fast and slow acoustic modulations showed that the amplitude of the N1 was larger for fast compared to slow modulation frequencies. On the contrary the amplitude of the P2 was larger for slow when compared to fast acoustic modulations. These effects did not differ between age groups (please also see, **Table 1** for an overview).

Time-frequency analysis

The AEP-analysis reported so far is sensitive only to the onset of the stimuli. To find out whether the differential stimulus features (slow versus fast modulations) elicit a sustained response over the full stimulation period we performed a TFA on the EEG data. Sustained differential synchronizations and desynchronizations have been reported in response to stimulus features in a number of systems in adults (e.g., gamma-synchronization and alpha-desynchronization in the visual system (Koch et al., 2009). Since there are very few reports on TFA in very young infants (e.g., Csibra et al., 2000; Pena et al., 2010) this analysis was explorative in nature and we could not make strong predictions to whether de- or synchronizations were to be expected and in which frequency bands such modulations should be seen. Therefore a TFA over a wide spectral range (4–70 Hz) was performed separately for fast and slow acoustic modulations and in all five ROIs: left-medial: Fp1/F3/C3/P3, right-medial: Fp2/F4/C4/P4, left-lateral: F7/T7/F9, right-lateral: F8/T8/F10, central: Fz/Cz/Pz.

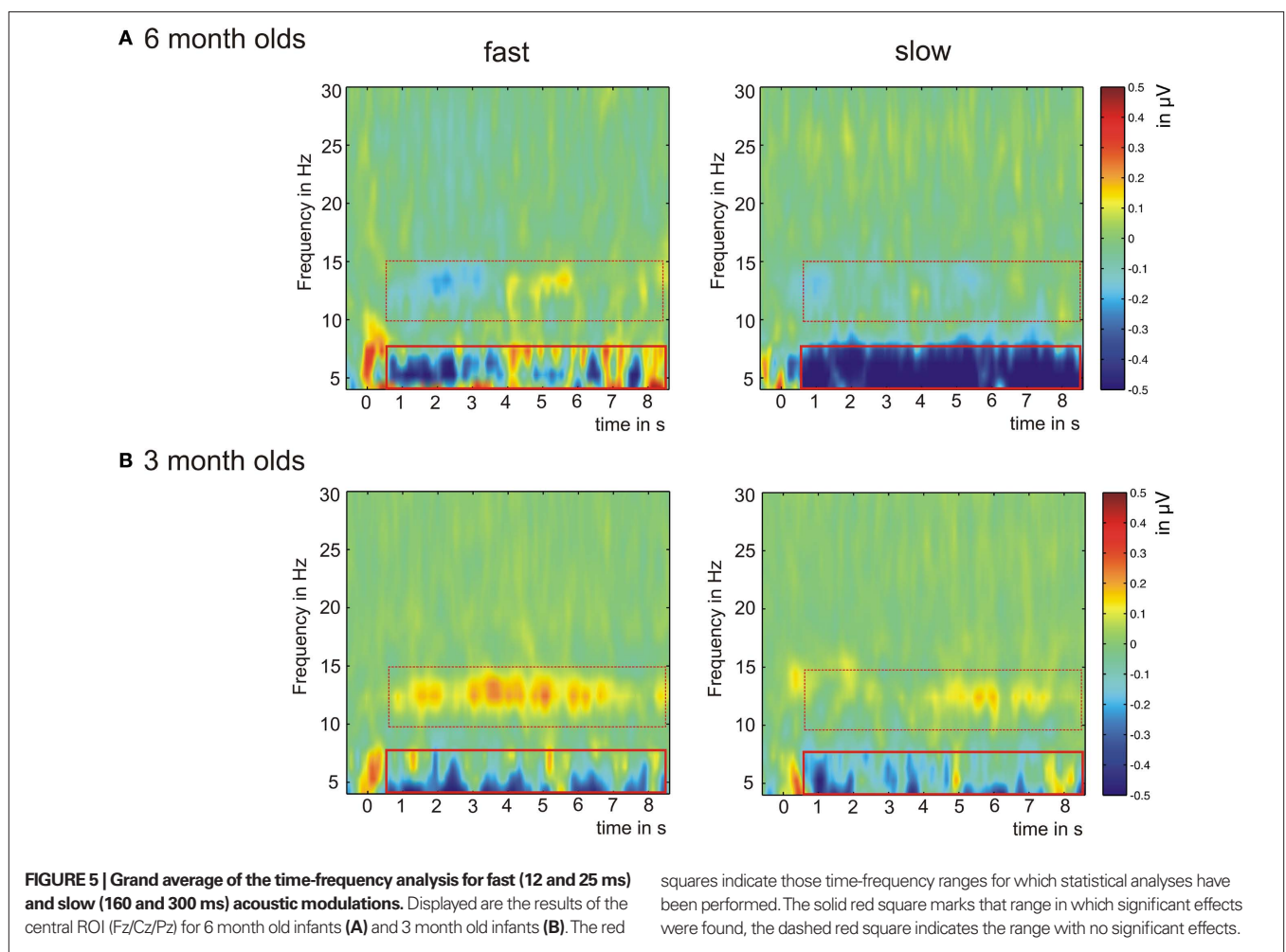
Figure 5 shows modulations in two frequency bands (averaged across all ROIs), that is, in the theta (4–8 Hz) and alpha-range (10–15 Hz).

Fast and slow acoustic modulations elicited synchronization in the theta-range during the first 300 ms after stimulus onset in both age-groups. Furthermore, ~500 ms after stimulus onset 6 month olds showed a sustained desynchronization in response to the slowly varying stimuli in the theta-range. This desynchronization was less pronounced for the fast modulated stimuli. The 3 month olds did not show this effect. A similar but weaker desynchronization was seen for both conditions in this younger age group. With regard to the higher frequency band (10–15 Hz) only the 3 month olds showed a difference between the conditions. In this frequency band fast modulated stimuli elicited a stronger synchronization when compared to the slowly modulated stimuli. In the 6 month olds a small and unstable desynchronization was seen in this higher frequency range. In all higher frequency bands (including gamma) no de/synchronization was seen in the time-frequency plots, and statistical analysis confirmed this result.

Table 1 | Overview of the statistically significant EEG and NIRS results.

		AEP				TFA	NIRS
		Mean amplitude		Peak amplitude		Desynchronization (4–8 Hz)	Activation (oxy and/or deoxy)
		N1	P2	N1	P2		
6 mo	Fast	+	=	+	–	–	LH > RH (3) RH > LH (6)
	Slow	–	=	–	+	+	RH > LH (1, 6)
3 mo	Fast	+	=	+	–	–	LH > RH (2, 3, 5)
	Slow	–	=	–	+	–	LH > RH (2, 5)

Displays the comparison of fast and slow acoustic modulations for the two age groups separately. 6 mo, 6 month old infants; 3 mo, 3 month old infants; AEP, auditory-evoked potential, TFA, time-frequency analysis; NIRS, near-infrared spectroscopy; N1, first negativity; P2, second positivity; LH, left hemisphere; RH, right hemisphere; (+), increase; (-), decrease; (=), no difference. Location of the areas measured by means of NIRS are described by the number in brackets (please also refer to Figure 1): (1) inferior frontal, (2) superior frontal, (3) inferior temporal, (4) superior temporal, (5) posterior temporal, and (6) temporo-parietal.



For statistical analysis of the two lower frequency bands we averaged the power changes in the theta-range (4–8 Hz) and the alpha range (10–15 Hz) from 500 to 8900 ms. We chose this time window because early synchronization effects during the first hundreds of milliseconds after stimulus onset are likely due to the evoked response

(see Materials and Methods). The resulting changes in oscillatory amplitude in each ROI entered a repeated measures ANOVA with the within-subject factors *condition*, and *hemisphere*, and *age group* as between-subject factor. The *central* ROI was analyzed using a repeated measures ANOVA with the factors *condition* and *age group*.

The theta desynchronization yielded a significant interaction of *condition* \times *age group* in every ROI: *left- and right-medial*: $F(1,43) = 8.09, p < 0.007$, *left- and right-lateral*: $F(1,42) = 6.37, p < 0.02$, and *central*: $F(1,43) = 7.68, p < 0.008$. Furthermore, we found a significant main effect of *condition*: *left- and right-medial*: $F(1,43) = 6.13, p < 0.02$, *left- and right-lateral*: $F(1,42) = 5.85, p < 0.02$, and *central*: $F(1,43) = 8.76, p < 0.005$. Based on the results, we computed *post hoc* paired *t*-tests for the two age groups separately, to compare the oscillatory activity for fast and slow acoustic modulations. **Figure 5** shows the results of the TFA analysis for the central ROI exemplarily.

In 6 month olds the paired *t*-tests showed that slow acoustic modulations elicited stronger desynchronization in the theta-range when compared to fast modulations (*lateral*: $t_{(22)} = 3.11, p = 0.005$; *medial*: $t_{(22)} = 2.91, p = 0.008$; *central*: $t_{(22)} = 3.54, p = 0.002$). For the 3 month olds the paired *t*-tests did not reveal significant differences between fast and slow acoustic modulations in the theta-range. In the alpha-range significant effects were found in neither age group and none of the ROIs.

To summarize, slowly modulated stimuli elicit a sustained desynchronization in the theta-range (4–8 Hz) in 6 month old infants. This desynchronization is statistically larger than the response to the fast modulated stimuli. The effect was not seen in the younger infants (please see, **Table 1** for an overview of the results).

DISCUSSION

HEMODYNAMIC RESPONSES

Our results show that subtle auditory differences during the processing of complex auditory stimuli elicit a differential pattern of brain activation in infants. NIRS revealed a lateralized brain response for 6 month old infants, similar to the reported findings newborns (Telkemeyer et al., 2009) and in adults (Boemio et al., 2005). Fast acoustic modulations (12 and 25 ms) lead to an activation of bilateral temporal brain regions. On the contrary slow acoustic modulations (160 and 300 ms) resulted in a greater right-lateralized hemodynamic response in the temporal region. These results are in line with the assumptions of the multi-time resolution model linking hemispheric specialization for language features to an asymmetry in cortical tuning (Hickok and Poeppel, 2007; Poeppel et al., 2008). The model proposes that hemispheric lateralization during language perception partially results from the temporal features in the speech signal. Thereby, left and right auditory cortices are differentially specialized for the acoustic analysis in at least two different temporal integration windows (Poeppel, 2003; Poeppel et al., 2008). Bilateral auditory cortex areas are thought to decode fast acoustic modulations, specifically relevant for the decoding of segmental such as phonological information within the speech stream. Slow acoustic modulations, relevant for the perception of suprasegmental language features, like prosodic information, are mainly processed in right hemispheric cortical brain regions. A recent NIRS study in 4 month old infants comparing different speech and non-speech conditions observed right hemispheric activation for slowly modulated emotional voices, whereas speech sounds and scrambled non-speech sounds, both comprising fast acoustic variations, elicited leftward activation (Minagawa-Kawai et al., 2011). Similar to our results in newborns and 6 month

olds, the authors conclude that the observed lateralization might be driven by basic acoustic features. Interestingly, their results also emphasize the influence of linguistic features *per se* (i.e., exposure to the native language) on the modulation of cortical brain responses, because they found stronger left-hemispheric activation during native compared to non-native speech sounds.

It should be noted, that we failed to confirm the right hemispheric specialization for processing slow acoustic modulations in 3 month old infants. The results revealed dominant left-hemispheric responses for fast and slow acoustic modulations. In the light of our previous results in newborns (Telkemeyer et al., 2009) and considering the results in the above mentioned studies in 4 month old infants (Minagawa-Kawai et al., 2011), and adults (Boemio et al., 2005) we do not believe that this result proves a discontinuity of developmental lateralization with regard to complex auditory feature processing. Rather we consider experimental factors constitutive for this negative finding. The different levels of motor activity in infants during development and even more the ability to control and withhold from movement is one factor. Also inter-individual differences in response magnitude and optical parameters contributing to background optical properties vary greatly in adults and infants. Notably the analysis of the NIRS results in 3 month olds in our present study revealed significant results for oxy-Hb only. Both, oxy-Hb increases and deoxy-Hb decreases, are associated with neuronal activation (Obrig and Villringer, 2003). However, oxy-Hb is typically characterized by a larger amplitude compared to deoxy-Hb and thus more likely to yield larger effects. On the other hand, concentration changes in oxy-Hb are more susceptible to extracerebral, systemic changes in the hemodynamics (Boden et al., 2007).

In sum, the NIRS data reported here yield less robust effects compared to our previously reported results in newborns. Beyond differences in movement-artifacts, poorer data quality may also suggest that shorter stimulation periods and more repetitions may be a special requirement in studies of these age groups. During the study design we favored identical stimulation paradigms to allow for a comparison with the data in newborns. However, we recommend the use of shorter stimulation durations for future studies on auditory processing, especially when longitudinal aspects over the first year of life are addressed.

To summarize, despite these limitations we consider the symmetric processing of fast and the asymmetric, right-lateralized processing of slow temporal modulations during the auditory analysis to be rather stable from early development. This lateralization may contribute to the lateralization of differential linguistic feature analyses in the incoming auditory stream evolving in parallel to language competence.

AUDITORY-EVOKED POTENTIALS

We simultaneously measured EEG response to the stimuli, which provides a superb temporal resolution allowing for an inquiry into temporal aspects of neural activity correlated with auditory processing.

Since temporal features of the stimuli may affect the waveforms of the evoked response we computed AEPs for the time period of 1 s after stimulus onset. We were interested in developmental changes

of the general AEPs in response to auditory stimulation. The averaged AEPs across all stimulus conditions for both age groups are characterized by an early negative component (N1) followed by a large positivity (P2), mainly in fronto-central positions. Our results show, that the latency of the N1 at around 60 ms did not differ between the two age groups, whereas the amplitude increases with age. Previous results also described such a negative component to be present in the AEPs in newborns and young infants (Novak et al., 1989; Wunderlich et al., 2006). Comparable to our data, the N1 increases in amplitude with development until a discrete component is clearly observed in adulthood (Sussman et al., 2008). Kujala and Näätänen (2010) suggest that the increased amplitude of the N1 reflects an increased fine-grained cortical mapping. However, whether the observed component in infants parallels the N1 component in adults remains under debate (Lippe et al., 2009).

In line with previous research our results furthermore indicate that the infant's AEPs are dominated by a large positivity, especially over fronto-central electrodes and show less discrete components compared to adults (Ceponiene et al., 2002; Kushnerenko et al., 2002; Picton and Taylor, 2007). Similar to previous findings (Wunderlich et al., 2006; Pena et al., 2010) the comparison of the two age groups showed, that the P2 decreases in latency, from a mean peak at 315 ms in 3 month olds to 226 ms in 6 month old infants. In adults the peak of the P2 is described at around 150–200 ms (Näätänen and Picton, 1987; Lippe et al., 2009), in infants it varies around 200–250 ms (Picton and Taylor, 2007). Hence, with increasing brain maturation the latency of the P2 decreases.

Besides these developmental effects on the morphology of the averaged AEPs, we investigated whether differences in the temporal features of the stimuli modulate the AEP-components. Thus, we compared the amplitude of the AEPs for fast and slow acoustic modulations in the two age groups. In both age groups we found an increased amplitude of the N1 for fast compared to slowly modulated stimuli, primarily in fronto-central electrodes. In adults, acoustic information is consciously perceived after around 80 ms after stimulus onset (Näätänen and Winkler, 1999). It has been proposed that at least in adults a discriminable change of any feature of a continuous sound would elicit an N1 (Näätänen and Winkler, 1999). Thus, the N1 is associated with sound detection and is sensitive to physical aspects of the auditory stimulus (Näätänen and Picton, 1987) including the temporal modulation of the total acoustic energy (Ceponiene et al., 2005). Hence, the enhanced response to fast acoustic modulations in our results might be associated with the higher number of acoustic changes during the fast condition (acoustic modulations occur every 12, and every 25 ms, respectively), compared to the slow modulation condition (every 160 and 300 ms).

In both age-groups the amplitude of the P2 was larger for slowly modulated stimuli. However, the functional role of this positivity is not fully understood. In contrast to the N1, the P2 is modulated by consciously perceived, stimulus-specific features, such as emotional content or the salience of the stimulus (Ceponiene et al., 2005; Spreckelmeyer et al., 2009). Derégner et al. (2000) reported an increased amplitude of the P2 elicited by maternal voice compared to a stranger's voice already in newborn infants, suggesting an effect of attention. Therefore, the here observed increased P2 during slow acoustic modulations may indicate an increased atten-

tion or a preference of the infants toward the slowly varying stimuli. Such slow acoustic variations can be found in prosodic features of the speech signal. Studies investigating language acquisition in infancy emphasize the role of suprasegmental, prosodic information (Gleitman and Wanner, 1982; Jusczyk, 1997) during language development as they aid the segmentation of the speech stream into smaller units such as words (Jusczyk et al., 1999). Behavioral studies demonstrated that infants prefer the so called infant-directed speech mode adults use when addressing infants which is characterized by accentuated prosodic features (Werker and McLeod, 1989; Cooper and Aslin, 1990). This finding suggests that infants are more attracted by prosodically modulated features.

Oscillatory responses

In contrast to the AEPs reflecting the effects of temporal variation during the early acoustic analysis, the TFA is a marker of the sustained electrophysiological response. In both age-groups fast and slow acoustic modulations elicited a theta synchronization during the first 300 ms after stimulus onset, which is probably associated with the AEP (Bruneau et al., 1993). This result is in line with Fujioka and Ross (2008) who compared a violin tone to noise-burst stimuli to 4–6 year old children while measuring MEG. The authors report a synchronized theta response during the first ~200 ms after stimulus onset without any difference between the two acoustic stimulations and between hemispheres. Further, the authors reported a desynchronization in the alpha range (8–12 Hz) starting ~400 ms after stimulus onset. These may be similar to the classical Berger-effect in the visual system (Berger, 1929). Our results also revealed a desynchronization beginning ~500 ms after stimulus onset. However, we found this desynchronization in lower frequencies (between 4 and 8 Hz). In 6 but not in 3 month old infants, slow acoustic modulations elicited a significantly stronger desynchronization compared to fast acoustic modulations in that frequency band, hence, suggesting an effect of development. Processing sounds with complex spectrotemporal structure might become more refined with age. A developmental study investigating the phase-locked oscillatory response to musical tones revealed an increase in phase-locking of theta oscillatory activity with age (Shahin et al., 2010). Furthermore, it has been demonstrated that the response to speech sounds in children matures more rapidly than response to non-speech sounds (Pang and Taylor, 2000). Therefore one could speculate, that 6 but not 3 month old infants perceive the slow acoustic modulations at least as more familiar sounds compared to the fast acoustic variations. We did not find oscillatory activity in higher frequency bands, probably due to the fact, that the power of spontaneous oscillations shifts from lower to higher frequencies over early development (Shahin et al., 2010).

CONCLUSION

The present study used simultaneous assessment of hemodynamic and electrophysiological brain responses to investigate the perception of temporal features of non-linguistic complex acoustic stimuli. Subtle auditory differences during the processing of complex auditory stimuli elicit a differential pattern of brain activation in infants. Our NIRS results support the notion that language-specific hemispheric asymmetries are partially driven by acoustic features of

the speech signal. Though the NIRS results in 3 month old infants were inconclusive, we believe that the hemispheric specialization for processing fast and slow temporal modulations during the auditory analysis is rather stable from birth. The AEPs to the onset of the averaged acoustic stimuli indicated an effect of brain maturation on the morphology of the AEPs in general. However, similar to the results of the NIRS no age effect was found in the differential AEP analysis of fast and slow modulations. The larger amplitude of the N1 for fast modulated stimuli may result from higher energy of the acoustic stimulus due to its rapid transitions between different noise bands. On the contrary, the following P2 is affected by more conscious, stimulus-specific features such as attention. Both age groups showed an increased amplitude of the P2 to slow acoustic modulations. Given the importance prosodic features, characterized by slow acoustic modulations, play especially during language acquisition, the increased amplitude might reflect an increased attention of the infants toward the slow modulations. Consistently, the TFA also reveals a stronger theta-band desynchronization for slowly modulated stimuli in the older age group. It is

unclear whether this is due to a more fine-grained processing of complex spectrotemporal sounds in general or whether it is related to effects of attention. To our knowledge, this is the first study investigating slow oscillatory responses to non-linguistic auditory stimulation in early infancy complementing recent results in the language domain (Pena et al., 2010). Though the rather explorative approach precludes a specific interpretation, analyses of the time-frequency representations in infants during language acquisition may shed new light on the way how infants reach instantaneous representations of complex sounds.

ACKNOWLEDGMENTS

Financial support of the EU (NEST 012778, EFRE 20002006 2/6, nEUROpt 201076), and BMBF (BNIC, Bernstein Center for Computational Neuroscience, German-Polish cooperation FK: 01GZ0710) are gratefully acknowledged. Isabell Wartenburger is supported by the Stifterverband für die Deutsche Wissenschaft (Claussen-Simon-Stiftung). We would like to express our gratitude to all parents and their children who participated in this study.

REFERENCES

- Benasich, A. A., and Tallal, P. (2002). Infant discrimination of rapid auditory cues predicts later language impairment. *Behav. Brain Res.* 136, 31–49.
- Berger, H. (1929). Über das Elektroenzephalogramm des Menschen. *Archiv für Psychiatrie und Nervenkrankheiten* 87, 527–550.
- Boden, S., Obrig, H., Kohnecke, C., Benav, H., Koch, S. P., and Steinbrink, J. (2007). The oxygenation response to functional stimulation: is there a physiological meaning to the lag between parameters? *Neuroimage* 36, 100–107.
- Boemio, A., Fromm, S., Braun, A., and Poeppel, D. (2005). Hierarchical and asymmetric temporal sensitivity in human auditory cortices. *Nat. Neurosci.* 8, 389–395.
- Boynton, G. M., Engel, S. A., Glover, G. H., and Heeger, D. J. (1996). Linear systems analysis of functional magnetic resonance imaging in human V1. *J. Neurosci.* 16, 4207–4221.
- Bruneau, N., Roux, S., Guerin, P., Garreau, B., and Lelord, G. (1993). Auditory stimulus intensity responses and frontal midline theta rhythm. *Electroencephalogr. Clin. Neurophysiol.* 86, 213–216.
- Ceponiene, R., Alku, P., Westerfield, M., Torki, M., and Townsend, J. (2005). ERPs differentiate syllable and non-phonetic sound processing in children and adults. *Psychophysiology* 42, 391–406.
- Ceponiene, R., Rinne, T., and Näätänen, R. (2002). Maturation of cortical sound processing as indexed by event-related potentials. *Clin. Neurophysiol.* 113, 870–882.
- Choudhury, N., and Benasich, A. A. (2010). Maturation of auditory evoked potentials from 6 to 48 months: prediction to 3 and 4 year language and cognitive abilities. *Clin. Neurophysiol.* 122, 320–338.
- Choudhury, N., Leppänen, P. H., Leevors, H. J., and Benasich, A. A. (2007). Infant information processing and family history of specific language impairment: converging evidence for RAP deficits from two paradigms. *Dev. Sci.* 10, 213–236.
- Cooper, R. P., and Aslin, R. N. (1990). Preference for infant-directed speech in the first month after birth. *Child Dev.* 61, 1584–1595.
- Cope, M., and Delpy, D. T. (1988). System for long-term measurement of cerebral blood and tissue oxygenation on newborn infants by near infra-red transillumination. *Med. Biol. Eng. Comput.* 26, 289–294.
- Csibra, G., Davis, G., Spratling, M. W., and Johnson, M. H. (2000). Gamma oscillations and object processing in the infant brain. *Science* 290, 1582–1585.
- Dehaene-Lambertz, G., Dehaene, S., and Hertz-Pannier, L. (2002). Functional neuroimaging of speech perception in infants. *Science* 298, 2013–2015.
- Dehaene-Lambertz, G., Hertz-Pannier, L., Dubois, J., Meriaux, S., Roche, A., Sigman, M., and Dehaene, S. (2006). Functional organization of perisylvian activation during presentation of sentences in preverbal infants. *Proc. Natl. Acad. Sci. U.S.A.* 103, 14240–14245.
- Deregner, R. A., Nelson, C. A., Thomas, K. M., Wewerka, S., and Georgieff, M. K. (2000). Neurophysiologic evaluation of auditory recognition memory in healthy newborn infants and infants of diabetic mothers. *J. Pediatr.* 137, 777–784.
- Fox, P. T., and Raichle, M. E. (1986). Focal physiological uncoupling of cerebral blood flow and oxidative metabolism during somatosensory stimulation in human subjects. *Proc. Natl. Acad. Sci. U.S.A.* 83, 1140–1144.
- Friederici, A. D. (2005). Neurophysiological markers of early language acquisition: from syllables to sentences. *Trends Cogn. Sci.* 9, 481–488.
- Fujioka, T., and Ross, B. (2008). Auditory processing indexed by stimulus-induced alpha desynchronization in children. *Int. J. Psychophysiol.* 68, 130–140.
- Gervain, J., and Mehler, J. (2010). Speech perception and language acquisition in the first year of life. *Annu. Rev. Psychol.* 61, 191–218.
- Giraud, A. L., Kleinschmidt, A., Poeppel, D., Lund, T. E., Frackowiak, R. S., and Laufs, H. (2007). Endogenous cortical rhythms determine cerebral specialization for speech perception and production. *Neuron* 56, 1127–1134.
- Gleitman, L. R., and Wanner, E. (1982). “Language acquisition: the state of the state of the art,” in *Language Acquisition: The State of the Art*, eds E. Wanner and L. R. Gleitman (New York: Cambridge University Press), 3–48.
- Greenhouse, S. W., and Geisser, S. (1959). On methods in the analysis of profile data. *Psychometrika* 24, 95–112.
- Hickok, G., and Poeppel, D. (2007). The cortical organization of speech processing. *Nat. Rev. Neurosci.* 8, 393–402.
- Holland, S. K., Plante, E., Weber, B. A., Strawsburg, R. H., Schmithorst, V. J., and Ball, W. S. Jr. (2001). Normal fMRI brain activation patterns in children performing a verb generation task. *Neuroimage* 14, 837–843.
- Homae, F., Watanabe, H., Nakano, T., Asakawa, K., and Taga, G. (2006). The right hemisphere of sleeping infant perceives sentential prosody. *Neurosci. Res.* 54, 276–280.
- Hoogenboom, N., Schoffelen, J. M., Oostenveld, R., Parkes, L. M., and Fries, P. (2006). Localizing human visual gamma-band activity in frequency, time and space. *Neuroimage* 29, 764–773.
- Jensen, O., Hari, R., and Kaila, K. (2002). Visually evoked gamma responses in the human brain are enhanced during voluntary hyperventilation. *Neuroimage* 15, 575–586.
- Jusczyk, P. W. (1997). *The Discovery of Spoken Language*. Cambridge, MA: MIT Press.
- Jusczyk, P. W., Houston, D. M., and Newsome, M. (1999). The beginnings of word segmentation in English-learning infants. *Cogn. Psychol.* 39, 159–207.
- Kleinschmidt, A., Obrig, H., Requardt, M., Merboldt, K. D., Dirnagl, U., Villringer, A., and Frahm, J. (1996). Simultaneous recording of cerebral blood oxygenation changes during human brain activation by magnetic resonance imaging and near-infrared spectroscopy. *J. Cereb. Blood Flow Metab.* 16, 817–826.
- Koch, S. P., Werner, P., Steinbrink, J., Fries, P., and Obrig, H. (2009). Stimulus-induced and state-dependent sustained gamma activity is tightly coupled to the hemodynamic response in humans. *J. Neurosci.* 29, 13962–13970.
- Kuhl, P., and Rivera-Gaxiola, M. (2008). Neural substrates of language acquisition. *Annu. Rev. Neurosci.* 31, 511–534.
- Kuhl, P. K. (2004). Early language acquisition: cracking the speech code. *Nat. Rev. Neurosci.* 5, 831–843.

- Kuhl, P. K., Williams, K. A., Lacerda, F., Stevens, K. N., and Lindblom, B. (1992). Linguistic experience alters phonetic perception in infants by 6 months of age. *Science* 255, 606–608.
- Kujala, T., and Näätänen, R. (2010). The adaptive brain: a neurophysiological perspective. *Prog. Neurobiol.* 91, 55–67.
- Kushnerenko, E., Ceponiene, R., Balan, P., Fellman, V., Huottilainen, M., and Näätänen, R. (2002). Maturation of the auditory event-related potentials during the first year of life. *Neuroreport* 13, 47–51.
- Lippe, S., Martinez-Montes, E., Arcand, C., and Lassonde, M. (2009). Electrophysiological study of auditory development. *Neuroscience* 164, 1108–1118.
- Lopes da Silva, F. (1991). Neural mechanisms underlying brain waves: from neural membranes to networks. *Electroencephalogr. Clin. Neurophysiol.* 79, 81–93.
- Mehler, J., Sebastian-Galles, N., and Nespore, M. (2004). “Biological foundations of language: language acquisition, cues for parameter setting and the bilingual infant,” in *The New Cognitive Neuroscience*, ed. M. Gazzaniga (Cambridge, MA: MIT Press), 825–836.
- Minagawa-Kawai, Y., Mori, K., Naoi, N., and Kojima, S. (2007). Neural attunement processes in infants during the acquisition of a language-specific phonemic contrast. *J. Neurosci.* 27, 315–321.
- Minagawa-Kawai, Y., van der Lely, H., Ramus, E., Sato, Y., Mazuka, R., and Dupoux, E. (2011). Optical brain imaging reveals general auditory and language-specific processing in early infant development. *Cereb. Cortex* 21, 254–261.
- Näätänen, R., and Picton, T. (1987). The N1 wave of the human electric and magnetic response to sound: a review and an analysis of the component structure. *Psychophysiology* 24, 375–425.
- Näätänen, R., and Winkler, I. (1999). The concept of auditory stimulus representation in cognitive neuroscience. *Psychol. Bull.* 125, 826–859.
- Nierhaus, T., Schön, T., Becker, R., Ritter, P., and Villringer, A. (2009). Background and evoked activity and their interaction in the human brain. *Magn. Reson. Imaging* 27, 1140–1150.
- Novak, G. P., Kurtzberg, D., Kreuzer, J. A., and Vaughan, H. G. Jr. (1989). Cortical responses to speech sounds and their formants in normal infants: maturational sequence and spatiotemporal analysis. *Electroencephalogr. Clin. Neurophysiol.* 73, 295–305.
- Obrig, H., Rossi, S., Telkemeyer, S., and Wartenburger, I. (2010). From acoustic segmentation to language processing: evidence from optical imaging. *Front. Neuroenergetics* 2:13. doi: 10.3389/fnene.2010.00013
- Obrig, H., and Villringer, A. (2003). Beyond the visible-imaging the human brain with light. *J. Cereb. Blood Flow Metab.* 23, 1–18.
- Pang, E. W., and Taylor, M. J. (2000). Tracking the development of the N1 from age 3 to adulthood: an examination of speech and non-speech stimuli. *Clin. Neurophysiol.* 111, 388–397.
- Pena, M., Maki, A., Kovacic, D., Dehaene-Lambertz, G., Koizumi, H., Bouquet, F., and Mehler, J. (2003). Sounds and silence: an optical topography study of language recognition at birth. *Proc. Natl. Acad. Sci. U.S.A.* 100, 11702–11705.
- Pena, M., Pittaluga, E., and Mehler, J. (2010). Language acquisition in premature and full-term infants. *Proc. Natl. Acad. Sci. U.S.A.* 107, 3823–3828.
- Perani, D., Saccuman, M. C., Scifo, P., Spada, D., Andreolli, G., Rovelli, R., Baldoli, C., and Koelsch, S. (2010). Functional specializations for music processing in the human newborn brain. *Proc. Natl. Acad. Sci. U.S.A.* 107, 4758–4763.
- Pfurtscheller, G., and Lopes da Silva, F. H. (1999). Event-related EEG/MEG synchronization and desynchronization: basic principles. *Clin. Neurophysiol.* 110, 1842–1857.
- Picton, T. W., and Taylor, M. J. (2007). Electrophysiological evaluation of human brain development. *Dev. Neuropsychol.* 31, 249–278.
- Poeppel, D. (2003). The analysis of speech in different temporal integration windows: cerebral lateralization as ‘asymmetric sampling in time’. *Speech Commun.* 41, 245–255.
- Poeppel, D., Idsardi, W. J., and van Wassenhove, V. (2008). Speech perception at the interface of neurobiology and linguistics. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 363, 1071–1086.
- Rossi, S., Jürgenson, I. B., Hanulíková, A., Telkemeyer, S., Wartenburger, I., and Obrig, H. (2010). Implicit processing of phonotactic cues: evidence from electrophysiological and vascular responses. *J. Cogn. Neurosci.* 1, 1–13.
- Schönwiesner, M., Rübsem, R., and von Cramon, D. Y. (2005). Hemispheric asymmetry for spectral and temporal processing in the human antero-lateral auditory belt cortex. *Eur. J. Neurosci.* 22, 1521–1528.
- Shahin, A. J., Trainor, L. J., Roberts, L. E., Backer, K. C., and Miller, L. M. (2010). Development of auditory phase-locked activity for music sounds. *J. Neurophysiol.* 103, 218–229.
- Sharbrough, F., Chatrian, G. E., Lesser, R. P., Lüders, H., Nuwer, M., and Picton, T. W. (1991). American Electroencephalographic Society guidelines for standard electrode position nomenclature. *J. Clin. Neurophysiol.* 8, 200–202.
- Singer, W. (1999). Time as coding space? *Curr. Opin. Neurobiol.* 9, 189–194.
- Singer, W. (1993). Synchronization of cortical activity and its putative role in information processing and learning. *Annu. Rev. Physiol.* 55, 349–374.
- Spreckelmeyer, K. N., Kutas, M., Urbach, T., Altenmüller, E., and Münte, T. F. (2009). Neural processing of vocal emotion and identity. *Brain Cogn.* 69, 121–126.
- Stefanics, G., Haden, G., Huottilainen, M., Balazs, L., Sziller, I., Beke, A., Fellman, V., and Winkler, I. (2007). Auditory temporal grouping in newborn infants. *Psychophysiology* 44, 697–702.
- Steinbrink, J., Villringer, A., Kempf, F., Haux, D., Boden, S., and Obrig, H. (2006). Illuminating the BOLD signal: combined fMRI-fNIRS studies. *Magn. Reson. Imaging* 24, 495–505.
- Stevens, K. N. (1998). *Acoustic Phonetics*. Cambridge, MA: MIT Press.
- Sussman, E., Steinschneider, M., Gumenyuk, V., Grushko, J., and Lawson, K. (2008). The maturation of human evoked brain potentials to sounds presented at different stimulus rates. *Hear Res.* 236, 61–79.
- Taga, G., Asakawa, K., Maki, A., Konishi, Y., and Koizumi, H. (2003). Brain imaging in awake infants by near-infrared optical topography. *Proc. Natl. Acad. Sci. U.S.A.* 100, 10722–10727.
- Tallon-Baudry, C., and Bertrand, O. (1999). Oscillatory gamma activity in humans and its role in object representation. *Trends Cogn. Sci.* 3, 151–162.
- Telkemeyer, S., Rossi, S., Koch, S. P., Nierhaus, T., Steinbrink, J., Poeppel, D., Obrig, H., and Wartenburger, I. (2009). Sensitivity of newborn auditory cortex to the temporal structure of sounds. *J. Neurosci.* 29, 14726–14733.
- Werker, J. F., and McLeod, P. J. (1989). Infant preference for both male and female infant-directed talk: a developmental study of attentional and affective responsiveness. *Can. J. Psychol.* 43, 230–246.
- Wunderlich, J. L., Cone-Wesson, B. K., and Shepherd, R. (2006). Maturation of the cortical auditory evoked potential in infants and young children. *Hear Res.* 212, 185–202.
- Zaehle, T., Jancke, L., Herrmann, C. S., and Meyer, M. (2009). Pre-attentive spectro-temporal feature processing in the human auditory system. *Brain Topogr.* 22, 97–108.
- Zaehle, T., Jancke, L., and Meyer, M. (2007). Electrical brain imaging evidences left auditory cortex involvement in speech and non-speech discrimination based on temporal features. *Behav. Brain Funct.* 3, 63.
- Zaehle, T., Wüstenberg, T., Meyer, M., and Jancke, L. (2004). Evidence for rapid auditory perception as the foundation of speech processing: a sparse temporal sampling fMRI study. *Eur. J. Neurosci.* 20, 2447–2456.
- Zatorre, R. J., and Belin, P. (2001). Spectral and temporal processing in human auditory cortex. *Cereb. Cortex* 11, 946–953.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 12 November 2010; paper pending published: 24 December 2010; accepted: 25 March 2011; published online: 09 April 2011.

Citation: Telkemeyer S, Rossi S, Nierhaus T, Steinbrink J, Obrig H and Wartenburger I (2011) Acoustic processing of temporally modulated sounds in infants: evidence from a combined near-infrared spectroscopy and EEG study. *Front. Psychol.* 2:62. doi: 10.3389/fpsyg.2011.00062

This article was submitted to *Frontiers in Language Sciences*, a specialty of *Frontiers in Psychology*.

Copyright © 2011 Telkemeyer, Rossi, Nierhaus, Steinbrink, Obrig and Wartenburger. This is an open-access article subject to a non-exclusive license between the authors and Frontiers Media SA, which permits use, distribution and reproduction in other forums, provided the original authors and source are credited and other Frontiers conditions are complied with.



Assessing signal-driven mechanisms in neonates: brain responses to temporally and spectrally different sounds

Yasuyo Minagawa-Kawai¹, Alejandrina Cristià², Inga Vendelin², Dominique Cabrol³ and Emmanuel Dupoux^{2*}

¹ Graduate School of Human Relations, Global-COE, CARLS, Keio University, Tokyo, Japan

² Laboratoire de Sciences Cognitives et Psycholinguistique, EHESS, ENS-DEC, CNRS, Paris, France

³ Department of Gynecology and Obstetrics, AP-HP Cochin Port Royal, Paris, France

Edited by:

Judit Gervain, CNRS – Université Paris Descartes, France

Reviewed by:

Silke Telkemeyer, Free University

Berlin, Germany

Gábor Péter Háden, Institute for Psychology, Hungarian Academy of Sciences, Hungary

*Correspondence:

Emmanuel Dupoux, Laboratoire de Sciences Cognitives et Psycholinguistique, EHESS, ENS-DEC, CNRS, 29, rue d'Ulm, 75005 Paris, France.

e-mail: emmanuel.dupoux@gmail.com

Past studies have found that, in adults, the acoustic properties of sound signals (such as fast versus slow temporal features) differentially activate the left and right hemispheres, and some have hypothesized that left-lateralization for speech processing may follow from left-lateralization to rapidly changing signals. Here, we tested whether newborns' brains show some evidence of signal-specific lateralization responses using near-infrared spectroscopy (NIRS) and auditory stimuli that elicits lateralized responses in adults, composed of segments that vary in duration and spectral diversity. We found significantly greater bilateral responses of oxygenated hemoglobin (oxy-Hb) in the temporal areas for stimuli with a minimum segment duration of 21 ms, than stimuli with a minimum segment duration of 667 ms. However, we found no evidence for hemispheric asymmetries dependent on the stimulus characteristics. We hypothesize that acoustic-based functional brain asymmetries may develop throughout early infancy, and discuss their possible relationship with brain asymmetries for language.

Keywords: near-infrared spectroscopy, neonates, functional lateralization, auditory cortex, hemispheric specialization, development

INTRODUCTION

There is a considerable gap between what we know about the language brain networks in adulthood and the development of this network in infancy. At the root of this gap is the lack of non-invasive, infant-friendly imaging tools. In the past decade, multi-channel near-infrared spectroscopy (NIRS) has generated a lot of interest, as it can measure infants' cerebral cortex hemodynamic responses with a spatial resolution of about 2–3 cm and is as innocuous as electroencephalography (EEG; Minagawa-Kawai et al., 2008 for a review). One of the issues that remain largely unsettled is the developmental origin of the left lateralization for language that is observed in adults (see Minagawa-Kawai et al., 2011a). At least two general hypotheses have been put forward; the *domain-driven hypothesis* postulates that there brain networks specifically dedicated to language (see Dehaene-Lambertz et al., 2006, 2010). The *signal-driven hypothesis* holds that specialization for higher cognitive functions emerges out of interactions between sensory signals and general learning principles (see Elman et al., 1997). In particular, the left–right asymmetry for language found in adults has been tied to a structural preference of left cortices for fast changing signals as opposed to slow changing signals. Such patterns have been supported by brain data collected on human adults (Zatorre and Belin, 2001; Boemio et al., 2005; Schonwiesner et al., 2005; Jamison et al., 2006), animals (Wetzel et al., 1998; Rybalko et al., 2006), and, more recently, infants (Homae et al., 2006). In this general instantiation, the left hemisphere is preferentially involved in processing rapid changes such as those that distinguish phonemes, whereas the right hemisphere is more engaged in spectral processing such as that required for the discrimination of prosodic changes. The fast/slow distinction in itself cannot be sufficient to account for the

lateralization of language in adults. Indeed, language is a very complex signal defined at several hierarchical levels (features, phonemes, morphemes, utterances), the acoustic correlates of which use the entire spectrum of cues, from very short temporal events to complex and rather slow spectral patterns (see also Minagawa-Kawai et al., 2011a). Nonetheless, the idea that signal characteristics provide an *initial* bias for left/right lateralization is theoretically attractive, as it would reduce the number of specialized brain networks stipulated in infants, and makes important predictions regarding the development of functional specialization for speech on the one hand, and the biases affecting auditory processing even in adults.

Therefore, it is not surprising that this view has inspired a great deal of research over the past 10 years. Over this period of time, a number of operational definitions of precisely what in the acoustic signal determines lateralization have been proposed. Some of the acoustic candidates proposed to determine the hemispheric dominance include the rate of change (Belin et al., 1998), the duration of constituent events (shorter leads to left dominant; Bedoin et al., 2010), the type of encoding (temporal versus spectral, Zatorre and Belin, 2001), and the size of the analysis window (Poeppel et al., 2003). Among these many hypotheses, two dominant views are emerging. The *multi-time-resolution model* (Poeppel et al., 2003) stipulates that the left and right hemispheres are tuned to different frequencies for sampling and integrating information. Specifically, the left (or bilateral) cortices perform rapid sampling of small integration windows, whereas the right cortices perform slow sampling over a rather large integration window (Hickok and Poeppel, 2007; Poeppel et al., 2008). The *temporal versus spectral hypothesis* (Zatorre and Belin, 2001) emphasizes the trade-off that arises between the representation of speech that have good temporal resolution and

those that have good spectral resolution. Temporal information would be represented most accurately in the left hemisphere and spectral information in the right hemisphere. Even though these hypotheses have a family resemblance, they are conceptually distinct, and have given rise to distinct sets of stimuli, which vary in the lateralization patterns they elicit in different brain structures.

In general, signal-driven asymmetries have most frequently been observed in the superior temporal gyrus (STG), particularly the auditory cortex (Belin et al., 1998; Zatorre and Belin, 2001; Jamison et al., 2006). In addition, anterior and posterior superior temporal sulcus (STS) (Zatorre and Belin, 2001; Boemio et al., 2005), as well as the temporo-parietal region (Telkemeyer et al., 2009), have been reported to be sensitive to the acoustic properties of stimuli. Regarding lateralization, some studies document left dominance for fast stimuli and right dominance for slow ones, whereas other studies only report half of this pattern (typically, only right dominance for slow stimuli; e.g., Boemio et al., 2005; Britton et al., 2009). We believe that some of this diversity could be due to using different stimuli and different fractioning of regions adjacent to the auditory cortex (see Belin et al., 1998; Giraud et al., 2007). Therefore, it becomes imperative to assess the existence of signal-driven asymmetries early on in development using a range of different stimuli.

The first infant studies documenting variable lateralization patterns in response to auditory stimulation typically focused on the contrast between speech and non-speech. For example, Peña et al. (2003) found that the newborn brain exhibited greater activation in the left hemisphere when presented with forward than with backward speech. Sato et al. (2006) replicated this effect for the newborns' native language, but documented symmetrical temporal responses to forward and backward speech in a language not spoken by the newborns' mothers. These results suggest that experience may modulate left dominance in response to speech; however, since forward and backward speech (native or non-native) are similar in terms of the fast/slow distinction, these results cannot shed light on the possible existence of signal-driven asymmetries.

In fact, only one published study has tested one implementation of the signal-driven hypothesis in newborns. Telkemeyer et al. (2009) presented newborns with stimuli previously used on adults by Boemio et al. (2005). The stimuli set used in the newborn study consisted in static tones concatenated differently in the different conditions, so as to vary their temporal structure. There were two *fast* conditions, obtained by concatenating segments of either 12 or 25 ms of length; and two *slow* conditions, where segments were either 160 ms or 300 ms of duration (the adult version had two additional conditions, where segments were 45 and 85 ms in duration; and a second set of frequency modulated tones with the same temporal manipulations). The adult fMRI study by Boemio et al. (2005) had documented a bilateral response to the temporal structure, such that there were linearly greater activation with greater segment length in both left and right STS and STG. Telkemeyer et al. (2009) measured oxy- and deoxy-Hb changes using six NIRS channels placed over temporal and parietal areas in each hemisphere. The results reported only included deoxy-Hb signals on three NIRS channels in each hemisphere; one of these three channels exhibited a greater activation over the right temporo-parietal area for the 160 and 300 ms conditions. This last result, however, was not

particularly strong statistically, as the greater activation in the right hemisphere was only significant on a one-tailed, uncorrected *t*-test. In addition, results showed an overall stronger bilateral response for stimuli using 25 ms modulations compared to all other durations. Thus, in neither the adult nor newborn data did those stimuli modulate the activation of the *left* hemisphere. Both of these results are in line with a recent version of the multi-temporal-resolution model that proposes that temporal structure modulates activation *bilaterally* (Poeppel et al., 2008).

However, as we mentioned above, this is not the only instantiation of the signal-based hypothesis, and different sets of stimuli have led to more or less clear-cut lateralization patterns. In the study presented here, we re-evaluate the possibility that newborns' functional lateralization can be modulated by acoustic characteristics of the auditory stimulation using a set of stimuli that vary in *both* temporal and spectral complexity; and which has elicited a clear left-right asymmetry in adults according to both PET (Zatorre and Belin, 2001) and fMRI (Jamison et al., 2006). These adult data suggested that left temporal cortices were most responsive to fast, spectrally simple stimuli; whereas spectrally variable, slow stimuli elicited greater activation in right temporal cortices. As noted above, newborn studies sometimes lack statistical power. We tried to maximize the statistical power in the present study in several ways. First, we selected a small subset of 3 out of the original 11 conditions used by Zatorre and Belin (2001), selecting an intermediate case with long duration and little variability ("control") as well as the two most extreme conditions, namely brief, 21 ms segments with little spectral variability across them (fixed at 1 octave jumps; "temporal"); and long, 667 ms segments with greater spectral variability across subsequent segments (any of eight different frequencies could occur; "spectral"). Fewer conditions ensured more trials per condition; and the choice of extreme temporal and spectral characteristics enhanced the probability of finding signal-driven asymmetries. To further increase the power, the artifacts were rejected so as to keep as much good data as possible (for instance, an artifact in the middle of a trial will not necessarily yield a removal of the entire trial, but only of the section of signal that was artifacted). In addition, we measured signals from both small (2.5 cm) and large (5.6 cm) separations between probes and detectors, potentially opening up the possibility of measuring "deeper" brain regions; and we provide data for a total of 14 channels per hemisphere. With this configuration, we could sample at several of the potential candidates for asymmetries: STG involving auditory cortex (one shallow and deep channels), anterior STG and STS (two channels), and temporo-parietal (one channel – similar to the one showing a trend toward an asymmetrical activation in Telkemeyer et al., 2009). Finally, we have investigated both oxygenated hemoglobin (oxy-Hb) and deoxy-Hb, while only deoxy-Hb was reported in the previous newborn study looking at the question. Although it has been argued that deoxy-Hb is less affected by extracerebral artifacts than oxy-Hb (see Obrig et al., 2010; and Obrig and Villringer, 2003, for important considerations regarding the two measures), most previous NIRS research on infants focuses on oxy-Hb, which appears to be more stable across children. In a recent review, Lloyd-Fox et al. (2010) report that of all infant NIRS studies published up to 2009, 83% reported increases in

oxy with stimulation, 6% decreases, 1% no change, and 3% did not report it; in contrast, 28% reported increases in deoxy with stimulation, 17% decreases, 19% no change, and 36% did not report it altogether. Our analyses below concentrate on oxy-Hb; results with deoxy-Hb, showing the same pattern but with weaker significance levels, are reported in **Table A1** in Appendix.

We expected that through the combination of these strategies, we would be able to document signal-based asymmetries as elicited by maximally different stimuli in terms of their spectral and temporal composition. Based on the previous adult work carried out with these same stimuli, we predicted that, if temporal and spectral characteristics drove lateralization early on in development, temporal stimuli should evoke larger activations than control stimuli in the LH, and spectral stimuli should evoke larger activations in the RH when compared to control stimuli.

MATERIALS AND METHODS

PARTICIPANTS

Thirty-eight newborns were tested; we excluded nine infants who did not have enough data due to movement artifacts, loose attachment, or test interruption due to infant discomfort. The final sample consisted of 29 infants (15 boys), tested between 0 and 5 days of age ($M = 2.41$), for which we obtained the equivalent of at least 10 full analyzable trials per condition, for more than 50% of the channels. Parental report indicated that both parents were left handed in one case (<5%), and one parent was left handed in seven cases (<25%). All infants were full-term without medical problems. Consent forms were obtained from parents before the infants' participation. This study was approved by the CPP Ile de France III Committee (No. ID RCB (AFSSAPS) 2007-A01142-51).

Table 1 | Levels of activations (oxy-Hb changes) for temporal, control, and spectral conditions.

Channel	Temporal					Control					Spectral				
	β	T	N	p (unc)	p (corr)	β	T	N	p (unc)	p (corr)	β	T	N	p (unc)	p (corr)
LEFT SHALLOW															
1 MFG/IFG	0.0129	2.49	28	0.022	ns										
2 IFG	0.0122	2.11	29	0.046	ns	0.0111	1.75	29	0.088	ns					
3 aSTG	0.0208	3.20	28	0.004	~										
4 STG/IFG	0.0223	3.99	29	0.001	0.008										
5 Precentral	0.0170	2.59	29	0.017	ns										
6 STG/STS	0.0154	1.86	29	0.072	ns	0.0117	2.09	29	0.048	ns	0.0117	1.95	29	0.063	ns
7 STG	0.0165	2.96	29	0.008	ns	0.0150	2.37	29	0.028	ns	0.0208	2.60	29	0.017	ns
8 SMG	0.0208	3.16	29	0.005	~	0.0156	2.07	29	0.050	ns					
10 SMG/pSTG	0.0156	2.64	29	0.016	ns										
LEFT DEEP															
4a FG/TG	0.0162	2.92	29	0.008	ns	0.0118	1.78	29	0.084	ns	0.0122	1.75	29	0.088	ns
4b FG/postcent	0.0149	2.87	28	0.009	ns	0.0099	2.11	28	0.047	ns					
7a STG/SMG	0.0265	3.43	28	0.002	0.037	0.0122	2.41	29	0.025	ns					
postcentral/															
7b TG/FG	0.0160	2.24	29	0.036	ns						0.0181	2.36	28	0.028	ns
RIGHT SHALLOW															
1 MFG/IFG	0.0173	2.18	29	0.041	ns						-0.011	-1.87	28	0.072	ns
2 IFG	0.0200	2.95	29	0.008	ns										
4 STG/IFG	0.0259	5.27	29	0.000	0.000	0.0163	2.49	29	0.021	ns	0.0178	2.43	29	0.024	ns
5 precentral	0.0131	1.90	29	0.068	ns	0.0221	3.34	29	0.003	0.044					
6 STG/STS	0.0194	3.14	29	0.005	~										
7 STG	0.0218	3.74	29	0.001	0.016	0.0155	2.14	29	0.044	ns	0.0135	2.05	29	0.052	ns
8 SMG	0.0229	3.85	29	0.001	0.011										
10 SMG/pSTG											0.0131	2.30	28	0.032	ns
RIGHT DEEP															
4a FG/TG	0.0178	2.82	29	0.011	ns										
4b FG/postcent	0.0169	3.31	29	0.003	0.050										
7a STG/SMG	0.0252	4.84	29	0.000	0.001	0.0172	2.62	28	0.016	ns					
postcentral															
7b/TG/FG	0.0134	2.93	29	0.008	ns										

N, number of children contributing data for that channel and condition; *unc*, uncorrected; *corr*, corrected. MFG, middle frontal gyrus; IFG, inferior frontal gyrus; *a*, anterior; STG, superior temporal gyrus; STS, superior temporal sulcus; *p*, posterior; MTG, middle temporal gyrus; FG, frontal gyrus; TG, temporal gyrus.

STIMULI

The stimuli were pure-tone patterns which varied in their spectral and temporal properties as shown in **Figure 1** using two parameters: f , the frequency separation between two adjacent tones, and t , the base duration of a tone segment. The original stimuli in Zatorre and Belin (2001) included a five-step continuum of stimuli varying in spectral complexity, a five-step continuum of stimuli varying in temporal complexity, and a control condition (see **Figure 1** of Zatorre and Belin, 2001 for more details). Among them, we selected two conditions that constituted the extreme exemplars of the spectral and temporal continua, respectively, plus the control condition (Zatorre and Belin, 2001). The acoustic parameters of the stimuli in these three conditions are as follows. The *control* stimuli (low spectral complexity – long duration) consisted in randomly selected pure tones of either 500 or 1000 Hz (a frequency separation of one octave, i.e., $f = 1200$ cents) and a long base duration ($t = 667$ ms). These had low spectral complexity because only two frequencies were used; this is evident in the long term spectrum shown in **Figure 1** (top panel). *Spectral* stimuli (high spectral complexity – long duration) had the same temporal properties as the control stimuli ($t = 667$ ms), but consisted of tones selected from a set of 33 frequencies, roughly spaced by $1/32$ of an octave, between 500 and 1000 Hz (minimum separation $f = 37.5$ cents, see the more complex long term spectrum, **Figure 1**, bottom right). *Temporal* stimuli (low spectral complexity – short duration) kept the spectral properties identical to those in the control condition, with randomly selected tones of either 500 or 1000 Hz ($f = 1200$ cents), but used a much shorter base duration ($t = 21$ ms). The stimuli were generated through a random sampling procedure respecting the following distribution: if the base duration was t for a condition, the actual duration of a tone of a given frequency varied from t to nt , where n is an integer, with probability of $1/2^n$. Therefore, the highest probability of occurrence ($1/2^1 = 0.5$) is equal to the shortest duration (t). The next longer tone would have a duration of $2t$, and would occur with a probability of $1/2^2 = 0.25$ (a probability for $3t$ is $1/2^3 = 0.125$

and so forth). Therefore, the stimuli with a base duration of t contained segments of varying duration, whose average was $\sum_n n \cdot 2^{-n} / \sum_n 2^{-n} \sim 2t$. This random sampling procedure was chosen in order to avoid the appearance of a low frequency spectral component that would have arose if a strict alternation had been used. We truncated the original stimuli in Zatorre and Belin (2001) to last 10 s (1 trial = 10 s stimulation). Trials belonging to the three conditions were presented in random alternations. A jittered silence period was inserted between the trials (8–14 s). A run was the concatenation of eight trials of each type (i.e., a total of 24 trials per run). Infants we presented with 2 or 3 runs depending on their state; for a total of 24 trials per condition, or 72 trials in overall total. Specifically, if the infant was calm enough without too much movement even after the two runs, we kept on recording. Twenty-one out of the 29 infants included finished all 3 runs; in average, infants completed 2.7 runs.

PROCEDURE

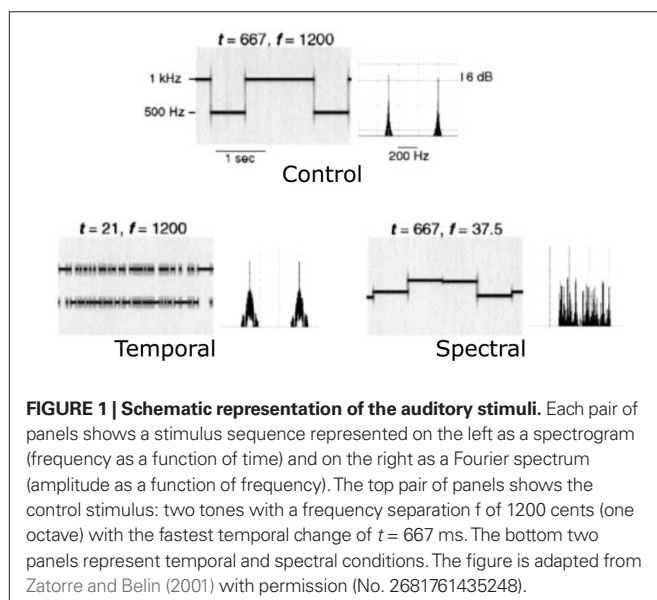
Changes in hemoglobin concentrations and oxygenation levels in the bilateral temporal and frontal areas were recorded using NIRS. Our NIRS system (UCL-NTS, Department of Medical Physics and Bioengineering, UCL, London, UK; Everdell et al., 2005) continuously emits near-infrared lasers of two wavelengths (670 and 850 nm) from four sources on each pad, each of them being modulated by a specific frequency for source separation. The light transmitted through scalp and brain tissues is measured by four detectors on each pad, as shown in **Figure 2**. This configuration provides data from 10 channels between adjacent source-detector pairs, at a distance of 25 mm; and four channels between non-adjacent pairs, at a distance of 56 mm. According to Fukui et al. (2003)'s model on light penetration in newborns, the NIR light remains in a shallow area of the gray matter with separations of 20–30 mm, but it may go into white matter with 50 mm separations. Therefore, we will call the first 10 “shallow,” and the other 4 “deep.” Channel numbering goes from anterior to posterior, with deep channels coded with the same number, as a shallow counterpart that travels in the same region and a letter (a/b).

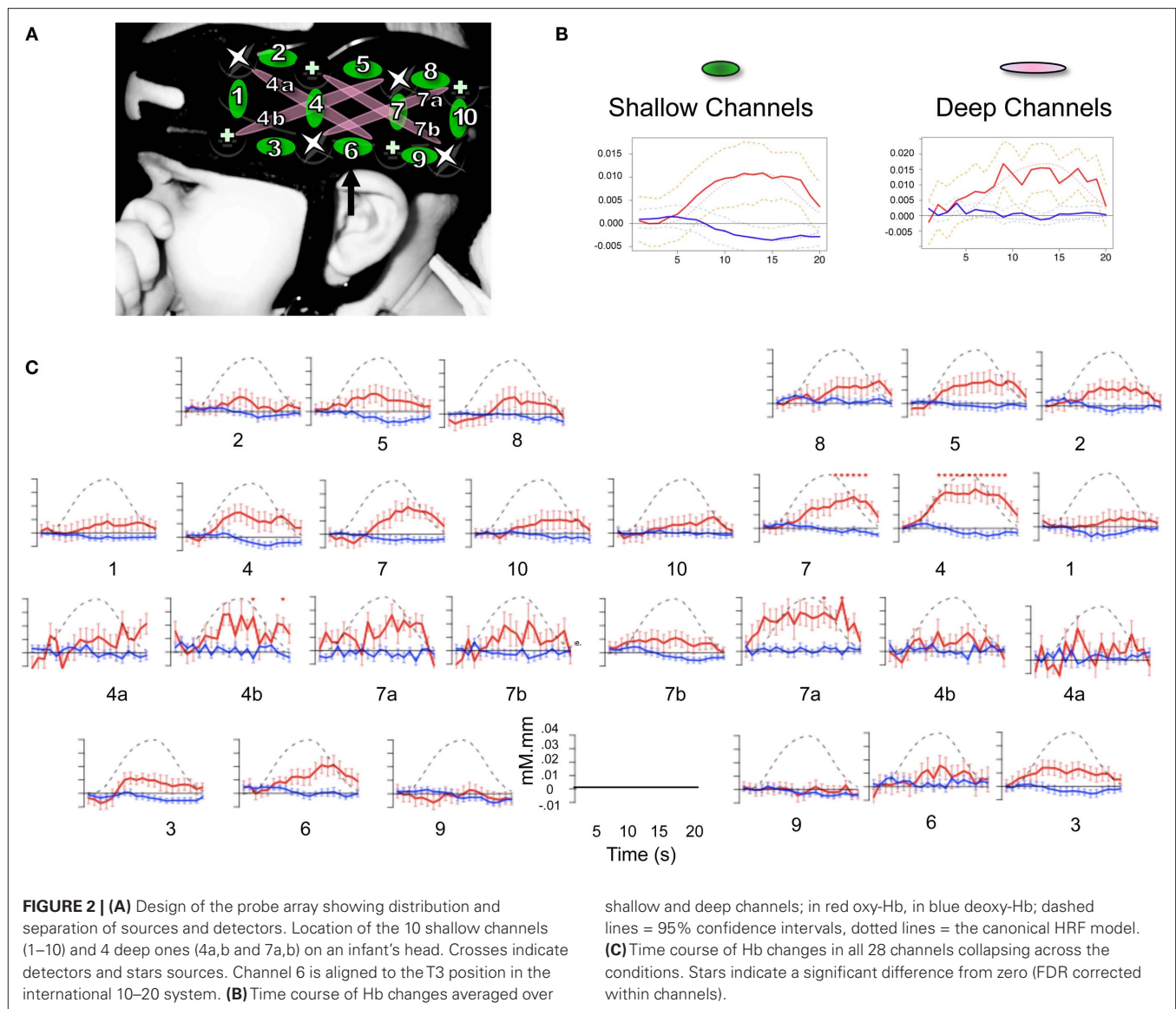
The NIRS recording was performed in a sound-attenuated room. In attaching the probes, we used the international 10–20 system as a reference, aligning the bottom of the pad with the T3–T5 line using anatomical landmarks. Once the cap was fit, the stimuli were presented from a loudspeaker at about 80 dB measured at the approximate location of the infant's head. The infants were inside their cots while they listened to the stimuli; all of them were asleep for most or all of the study.

DATA ANALYSIS

Preprocessing

We used the modified Beer Lambert Law to convert intensity signals into oxy- and deoxy-Hb concentration. Movement artifacts are typically detected through the application of a threshold on signal change measured by the difference between successive samples for each channel (e.g., Peña et al., 2003; Gervain et al., 2008). Unfortunately, this strategy cannot be applied to deep channels, for which the signal-to-noise is very low, due to high signal attenuation because of a longer path length through brain tissues. An application of the standard thresholds for deep channels would result in





removing almost all of the signal. We therefore developed a probe-based artifact rejection technique based on the consideration that movement artifacts should theoretically affect probes (i.e., poor probe attachment where a given source or detector is not close enough to the skin), rather than individual channels. Therefore, we applied artifact detection not on the individual channels signals, but to the average of all of the signals relevant for a given probe. Naturally, movement artifacts could affect multiple probes; but then they will also affect all the individual channels associated with each of those probes. As in Gervain et al. (2008) and Peña et al. (2003), we used total-Hb (previously band-pass filtered between 0.02 and 0.7 Hz). Total-Hb is used with the reasoning that a movement artifact should affect both wavelengths, and therefore both oxy and deoxy-Hb. A segment of signal was labeled as artifacted if two successive samples (separated by 100 ms because of the 10 Hz sampling rate) differed by more than 1.5 mM-mm (as in Peña et al., 2003). This threshold operates under the assumption that hemodynamic responses normally do not change more than

1.5 mM-mm in 100 ms. The channel-based movement artifacts were computed by taking the set theoretic union of the movement artifacts of the two probes (sources and detectors) defining a particular channel. Additionally, if there were less than 20 s between two regions of artifact, the intervening signal was esteemed to be too short for a proper analysis (20 s is the duration of a typical hemodynamic response from the brain) and was also coded as artifacted. In brief, this method profits from the relatively good signal-to-noise ratio of the surface channels to detect probe-based movement artifacts, and then applies the temporal definition of these artifacts to the analysis of both surface and deep channels.

Non-linear detrending was achieved through applying a general linear model (GLM) to the data of each channel, introducing sine and cosine regressors of periods between 2 min and the duration of the run (8 min), plus a boxcar regressor for each of stretch of non-artifacted data (to model possible changes in baseline after a movement artifact). The artifacted regions were silenced (that is, given a weight of zero in the regression; this is mathematically

equivalent to removing the artifactual signals prior to running the regression). The residuals of this analysis were taken as primary data on subsequent analyses.

HRF reconstruction

In order to reconstruct the hemodynamic response function for each infant and channel, a linear model was fitted with twenty 1-s boxcar regressors time-shifted by 0, 1, ..., 19 s respectively from stimulus onset. These reconstructed HRF were averaged across all channels, and compared to an adult HRF model at various delays using a linear regression. The phase was estimated as the delay that yields the best regression coefficient within the range of (−2, 6). Finally, bootstrap resampling of the individual subjects' data (Westfall and Young, 1993; $N = 10,000$) was used to generate 95% confidence intervals for this optimal phase.

Analyses of activations

Experimental effects were analyzed by introducing three regressors, one for each of the three conditions (temporal, spectral, control). Each of these regressors consisted of a boxcar on for the stimulation duration (with a value of 1) and off (with a value of zero) everywhere else, convolved with the HRF response using the optimal phase as described in the previous steps. As before, the regions of artifacts were silenced. The beta values for each channel and each conditions resulting from this GLM were tested against zero using a t -test. The t -test significance levels were corrected for multiple comparisons channel-wise using Monte Carlo resampling. This method consists in estimating the sampling distribution of the maximum of the t value across channels under the null hypothesis (t_{\max}). The null hypothesis was obtained by flipping the sign of all beta values across channels randomly for each subject, computing the t -test across subjects, and finding the maximum of this statistics for that particular simulation. This procedure was repeated 10,000 times, providing an estimate of the t_{\max} distribution, from which we derived the corrected p -values. To answer the main research question, the beta values were entered into Analyses of Variance in order to test whether the intercept was different from zero and whether there were effects of Condition (Temporal, Control, Spectral) or with Hemisphere (Left, Right). Three ANOVAs were carried out. The first included all of the channels averaged across channel position for each hemisphere and depth separately; thus, there were three within-subject factors: Condition (Temporal, Control, Spectral); Hemisphere (Left, Right); and Depth (Deep, Shallow). However, since deep and shallow channels have different light path length, we also carried out separate ANOVAs on shallow and deep channels. In the latter, there were only two within-subject factors: Condition (Temporal, Control, Spectral); and Hemisphere (Left, Right). Analyses of regions of interest (ROI) are frequently used in research on the signal-based hypothesis, and, as noted in the Introduction, these analyses highlight the role of STG, STS, and inferior parietal area. Consequently, we also report mean and SE for each of the six channels that tap those regions (3, 6, 7, 7a, 8, 10). For spatial estimation of channel location in the brain, we employed the virtual registration method (Tsuzuki et al., 2007) to map NIRS data onto the MNI standard brain space. Although this method is basically applicable to adult brains, we adapted it for

evaluation of infants' brain activity by adjusting for differences in head size and the emitter-detector separation length (inter-probe separation) between adults and neonates.

Results

As shown in **Figure 2**, the infant HRF appears to be delayed by about 2 s with respect to the adult one for the oxy-Hb signal (95% confidence interval: [0.0, 2.0]) and by about 4 s for deoxy-Hb (95% confidence interval: [1.4, 5.8]). In shallow channels, the fit between the reconstructed oxy- and deoxy-Hb and the adult HRF model was of $R = 0.97$ and $R = 0.96$, respectively. The fit remained relatively good for deep channels for oxy-Hb ($R = 0.84$); but not so for deoxy-Hb ($R = 0.62$). Given that there is a better correlation with the adult model for oxy-Hb, the rest of the analyses take into account only this variable. Results on deoxy-Hb showed a similar, but statistically weaker pattern; they are reported in **Table A1** in Appendix. As in Minagawa-Kawai et al. (2011b), we have carried out the rest of the analyses for oxy-Hb with the adult HRF model shifted by the average phase lag (2 s). Activation results in the three conditions are represented in **Figure 3** and summarized in **Table 1**. There were significant activations for the Temporal condition in two channels in the Left STG area (channels 4 and 7b, $p < 0.05$ corrected for multiple comparisons), and in five channels in homologous regions in the Right (channels 4, 4a, 7, 7a, and 8, $p < 0.05$ corrected). For the Control condition, only one channel survives multiple comparison (Right channel 5, $p < 0.05$ corrected). No channel was found to be reliably activated in the Spectral condition.

An Analysis of Variance with Hemisphere, Condition, and Depth as factors revealed a significant intercept [$F(1, 28) = 17.4$, $p < 0.001$] and an effect of Condition [$F(2, 56) = 3.6$, $p = 0.03$] but no other main effects or interactions [all other F s < 1 except for Hemisphere \times Condition \times Depth, $F(2, 56) = 1.9$, $p > 0.1$; **Figure 3**]. The main effect of Condition was due to the Temporal stimuli evoking greater activation, as evident on **Figure 3**. The separate ANOVAs suggested this effect may be stronger in shallow [$F(2, 56) = 2.7$, $p = 0.08$] than in deep [$F(2, 56) = 4.0$, $p < 0.03$] channels, but neither revealed any significant or marginal effects of or interactions with Hemisphere (all F s < 0.6 , $p > 0.55$). Moreover, even a preliminary inspection of channels selected on the basis of previously used ROIs reveals little in the way of an interaction between hemisphere and condition, as shown in **Figure 4** (a number of children did not provide data for some of these comparisons: 1 for channels 3 and 4a; 2 for channels 7a, 9, and 10). Although Channel 3 appears to follow half of the prediction (greater activation on the LH), the Condition \times Hemisphere interaction is also non-significant in an ANOVA specific to this channel [$F(2, 54) = 1.3$, $p > 0.1$]. The same ANOVA was applied to deoxy-Hb data but none of the channels showed significant main effects nor interactions after correction for multiple comparisons.

DISCUSSION

In this study, we have tested the possibility that newborns' brain activation responds to the rate of change or spectral variability of the stimuli, and if so, whether briefer, spectrally simpler segments (*Temporal*) evoke left-dominant activations, whereas longer (*Control*), spectrally more variable (*Spectral*) stimuli capture the RH to a greater extent. The answer was clear in both cases: Activation

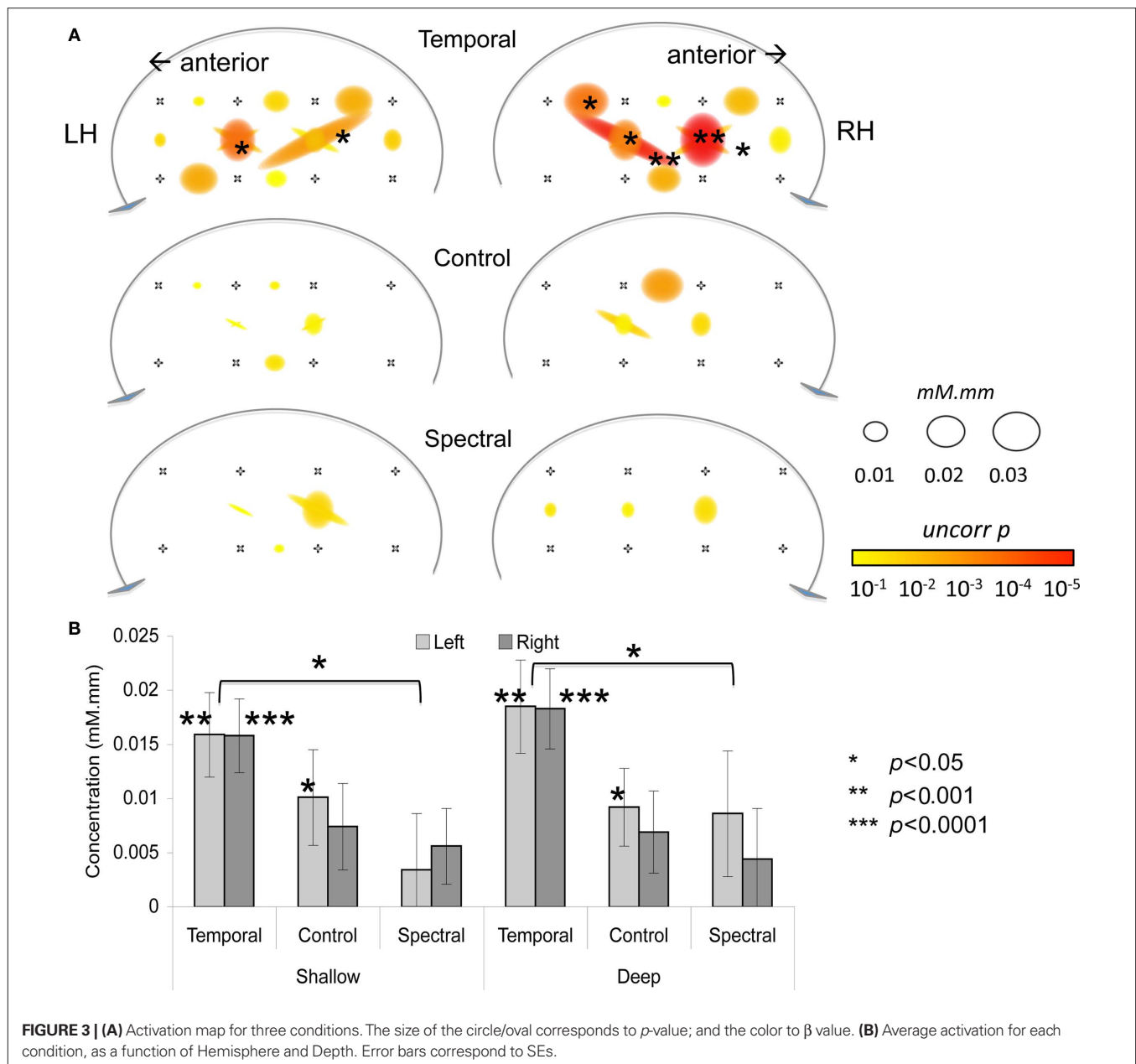


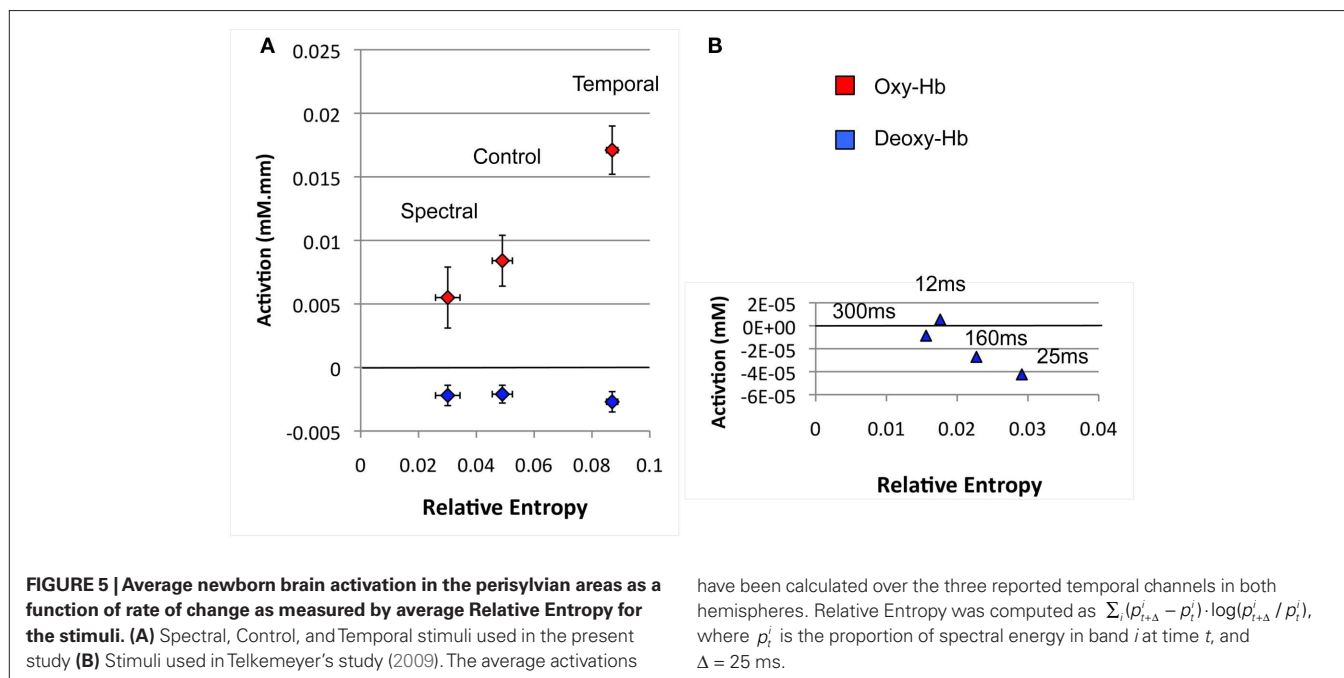
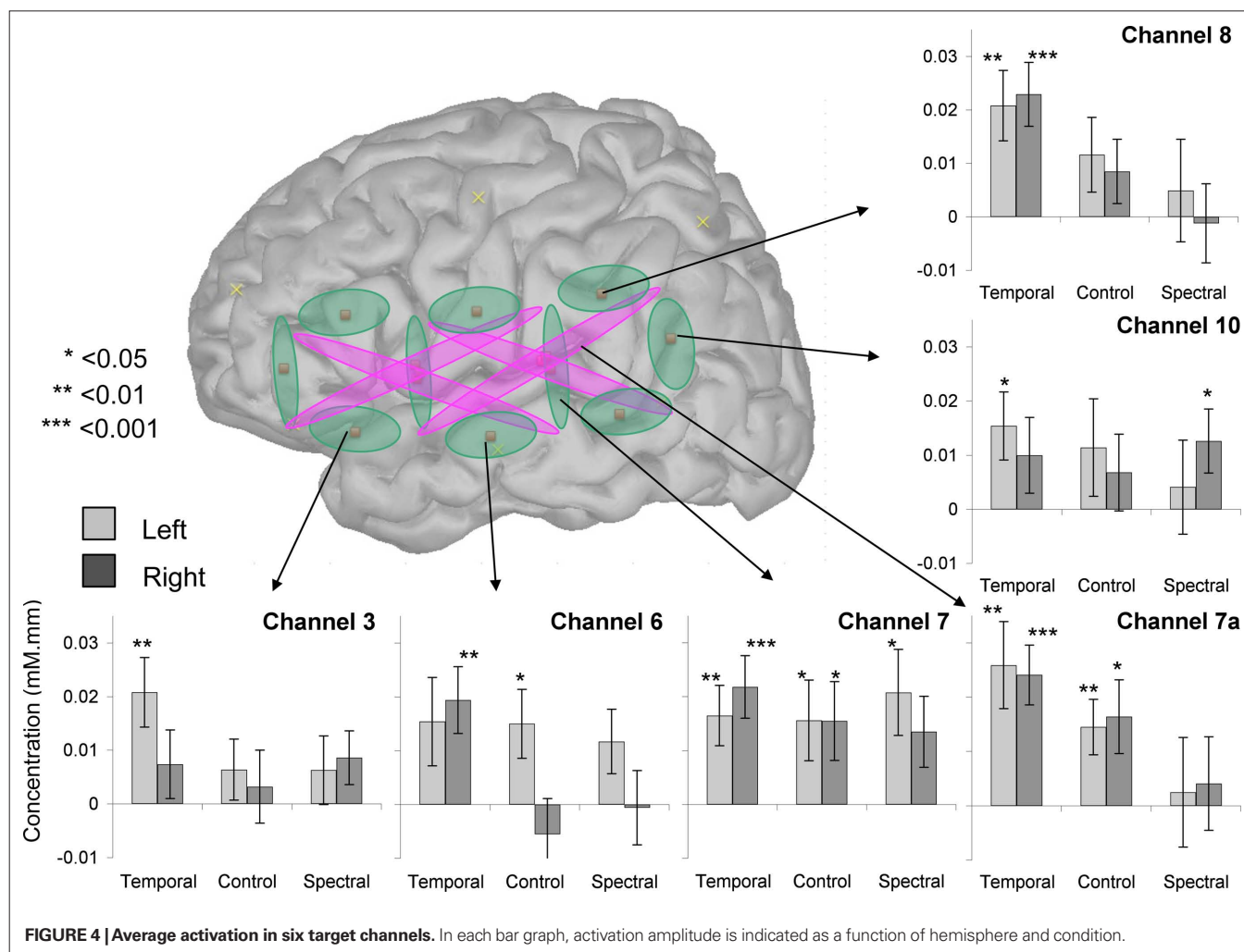
FIGURE 3 | (A) Activation map for three conditions. The size of the circle/oval corresponds to p -value; and the color to β value. **(B)** Average activation for each condition, as a function of Hemisphere and Depth. Error bars correspond to SEs.

was greater *bilaterally* for the Temporal stimulation than either Spectral or Control; but there was no evidence for hemispheric asymmetries dependent on the stimulus characteristics.

To the best of our knowledge, little attention has been devoted to overall changes in response depending on the rate of change in adults, a robust finding in the present study. This pattern had already been noted in the previous newborn study by Telkemeyer et al. (2009). In that work, five of the six channels exhibit greater activation in response to 25 ms segments than any of the other stimuli tested (Figure 4, p. 14730). One may interpret this result as a predisposition of the infant brain to respond to stimuli whose rate of change approximates that of speech, estimated at 25–35 ms by previous work (e.g., Poeppel and Hickok, 2004). In **Figure 5**, we present the average newborns' brain responses as a function of

rate of change. This rate of change was quantified as the Relative Entropy computed over adjacent 25 ms frames of a cochleogram's representation of the stimuli. Both Telkemeyer et al. (2009) and our results fit a linear increase in brain activation as a function of Relative Entropy (stimuli that change too fast, as the 12 ms stimuli of Telkemeyer have actually a Lower Relative Entropy than stimuli that change at the 25 ms rate).

Thus, these findings are in good agreement with the recently revised version of the time-resolution model proposed by Poeppel et al. (2008), where it is hypothesized that fast changing stimuli elicit larger responses which are largely symmetrical. Our data complies with this prediction, as newborns had greater bilateral activations for the temporal condition than for the other conditions. Interestingly, the same is not always found in adults. Careful



inspection of previous research results also shows greater activation for temporally complex stimuli in some cases (Jamison et al., 2006, Table 1, p. 1270: The Temporal > Spectral contrast activates 1354 voxels in HG and STG, and the Spectral > Temporal contrast, 197 voxels in STG; so at least eight times as many voxels activated for Temporal than Spectral); but not in others (Zatorre and Belin, 2001, Figure 3, p. 950: total average CBF in Anterior STG was identical for the Temporal and Spectral conditions; Boemio et al., 2005, Figure 4, p. 392, activation in STG *decreases* linearly with temporal rate). This is certainly a question that deserves further investigation with both infant and adult populations.

As for the question of asymmetries induced by the present stimuli, it is clear that the current study on newborns fails to replicate the PET (Zatorre and Belin, 2001) and fMRI (Jamison et al., 2006) adult results found with a superset of the same stimuli. However, our results line up with the previous newborn study, where the evidence for lateralization was scarce (one-tailed *t*-test on 1 channel selected among 6, $0.03 < p_{\text{uncorr}} < 0.05$; Telkemeyer et al., 2009). Given that two studies on a comparable population yielded similar results, this may indicate initially weak signal-driven asymmetries. Such a result may not be too surprising, given that the hypothesized etiologies for these asymmetries do indeed strengthen with development. Previous research has suggested three possible bases underlying hemispheric signal-driven specialization: myelination and column spread (Zatorre and Belin, 2001), structural connectivity between STS and STG (Boemio et al., 2005), and functional connectivity among auditory areas (Poeppel and Hickok, 2004). A good deal of research has documented the dramatic changes in myelination and structural connectivity as a function of development (see for example, Paus et al., 2001; Haynes et al., 2005; Dubois et al., 2006), and functional connectivity from EEG, MRI, and NIRS studies begin to draw a picture of rapidly changing, increasingly complex connections (e.g., Johnson et al., 2005; Lin et al., 2008; Gao et al., 2009; Homae et al., 2010). As Homae et al. (2010) put it, “the infant brain is not a miniature version of the adult brain but a continuously self-organizing system that forms functional regions and networks among multiple regions via short-range and long-range connectivity” (p. 4877). Efficient hemispheric specialization relies on synaptic bases for rapid inter-hemispheric communication of neural signals. This system may not be sufficiently developed in neonates. Indeed, recent work documents that long-range functional connectivity across hemispheres is not present in the temporal and parietal areas in young infants (Homae et al., 2010).

Nonetheless, there is an alternative explanation, according to which all of the necessary components for signal-based asymmetries are present and functional by full-term birth, but our study simply failed to reveal them, for example due to problems with the stimuli. To begin with, given that studies using language or music stimuli have been able to uncover hemispheric asymmetries in newborns (Peña et al., 2003; Gervain et al., 2008; Perani et al., 2010; Arimitsu et al., this volume), it could be the case that our non-speech, pure-tone stimuli fails to evoke a strong enough response that would allow the assessment of interactions of stimuli type by hemisphere. In line with this explanation, notice that in Peña et al. (2003), backward speech failed to activate auditory areas (a result replicated in Sato et al., 2006), suggesting that not

all stimuli are able to evoke activations measurable with NIRS. Moreover, one previous adult study suggests that simple tones evoke activation in deeper areas of the brain (Jamison et al., 2006). Although we have made an attempt to capture such activations by analyzing light absorption over long-distance channels, the paths traveled by this light may not have included the relevant structures. Additionally, a reviewer pointed out that, due to limitations in infants' frequency resolution, the stimuli should have been altered to provide a better match with the previous adult work. Novitski et al. (2007) have documented that newborns' ERPs did not evidence detecting a change in a sequence of 100 ms tones separated by 700 ms inter-stimulus interval when the oddball deviated from the standard by only 5%; that is, 12.5 Hz away from a 250 Hz standard or 50 Hz away from a 1000 Hz standard. However, the response was reliable when the oddball differed by 20% (50 Hz from the 250 Hz, 200 from the 1000 Hz). Newborns may not have been able to detect some of the spectral changes in the Spectral condition, since our stimuli were sampled at 32 steps between 500 and 1000 Hz. Yet, the probability of two subsequent segments being less than 5% away is only 20%; therefore, 80% of the transitions could probably be detected by the newborns. While expanding the range of variation to two octaves and sampling at larger intervals is certainly an avenue worth exploring, we remain agnostic as to the power of this strategy, given that *no channel in either hemisphere* was activated either during this condition, nor the Control condition, where there were easily detectable 500-Hz frequency changes.

Another way in which an underlying asymmetry could have escaped detection is due to technical limitations in NIRS. In addition to the problem of detecting activation below the cortical surface mentioned above, NIRS channels necessarily travel in a limited path, and due to their spacing, low-density NIRS may fail to capture particular activation foci. Indeed, as pointed out by Rutten et al. (2002), lateralization is more accurately measured by limiting the analyses to specified regions, which is not feasible with low-density NIRS as used here. The activation area captured with NIRS appears broader than it really is, and thus NIRS fails to capture strong signals from discrete activation foci.

An even larger question that also awaits further data is whether signal-driven hypotheses still stand a good chance to explain left lateralization for language. These hypotheses face a number of challenges, the first of which, as briefly stated in the Introduction, is the definition of the stimulus characteristics yielding robust lateralization patterns. Disentangling the precise physical characteristics that drive the hemispheric dominance patterns could be resolved by further experimentation with more minimalistic stimuli.

The second challenge for the explanatory power of this class of hypotheses for language relates to the fit between these distinguishing characteristics and speech itself. With the gain in precision of the implementation of the hypothesis, our work faces the loss of its explanatory power, as it is clear that speech has both short and long segments, fast and slow transitions, temporal and spectral contrasts, etc. It is, nonetheless, possible that some combination of these factors succeeds in perfectly describing language, though this is also an empirical problem, which could be addressed as a first approach by doing multiple regressions on speech with the proposed factors as regressors.

Naturally, the final challenge of these hypotheses lies not in their potential explanatory power, but in their empirical fit. On the one hand, it is clear that a substantial number of researchers have abandoned signal-driven explanations for the *left* lateralization of speech processing, and instead propose that acoustic characteristics can only modulate overall activation or *right* dominance (Poeppel et al., 2008). On the other hand, we know that, at some point, acoustic characteristics give way to linguistic characteristics, such that vowel and pitch contrasts (spectral, slow, lower frequency...) elicit strong left dominance, just as consonants do, but only if they belong to the listener's language. Zatorre and Belin (2001) have proposed that the signal driven hypothesis only provides an initial bias, but language later takes a life of its own. Indeed, we recently proposed a model of speech acquisition and brain development in which left dominance for language is the result not of a single bias, but a combination of factors whose significance is modulated by development (Minagawa-Kawai et al., 2007, 2011a). This model proposes that signal-driven biases dominate activation patterns at an early stage of speech acquisition, such that hemispheric lateralization is determined primarily by the acoustic features of the spoken stimuli. As infants gain in language experience, they begin to acquire sounds, sound patterns, word forms, and other frequent sequences, all of which involve learning mechanisms that are particularly efficient in the left hemisphere. Consequently, as linguistic categories are built, activation in response to speech is best described as domain-driven, as it results in an increased leftward lateralization exclusively for the first language categories. Although evidence on the development of lateralization continues to be sparse, current data on infants' listening to speech or non-speech are accurately captured by this model. For example, cerebral responses in young infants are bilateral to vowel contrasts (Minagawa-Kawai et al., 2007, 2009) and right-dominant or bilateral (Sato et al., 2010) for pitch accent, while in adulthood these contrasts evoke left-dominant activations in listeners whose language utilizes them, and right-dominant otherwise.

REFERENCES

- Bedoin, N., Ferragne, E., and Marsico, E. (2010). Hemispheric asymmetries depend on the phonetic feature: a dichotic study of place of articulation and voicing in French stops. *Brain Lang.* 115, 133–140.
- Belin, P., Zilbovicius, M., Crozier, S., Thivard, L., Fontaine, A., Masure, M. C., and Samson, Y. (1998). Lateralization of speech and auditory temporal processing. *J. Cogn. Neurosci.* 10, 536–540.
- Boemio, A., Fromm, S., Braun, A., and Poeppel, D. (2005). Hierarchical and asymmetric temporal sensitivity in human auditory cortices. *Nat. Neurosci.* 8, 389–395.
- Britton, B., Blumstein, S. E., Myers, E. B., and Grindrod, C. (2009). The role of spectral and durational properties on hemispheric asymmetries in vowel perception. *Neuropsychologia* 47, 1096–1106.
- Dehaene-Lambertz, G., Hertz-Pannier, L., and Dubois, J. (2006). Nature and nurture in language acquisition: anatomical and functional brain-imaging studies in infants. *Trends Neurosci.* 29, 367–373.
- Dehaene-Lambertz, G., Montavont, A., Jobert, A., Alliol, L., Dubois, J., Hertz-Pannier, L., and Dehaene, S. (2010). Language or music, mother or Mozart? Structural and environmental influences on infants' language networks. *Brain Lang.* 114, 53–65.
- Dubois, J., Hertz-Pannier, L., Dehaene-Lambertz, G., Cointepas, Y., and Le Bihan, D. (2006). Assessment of the early organization and maturation of infants' cerebral white matter fiber bundles: a feasibility study using quantitative diffusion tensor imaging and tractography. *Neuroimage* 30, 1121–1132.
- Elman, J., Bates, E., Johnson, M., Karmiloff-Smith, A., Parisi, D., and Plunkett, K. (1997). *Rethinking Immateness: A Connectionist Perspective on Development*. Cambridge, MA: MIT Press.
- Everdell, N. L., Gibson, A. P., Tullis, I. D. C., Vaithianathan, T., Hebden, J. C., and Delpy, D. T. (2005). A frequency multiplexed near-infrared topography system for imaging functional activation in the brain. *Rev. Sci. Instrum.* 76, 093705.
- Fukui, Y., Ajichi, Y., and Okada, E. (2003). Monte Carlo prediction of near-infrared light propagation in realistic adult and neonatal head models. *Appl. Opt.* 42, 2881–2887.
- Gao, W., Zhu, H., Giovanello, K. S., Smith, J. K., Shen, D., Gilmore, J. H., and Lin, W. (2009). Evidence on the emergence of the brain's default network from 2-week-old to 2-year-old healthy pediatric subjects. *Proc. Natl. Acad. Sci. USA* 106, 6790–6795.
- Gervain, J., Macagno, F., Cogoi, S., Pena, M., and Mehler, J. (2008). The neonate brain detects speech structure. *Proc. Natl. Acad. Sci. U.S.A.* 105, 14222–14227.
- Giraud, A. L., Kleinschmidt, A., Poeppel, D., Lund, T. E., Frackowiak, R. S., and Laufs, H. (2007). Endogenous cortical rhythms determine cerebral specialization for speech perception and production. *Neuron* 56, 1127–1134.
- Haynes, R. L., Borenstein, N. S., Desilva, T. M., Folkerth, R. D., Liu, L. G., Volpe, J. J., and Kinney, H. C. (2005). Axonal development in the cerebral white matter of the human fetus and infant. *J. Comp. Neurol.* 484, 156–167.
- Hickok, G., and Poeppel, D. (2007). The cortical organization of speech processing. *Nat. Rev. Neurosci.* 8, 393–402.
- Homae, F., Watanabe, H., Nakano, T., Asakawa, K., and Taga, G. (2006). The right hemisphere of sleeping infant perceives sentential prosody. *Neurosci. Res.* 54, 276–280.
- Homae, F., Watanabe, H., Otobe, T., Nakano, T., Go, T., Konishi, Y., and Taga, G. (2010). Development of global cortical networks in early infancy. *J. Neurosci.* 30, 4877–4882.
- Jamison, H. L., Watkins, K. E., Bishop, D. V., and Matthews, P. M. (2006).

We assume that it is through the learning process that the cerebral networks that process native vowel contrast shift from the bilateral, signal-driven pattern to a left dominant, linguistically-driven processing.

CONCLUSION

In brief summary, this study aimed at providing additional evidence on the modulation of newborns' brain activation in response to physical characteristics of auditory stimulations. While these data provided further support for an increased responsiveness for change rates similar to those found in speech, there was no evidence of hemispheric asymmetries. This may indicate that the functional organizations responsible for such asymmetries in the adult are not in place by birth, but other explanations are still open, and further data, both in the way of replications and extensions to other age groups, are needed before drawing any definite conclusions.

ACKNOWLEDGMENTS

This work was supported by EU's sixth frame work program (neuronal origins of language and communication: NEUROCOM, Project no. 012738), in part by a grant from the Agence Nationale pour la Recherche (ANR-09-BLAN-0327 SOCODEV), Grant-in-Aid for Scientific Research (A) (Project no. 21682002), Academic Frontier Project supported by Ministry of Education, Culture, Sports, Science and Technology (MEXT) and funding from the Ecole de Neurosciences de Paris, and the Fyssen Foundation. The authors thank R. Zatorre for use of the stimuli, J. Hebden, and N. Everdell for their comments and help with the NIRS system, K. Friston for illuminating suggestions regarding data analysis and H. van der Lely, A. Shestakova, E. Kushnerenko, and J. Meek for their support of this earlier version of this study. We thank L. Filippin for help with the acquisition and analysis software; A. Bachmann for designing and constructing the probe pad; I. Brunet, A. Elgellab, J. Gervain, and S. Margules for assistance with the infant recruitment and test.

- Hemispheric specialization for processing auditory nonspeech stimuli. *Cereb. Cortex* 16, 1266–1275.
- Johnson, M. H., Griffin, R., Cibra, G., Halit, H., Farroni, T., de Haan, M., Tucker, L. A., Baron-Cohen, S., and Richards, J. (2005). The emergence of the social brain network: evidence from typical and typical development. *Dev. Psychopathol.* 17, 599–619.
- Lin, W., Zhu, Q., Gao, W., Chen, Y., Toh, C.-H., Styner, M., Gerig, G., Smith, J.K., Biswal, B., and Gilmire, J.H. (2008). Functional connectivity MR imaging reveals cortical functional connectivity in the developing brain. *Am. J. Neuroradiol.* 29, 1883–1889.
- Lloyd-Fox, S., Blasi, A., and Elwell, C. E. (2010). Illuminating the developing brain: the past, present and future of functional near infrared spectroscopy. *Neurosci. Biobehav. Rev.* 34, 269–284.
- Minagawa-Kawai, Y., Cristià, A., and Dupoux, E. (2011a). Cerebral lateralization and early speech acquisition: developmental scenario. *J. Dev. Cogn. Neurosci.* 1, 217–232.
- Minagawa-Kawai, Y., van der Lely, H., Ramus, F., Sato, Y., Mazuka, R., and Dupoux, E. (2011b). Optical brain imaging reveals general auditory and language-specific processing in early infant development. *Cereb. Cortex* 21, 254–261.
- Minagawa-Kawai, Y., Mori, K., Hebden, J. C., and Dupoux, E. (2008). Optical imaging of infants' neurocognitive development: recent advances and perspectives. *Dev. Neurobiol.* 68, 712–728.
- Minagawa-Kawai, Y., Mori, K., Naoi, N., and Kojima, S. (2007). Neural attunement processes in infants during the acquisition of a language-specific phonemic contrast. *J. Neurosci.* 27, 315–321.
- Minagawa-Kawai, Y., Naoi, N., and Kojima, S. (2009). *New Approach to Functional Neuroimaging: Near Infrared Spectroscopy*. Tokyo: Keio University Press.
- Novitski, N., Huottilainen, M., Tervaniemi, M., Näätänen, R., and Fellman, V. (2007). Neonatal frequency discrimination in 250–4000Hz range: electrophysiological evidence. *Clin. Neurophysiol.* 118, 412–419.
- Obrig, H., Rossi, S., Telkemeyer, S., and Wartenburger, I. (2010). From acoustic segmentation to language processing: evidence from optical imaging. *Front. Neuroener.* 2:13. doi: 10.3389/fnene.2010.00013
- Obrig, H., and Villringer, A. (2003). Beyond the visible – imaging the human brain with light. *J. Cereb. Blood Flow Metab.* 23, 1–18.
- Paus, T., Collins, D. L., Evans, A. C., Leonard, G., Pike, B., and Zijdenbos, A. (2001). Maturation of white matter in the human brain: a review of magnetic resonance studies. *Brain Res. Bull.* 54, 255–266.
- Peña, M., Maki, A., Kovacic, D., Dehaene-Lambertz, G., Koizumi, H., Bouquet, F., and Mehler, J. (2003). Sounds and silence: an optical topography study of language recognition at birth. *Proc. Natl. Acad. Sci. U.S.A.* 100, 11702–11705.
- Perani, D., Saccuman, M.C., Scifo, P., Spada, D., Andreolli, G., Rovelli, R., Baldoli, C., and Koelsch, S. (2010). Functional specializations for music processing in the human newborn brain. *Proc. Natl. Acad. Sci. USA* 107, 4758–4763.
- Poeppl, D., Guillemin, A., Thompson, J., Fritz, J., Bavelier, D., and Braun, A. R. (2003). Auditory lexical decision, categorical perception, and FM direction discrimination differentially engage left and right auditory cortex. *Neuropsychologia* 42, 183–200.
- Poeppl, D., and Hickok, G. (2004). Towards a new functional anatomy of language. *Cognition* 92, 1–12.
- Poeppl, D., Idsardi, W. J., and van Wassenhove, V. (2008). Speech perception at the interface of neurobiology and linguistics. *Philos. Trans. R. Soc. Lond. Biol. Sci.* 363, 1071–1086.
- Rutten, G. J. M., Ramsey, N. F., van Rijen, P. C., and van Veelen, C. W. M. (2002). Reproducibility of fMRI-determined language lateralization in individual subjects. *Brain Lang.* 80, 421–437.
- Rybalko, N., Suta, D., Nwabueze-Ogbo, F., and Syka, J. (2006). Effect of auditory cortex lesions on the discrimination of frequency-modulated tones in rats. *Eur. J. Neurosci.* 23, 1614–1622.
- Sato, H., Hirabayashi, Y., Tsubokura, S., Kanai, M., Ashida, S., Konishi, I., Uchida, M., Hasegawa, T., Konishi, Y., and Maki, A. (2006). Cortical activation in newborns while listening to sounds of mother tongue and foreign language: an optical topography study. *Proc. Intl. Conf. Inf. Study*, 037–070.
- Sato, Y., Sogabe, Y., and Mazuka, R. (2010). Development of hemispheric specialization for lexical pitch-accent in Japanese infants. *J. Cogn. Neurosci.* 22, 2503–2513.
- Schonwiesner, M., Rubsamen, R., and von Cramon, D. Y. (2005). Hemispheric asymmetry for spectral and temporal processing in the human antero-lateral auditory belt cortex. *Eur. J. Neurosci.* 22, 1521–1528.
- Telkemeyer, S., Rossi, S., Koch, S. P., Nierhaus, T., Steinbrink, J., Poeppl, D., Obrig, H., and Wartenburger, I. (2009). Sensitivity of newborn auditory cortex to the temporal structure of sounds. *J. Neurosci.* 29, 14726–14733.
- Tsuzuki, D., Jurcak, V., Singh, A. K., Okamoto, M., Watanabe, E., and Dan, I. (2007). Virtual spatial registration of stand-alone fNIRS data to MNI space. *Neuroimage* 34, 1506–1518.
- Westfall, P. H., and Young, S.S. (1993). *Resampling-based multiple testing*. New York, NY: John Wiley & Co.
- Wetzel, W., Ohl, F. W., Wagner, T., and Scheich, H. (1998). Right auditory cortex lesion in Mongolian gerbils impairs discrimination of rising and falling frequency-modulated tones. *Neurosci. Lett.* 252, 115–118.
- Zatorre, R. J., and Belin, P. (2001). Spectral and temporal processing in human auditory cortex. *Cereb. Cortex* 11, 946–953.

Conflict of Interest Statement: The authors declare that this manuscript was prepared in the absence of any commercial and financial relationships that could be construed as a potential conflict of interests.

Received: 21 January 2011; accepted: 07 June 2011; published online: 16 June 2011.
 Citation: Minagawa-Kawai Y, Cristià A, Vendelin I, Cabrol D and Dupoux E (2011) Assessing signal-driven mechanisms in neonates: brain responses to temporally and spectrally different sounds. *Front. Psychology* 2:135. doi: 10.3389/fpsyg.2011.00135
 This article was submitted to *Frontiers in Language Sciences*, a specialty of *Frontiers in Psychology*.
 Copyright © 2011 Minagawa-Kawai, Cristià, Vendelin, Cabrol and Dupoux. This is an open-access article subject to a non-exclusive license between the authors and Frontiers Media SA, which permits use, distribution and reproduction in other forums, provided the original authors and source are credited and other Frontiers conditions are complied with.

APPENDIX

Table A1 | Levels of activations (deoxy-Hb changes) for temporal, control and spectral conditions.

Channel	Temporal					Control					Spectral				
	β	T	N	p (unc)	p (corr)	β	T	N	p (unc)	p (corr)	β	T	N	p (unc)	p (corr)
LEFT SHALLOW															
2 IFG	−0.0080	−2.14	29	0.044	ns										
3 aSTG											−0.0051	−2.39	29	0.027	ns
4 STG/IFG	−0.0067	−2.14	29	0.044	ns	−0.0057	−2.01	29	0.056	ns	−0.0050	−1.73	29	0.091	ns
5 precentral	−0.0061	−1.90	29	0.068	ns										
7 STG						−0.0064	−2.25	29	0.035	ns					
9 pSTG/STS						−0.0054	−2.19	29	0.040	ns	−0.0058	−2.81	28	0.011	ns
LEFT DEEP															
4b FG/ postcent postcentral	−0.0047	−2.80	28	0.011	ns										
7b/TG/FG	−0.0032	−1.84	29	0.076	ns						−0.0045	−2.69	28	0.014	ns
RIGHT SHALLOW															
1 MFG/IFG						−0.0052	−1.72	28	0.092	ns					
2 IFG						−0.0054	−2.48	29	0.022	ns					
3 aSTG	−0.0054	−2.06	29	0.051	ns	−0.0046	−1.94	29	0.063	ns					
4 STG/IFG	−0.0068	−2.47	29	0.023	ns	−0.0050	−1.92	29	0.066	ns					
5 Precentral	−0.0045	−1.78	29	0.083	ns										
6 STG/STS	−0.0054	−2.54	29	0.019	ns	−0.0063	−2.39	29	0.027	ns	−0.0081	−3.12	29	0.005	~
9 pSTG/STS	−0.0065	−2.44	29	0.024	ns										

N, number of children contributing data for that channel and condition; *unc*, uncorrected; *corr*, corrected; *MFG*, middle frontal gyrus; *IFG*, inferior frontal gyrus; *a*, anterior; *STG*, superior temporal gyrus; *STS*, superior temporal sulcus; *p*, posterior; *MTG*, middle temporal gyrus; *FG*, frontal gyrus; *TG*, temporal gyrus.



Functional hemispheric specialization in processing phonemic and prosodic auditory changes in neonates

Takeshi Arimitsu¹, Mariko Uchida-Ota^{2,3}, Tatsuhiko Yagihashi^{1,2}, Shozo Kojima^{2,4}, Shigeru Watanabe^{2,5}, Isamu Hokuto¹, Kazushige Ikeda^{1,2}, Takao Takahashi^{1,2} and Yasuyo Minagawa-Kawai^{2,5*}

¹ Department of Pediatrics, School of Medicine, Keio University, Tokyo, Japan

² Global COE program, Centre for Advanced Research on Logic and Sensibility, Keio University, Tokyo, Japan

³ Keio Advanced Research Center, Keio University, Tokyo, Japan

⁴ Academic Frontiers Project, Meijo University, Nagoya, Japan

⁵ Graduate School of Human Relations, Keio University, Tokyo, Japan

Edited by:

Judit Gervain, CNRS – Université Paris Descartes, France

Reviewed by:

Janet F. Werker, University of British Columbia, Canada

Ho Henny Yeung, Centre National de la Recherche Scientifique, France

*Correspondence:

Yasuyo Minagawa-Kawai, Graduate School of Human Relations, Keio University, 2-15-45 Mita, Minato-ku, Tokyo 108-8345, Japan.
e-mail: myasuyo@bea.hi-ho.ne.jp

This study focuses on the early cerebral base of speech perception by examining functional lateralization in neonates for processing segmental and suprasegmental features of speech. For this purpose, auditory evoked responses of full-term neonates to phonemic and prosodic contrasts were measured in their temporal area and part of the frontal and parietal areas using near-infrared spectroscopy (NIRS). Stimuli used here were phonemic contrast /itta/ and /itte/ and prosodic contrast of declarative and interrogative forms /itta/ and /itta?/. The results showed clear hemodynamic responses to both phonemic and prosodic changes in the temporal areas and part of the parietal and frontal regions. In particular, significantly higher hemoglobin (Hb) changes were observed for the prosodic change in the right temporal area than for that in the left one, whereas Hb responses to the vowel change were similarly elicited in bilateral temporal areas. However, Hb responses to the vowel contrast were asymmetrical in the parietal area (around supra marginal gyrus), with stronger activation in the left. These results suggest a specialized function of the right hemisphere in prosody processing, which is already present in neonates. The parietal activities during phonemic processing were discussed in relation to verbal-auditory short-term memory. On the basis of this study and previous studies on older infants, the developmental process of functional lateralization from birth to 2 years of age for vowel and prosody was summarized.

Keywords: phoneme, prosody, functional lateralization, neonates, NIRS, auditory area

INTRODUCTION

Speech consists of two dominant components, i.e., segments and suprasegments, which correspond respectively to phonemic and prosodic levels of structure. Although language comprehension involves various processes, the perceptual analysis of segmental and suprasegmental information constitutes a crucial first step in the overall process of successful encoding of lexical, syntactic and pragmatic levels. Indeed, it is well known that learning specific features that are associated respectively with these two components is an important initial step for language acquisition in the first year of life. Functional cerebral lateralization in processing these two kinds of information has been demonstrated in neuroimaging literature on adult speech perception: human adults tend to show a left hemispheric dominance for processing phonemes and a right hemispheric dominance for processing prosodic information (e.g., Zatorre et al., 1992; Furuya and Mori, 2003). However, brain development of this specialized system in infants remains poorly understood, in spite of the fact that developmental studies offer the potential for uncovering critical clues to understanding the cerebral basis of linguistic skill acquisition. Accordingly, the present study is designed to investigate brain lateralization in neonates with the aim of determining the degree of hemispheric specialization of segments and suprasegments in early infancy.

Phonetic cues, which are characterized by formant patterns, determine the phonological status of various linguistic segments (e.g., vowel or consonant type). In contrast, prosodic cues, which are realized by pitch contour, intensity, and duration, determine suprasegmental linguistic information. Prosodic cues contribute to accentuation and intonation and also convey para/non-linguistic information, such as emotional state and talker identity. They can affect single segments as well as whole syllables/sentences. Furthermore, many phonological theories hypothesize that segments and prosodies are separately represented in different levels (e.g., Goldsmith, 1990). In general, a differential pattern of acquisition processes is observed for segments and suprasegments, or phoneme and prosody. With respect to segments, infants are born with the ability to discriminate among a wide range of phonological features of segments. They become sensitive to native phonetic (formant) patterns only after 6 months of age (Kuhl, 2004). In other words, the perceptual ability to differentiate segments is universal initially, but with maturation and exposure to the maternal language, perceptual sensitivity narrows gradually to exhibit language specificity, which appears at around 6 months of age for vowels and 12 months of age for consonants (Kuhl, 2004). In many cases, such language-specific learning starts earlier for the sentence/phrase-level prosody than for segments (Mehler et al., 1988; Nazzi et al.,

2000). Furthermore, the prosodic organization of speech facilitates language acquisition in infants, and the acoustic saliency of prosody, even at the syllable-level (e.g., stressed syllable), may help draw the attention of infants to speech and its structures (Juszyk et al., 1999). Among such suprasegmentals, acquisition of lexical tones in Mandarin, which is a syllable-level prosody, develops in a manner relatively similar way as that of segmental categories (Mattock and Burnham, 2006; Mattock et al., 2008), whereas another syllable-level prosody, that is, Japanese pitch accent, shows quite different developmental patterns (Mugitani, 2009). The present study attempts to compare the differences in brain responses to segments and suprasegmentals, which have differential acoustic and linguistic natures, as reviewed here. Either stimuli used in the study are well controlled such that each segmental and suprasegmental difference is realized within a final syllable with the least acoustical manipulation. Consequently, we use intonation contour, which is syllable-level prosody, as suprasegmental stimuli.

In adults, speech-processing involves functional hemispheric specialization. Specifically, hemispheric specialization with adult speech is influenced by at least two factors: acoustic and linguistic sound properties. The evidence for the importance of acoustic properties of speech signal derives from a growing body of research indicating that different auditory features activate the two hemispheres. In particular, when acoustic information (e.g., spectral frequency changes) is modulated over time, rapid modulations appear to predominantly activate the left hemisphere, whereas slow and/or spectral modulations show cortical activity lateralized to the right hemisphere (Zatorre and Belin, 2001; Poeppel, 2003; Poeppel et al., 2008; note that the revised model of Poeppel hypothesizes bilateral engagement for fast stimuli). Here, we refer to this notion as signal-driven hemispheric activation or signal-driven hypothesis. Various types of acoustic definitions and classifications can explain such signal-driven asymmetry (Minagawa-Kawai et al., 2011a for a review). However, among them, two dominant trends seem to be a dichotomy of temporal versus spectral changes (Zatorre and Belin, 2001; Schönwiesner et al., 2005; Jamison et al., 2006) and stimuli of short versus long time scale (high versus low frequencies), which are processed by a different size of analysis window (Poeppel, 2003; Poeppel et al., 2008). The difference between the two is that the former emphasizes spectral richness to evoke rightward dominance, whereas the latter focuses on time scale of sounds. There are other variations, and because of such different definitions of critical acoustic properties, the definition of the dichotomy has also varied, such that what is deemed “fast/temporal” or “slow/spectral” often depends upon the experimental variables within a given study. Dichotic listening studies in adults have also confirmed hemispheric asymmetry and the dependence of this asymmetry upon acoustic properties of the speech signal (see Shtyrov et al., 2000 for a review). Although these studies lack a common set of definitions for relevant acoustic properties, they nonetheless, illustrate that signal-driven laterality is a crucial issue in investigating the cerebral basis of speech.

In light of the signal-driven hypothesis, hemispheric specialization derives from the acoustic features of segments as well as suprasegments. In phonemes, these features involve the richness of temporal variations, whereas in prosody, they are associated with frequency modulations such as $F0$ or richness of spectral

features. More precisely, phonemes may be further divided according to their physical properties, such as consonants having rapid and dynamic spectral features and vowels having rather steady-state spectral features. Hence, in the realm of segments, it has been claimed that consonants, which have more rapidly changing acoustic energies than vowels, tend to show leftward dominance in contrast to vowels, with less rapidly changing acoustic energies likely to exhibit bilateral cortical engagements (Shankweiler and Studdert-Kennedy, 1967; Weiss and House, 1973). In contrast to these segmental properties, prosodic changes can be described as slowly changing stimuli or spectrally rich stimuli as in tonal changes. These tend to be localized in the right hemisphere. Our segmental versus suprasegmental stimuli can be generally interpreted in the context of the signal-driven hypothesis. According to Poeppel (2003) and Poeppel et al. (2008), variations of formant transitions are preferentially processed on the left side, and pitch contours requiring higher spectral resolution are on the right side. However, the phonetic stimuli used here involve a vowel contrast that does not exhibit prominent rapid acoustic changes (i.e., relative to changes present in consonants). Consequently, these stimuli are characterized by steady-state formant frequencies. This acoustic property should induce bilateral activity in the temporal areas according to the signal-driven hypothesis. Furthermore, even though our prosodic stimulus is not long unlike general sentential prosody, it has richer spectral features than that of the phonemic contrast, which tends to induce the right dominance (Zatorre and Belin, 2001; Schönwiesner et al., 2005). More specifically, in contrast to the phonemic change with only $F1$ and $F2$ differences, the prosodic contrast involves more complex spectral changes as a result of manipulation of fundamental frequency affecting all the harmonic structures. This complex spectral change is likely to induce the rightward dominance.

Although it is generally agreed that in adults these acoustic-physical factors drive laterality, probably at the lower processing level, i.e., perceptual level, higher level factors (i.e., cognitive) related to linguistic knowledge also play a crucial role in explaining cerebral specialization. For instance, in adults leftward lateralization depends upon whether a particular stimulus is perceived as a linguistic element (Dehaene-Lambertz et al., 2005; Mottonen et al., 2006); similarly, if a vowel contrast is phonemically distinctive in a listeners native language, this also enhances chances that it will be lateralized leftward (Näätänen et al., 1997; Dehaene-Lambertz and Gliga, 2004). Even at the phonetic level, native phonological contrasts associated with vowels, consonants, phonotactics, and accents (but not non-native ones) are generally processed in left temporal regions (Jacquemot et al., 2003; Sato et al., 2007). Even though pitch accents or lexical tones have a slowly changing signal, they are processed predominantly on the left hemisphere by native-listeners (Gandour et al., 2000). These findings suggest that language learning is also a critical consideration for the left-dominant brain network. In short, evidence collected from adult listeners is best explained by a combination of acoustic features, linguistic, and learning factors (Minagawa-Kawai et al., 2011a). Consequently, several speech-processing models hypothesize that, to a large extent, lateralization of sounds depends upon the level of processing involved (Poeppel, 2003; Friederici and Alter, 2004; Zatorre and Gandour, 2008).

What exactly is the developmental process that leads to the functional hemispheric specialization in speech? In recent years multi-channel near-infrared spectroscopy (NIRS) has enabled examination of this issue because this methodology allows reliable localization of the focus of neural activity. In fact, recent NIRS studies have provided evidence regarding the cerebral response of infants to phonological contrasts. Minagawa-Kawai et al. (2007) compared the neural sensitivities of different age groups (five groups from 3- to 28-month-olds) to changes in phonemic category of long and short vowels and found that Japanese infants show a left-dominant temporal response to an across-category phonemic change only after 13 months of age. Similarly, NIRS analyses show that 10-month-old infants exhibit a left-lateralized cerebral response to a difference in lexical pitch accents (Sato et al., 2010). Because younger age groups in these studies did not show a left-dominant response, Sato et al. (2010) hypothesized that exposure of infants to first language (L1) modified the cortex of older infants through the construction of an L1-specific brain network that is located predominantly on the left side. In addition, electroencephalography (EEG) studies have shown emergence of a language-specific brain response after L1 exposure (Cheour et al., 1998), and recent NIRS studies revealed for the first time a developmental change in cerebral lateralization by showing the specific brain regions involved (Minagawa-Kawai et al., 2007; Sato et al., 2010).

Of special relevance to the present study is the research of Sato et al. (2003). These researchers assessed cerebral lateralization for both prosodic and phonemic contrasts using different age groups, ranging in age from 7 months to 5-years. Infants older than 11–12 months showed a significant lateralization that resembled that of adults in that the phonemic changes evoked a left-dominant response whereas prosodic contrasts evoked a right-dominant response. By contrast, for younger children (7–8; 9–10 months), hemispheric laterality indices for phonemic and prosodic conditions did not differ significantly (Sato et al., 2003). Although these results appear to indicate that brain regions required for decoding phonemic and prosodic information become more specific with maturation, detailed inspection of the laterality index (LI) in this study revealed tendencies in younger age groups toward right-dominance lateralization for the prosodic condition and a bilateral response for the phonemic condition. **Figure 1** shows these data. Note that the LI for younger age groups in the prosodic condition trends downward, below zero, indicating right hemispheric dominance. Sato et al. (2003) statistically concentrated upon the overall LI difference between the two stimulus conditions. However, on inspection we found that for the youngest group the laterality index in the prosodic condition was significantly below zero (i.e., zero indicating null hemispheric bias). This result suggests that the prosodic sensitivity of infants is already functionally specialized hemispherically by the age of 7- to 8-months-old. Furthermore, recent evidence based upon neonates' responses to presentations of frequency modulated non-speech sequences demonstrated a rightward dominance with spectral patterns having relatively slow (prosodic-like) modulations (Telkemeyer et al., 2009). These results suggest predominant right-hemisphere engagement in processing prosody from the beginning of life. To date, however, no study has investigated the inborn cerebral basis for processing prosody in real speech stimuli.

The present study is designed to examine this issue by contrasting two distinctive linguistic features (i.e., phonemic and prosodic contrasts) using real speech materials. To this end, this research employs speech materials used in previous studies (Furuya and Mori, 2003; Sato et al., 2003) in which different age groups including infants, children, and adults were examined. This paradigm enables an assessment of laterality for segments and suprasegments in newborn infants who have not been significantly exposed to language. Furthermore, comparisons of data from this study with that of Sato et al. (2003) will provide a broader perspective on developmental changes in the functional laterality in human infants as a function of age. This study also allows an indirect examination of the neonates' cortical basis for processing auditory stimuli containing fast and slow/spectrally rich acoustic changes similar to those which activate adults brain asymmetrically. However, as stated before the phonemic stimuli used here is a vowel contrast characterized by steady-state formant frequencies, which is expected to induce bilateral activity in the temporal areas according to the signal-driven hypothesis.

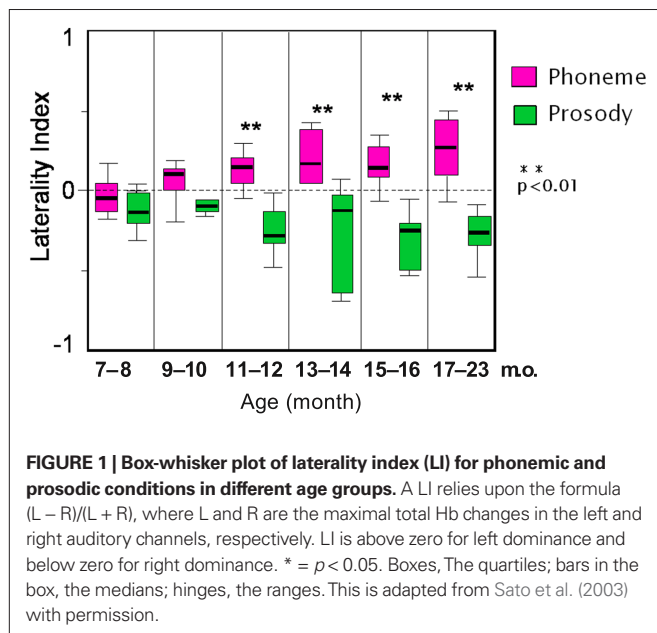
MATERIALS AND METHODS

PARTICIPANTS

Twenty Japanese neonates were tested with NIRS; four infants did not complete the protocol due to fussiness and excess movement; their data were excluded from further analyses. The final data set included data from 17 infants (average 4.8 days-old, range 3–8 days; 10 females). Among them, three infants failed to complete the phonemic condition and other two infants failed the prosodic condition, therefore the data set for each condition has different sets of participants ($N = 14$ for the phonemic condition and $N = 15$ for the prosodic condition). All neonates were full-term infants (averaged gestation: 271 days) with average birth weight of 2754 g (range: 1928–3298 g) and with no history of medical problems. All were from monolingual Japanese families. Consent forms were obtained from parents before the infants' participation. This study was approved by both of the ethic committees of Faculty of letters, Keio University (No. 09049), and Keio University hospital (No. 2009-189).

STIMULI AND CONDITIONS

Stimuli consisted of speech contexts, supplied by real words, which exhibited phonemic and prosodic differences. Three different stimulus patterns reflected respectively different forms of the Japanese verb /iku/(go); these were: An affirmative form /itta/(* has/have gone, * can be any subject), an imperative form /itte/(go away), and an interrogative form /itta?/(has/have * gone? Imaizumi et al., 1998). All stimuli were synthesized using ASL (Kay Elemetrics Corp., USA), an analysis-by-synthesis system based upon a speech signal produced by a male adult. Spectrograms of the stimuli are shown in **Figure 2**. Infant-directed speech was not used in the recording. The three stimuli have identical first syllables, and differ only in their final syllables. The duration of the first syllable /i/ is 80 ms followed by 200 ms of silent interval for geminate consonant /tt/ and the final vowel with the length of 80 ms. The phonemic contrast, consisting of pair members /iita/ versus /itte/, is based upon differences in the final vowel due to manipulation of formants 1 and 2 and their transitions; however both syllables have identical

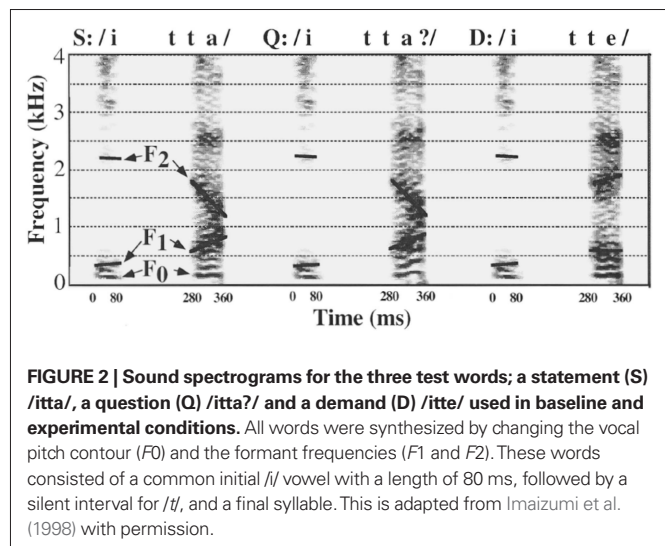


fundamental frequencies. Members of the prosodic contrasting pair /iita/ versus /itta?/ differ in pitch contours due to the manipulation of the fundamental frequency (F_0); specifically the interrogative form has a rising pitch on the final syllable, whereas the affirmative form has a slightly falling pitch on the last syllable (Figure 2).

Two main experimental conditions were: phonemic contrast and prosodic contrast. These were administered to respectively different groups of participants. Participants in both conditions received an identical baseline block of trials. In the phonemic condition, the stimulus /itta/ was repeated at 1-s intervals (trials) for a total of 15 s in the baseline block without any temporal variations; this block of trials was followed by another 15 s of presentations (trials) in the target block. In phonemic target block /itte/ and /itta/ were presented in a pseudo-random order at 1-s. intervals. In the prosodic condition, the same baseline condition was initially presented but it was followed a target block comprising a series presentations of /itta/ and /itta?/ randomized as in phonemic condition. The two blocks (baseline and target blocks) in each condition were alternated at least seven times for each condition. Presentation order of the two conditions was counterbalanced. Thus, as indicated above, the baseline for evaluating responses in experimental conditions was not silence but repetitions of the /itta/ stimulus that last 15 s. This use of non-silent baseline stimuli allowed us to extract those brain response components specific to differences in /a/ versus /e/ or to different pitch contours in each condition.

PROCEDURE

Near-infrared spectroscopy experiments were performed in a testing room at Keio University hospital. Evoked auditory responses in bilateral temporal area as well as a part of frontal and parietal regions were recorded using NIRS (ETG 4000, Hitachi Medical Corporation, Tokyo, Japan). This device emits 695 and 850 nm near-infrared lasers modulated at different frequencies and detects them with lock-in amplifiers to measure changes in the concentration and oxygenation of hemoglobin (Hb; Yamashita



et al., 1996). The recording channels resided in the optical path of the brain between the nearest pairs of incident and detection probes which were separated by 2 cm on the scalp surface. A silicon pad with five incident and four detection probes, arranged in 3×3 square lattice, was placed laterally on each side of the head. The total number of recording channels on each side was 12. The pad was attached to the head such that the center detector probe in the bottom of horizontal probe-line corresponded to the T3 or T5 position in the international 10/20 system. The bottom horizontal line of the probes was roughly aligned with the T3–Fp1–Fp2–T5 line. Stimuli were presented to neonates with amplitudes of approximately 67 dB via two speakers positioned 20–25 cm above from the infants' head. To prevent NIRS artifacts due to systemic vascular changes such as heart rate change and/or background sound changes, the stimulus sound levels were set relatively low. During the stimulation, the newborns were sleeping.

DATA ANALYSIS

Our analysis method consisted of two parts which involved, respectively, multiple channel analyses and analysis of cortical region of interest (ROI). Because previous NIRS studies on phoneme perception focused only upon the temporal area, this investigation used NIRS to widen the focus to include other brain regions which might be involved in early phonetic processing. First, we analyzed each channel separately to gauge localized activation levels. Channels showing strong activations were then compared with contra-lateral channel counterparts to assess laterality. Next, the ROI of the temporal region, determined according to the previous NIRS studies, was tested to assess the laterality effect.

Concentrations of oxygenated and deoxygenated Hb were calculated from the absorption of 695 and 830 nm laser beams sampled at 10 Hz, and smoothed with a 5-s moving average. Blocks of trials affected by movement artifacts were automatically removed after detecting rapid changes in oxy-Hb value, which had signal variations more than 0.7 mmol mm between successive samples (Rejection rate = 34.6%). The time-continuous data of Hb-signals for each channel were separated into analysis blocks, which consisted of 5 s baseline period followed by 15 s of the target block and

10 s of the baseline block. To eliminate long-term signal trends due to systemic vascular factors, a first-degree baseline fit was estimated for each channel using the first 4 and last 4 s of analysis block. The time course of Hb concentration changes of the analysis blocks were averaged more than five times for each of the stimulus conditions. To objectively set the time window for the analysis, we first calculated peak latency for all the sound conditions by averaging the Hb time course for all channels and participants. From the onset this latency was 11.1 s. Based on this value, a 5-s time window centered about the 11.1-s point, was determined for the target block (Watanabe et al., 2010; Minagawa-Kawai et al., 2011c). Five seconds prior to stimulus onset was used as a time window for the baseline block. The average concentration of oxy- and deoxy-Hb in each time window was calculated for all channels and for each subject. The significance of differences between Hb changes within the baseline and those within target blocks was determined using a *t*-test for each channel under two experimental conditions. Error rates were adjusted to accommodate multiple comparisons using a false discovery rate (FDR) for determination of statistical significance. Instead of the conventional family wise error correction procedure, a method of correction for multiple comparisons that has been shown suitable for NIRS studies (Benjamini and Hochberg, 1995; Singh and Dan, 2006) was applied to control for Type I and II errors. We set the value of *q* specifying the maximum FDR to 0.05, so that there were no more than 5% false positives on average in the number of significant channels.

Next, to assess laterality effects, we followed the same criteria as in previous studies. This entailed first defining a ROI of a vicinity of auditory area as CH6, 8, 9, and 11 on the left and CH19, 21, 22, and 24 on the right hemisphere. The averaged oxy-Hb values were calculated for each condition and hemisphere and then compared between hemispheres. Finally, we examined the laterality effect by employing an analysis procedure similar to that used in previous NIRS studies (Furuya and Mori, 2003; Sato et al., 2003, 2007; Minagawa-Kawai et al., 2005, 2007, 2009). This allows a direct comparison of results across different studies. For each participant, we selected one channel that showed the maximum oxy-Hb responses within a vicinity of auditory areas. This method has effectively revealed functional laterality of auditory processing between left-handers and right-handers (Furuya and Mori, 2003). The LI was calculated using the formula $(L - R)/(L + R)$, where L and R are peak values on left and right sides respectively.

For spatial estimation of channel location in the brain, we employed the virtual registration method (Tsuzuki et al., 2007) to map NIRS data onto the MNI standard brain space. Although this method is basically applicable to adult brains, we adapted it for evaluation of infants' brain activity by adjusting for differences in head size and the emitter–detector separation length (inter-probe separation) between adults and neonates. First, we calculated the average head size of neonates including the circumference (average, 33.8 cm; SD, 0.73), nasion–inion length (average, 20.8 cm; SD, 1.49), and length of preauricular points (average, 22.2 cm; SD, 1.22). The head size ratio of the adult to neonate was revealed to be similar to that of 30 mm inter-probe separation to 20 mm used for infants, with an error range of 2–3 mm. Because the error range of adults virtual registration for the same channel placement with an inter-probe separation of 30 mm was 4–8 mm, this registration can be

applied to our participants. Considering the differences of detailed anatomy in infants and adults such as relative brain position in terms of 10–20 system, we did not use the detailed anatomical labeling obtained from virtual registration. Instead, we used the approximate anatomical labeling.

RESULTS

Both phonemic and prosodic contrasts activated the neonates' brain in substantially broad areas involving superior temporal gyrus, inferior frontal gyrus, and inferior parietal regions. However, the two experimental conditions elicited respectively different time courses of Hb changes as well as revealing different activation foci. This is shown in **Figures 3, 4** and **Table 1**. **Figure 3** shows that Hb changes in the phonemic condition had 10.2 s of peak latency with an initial dip, whereas changes in the prosodic condition showed a peak of 12.1 s without an initial dip. There was no statistically significant difference between these peak times ($t = 0.69$, $p = 0.24$). Phonemic changes activated the inferior frontal, inferior parietal, and temporal areas with less parietal or superior part of activities on the right. In contrast, the prosodic changes evoked responses chiefly around temporal areas. Among these areas, activation foci whose *p*-value is below 0.01 (corrected) are CH6, CH22 (vicinity of auditory areas on the left and right) and CH5 (inferior parietal area) for the phonemic condition and CH24 [vicinity of auditory areas on the right, superior temporal sulcus (STS)/mid temporal] for the prosodic condition.

To examine laterality differences, averaged oxy-Hb values in the ROI of the auditory area as well as the non-auditory channel (CH5) registering strong activity ($p < 0.01$, corrected) were compared with counterpart regions, i.e., contra-lateral ROI and channel. As CH6, 22, and 24 were included in ROI, we did not test them individually. Results of a paired *t*-test showed a significantly strong activation in left-CH5 ($t = 2.29$, $p < 0.05$, corrected) for the phonemic condition. Although the ROI activations in the phonemic condition showed slightly rightward dominance, they did not have any significant hemispheric difference probably due to larger variance than that of the prosodic condition ($t = 0.84$, $p > 0.05$). In the prosodic condition, significantly stronger activation was found in the right ROI than in the left one ($t = 1.88$, $p < 0.05$, corrected; **Figure 5**). To compare the neonates' results with those from previous studies using similar methods, we applied the same

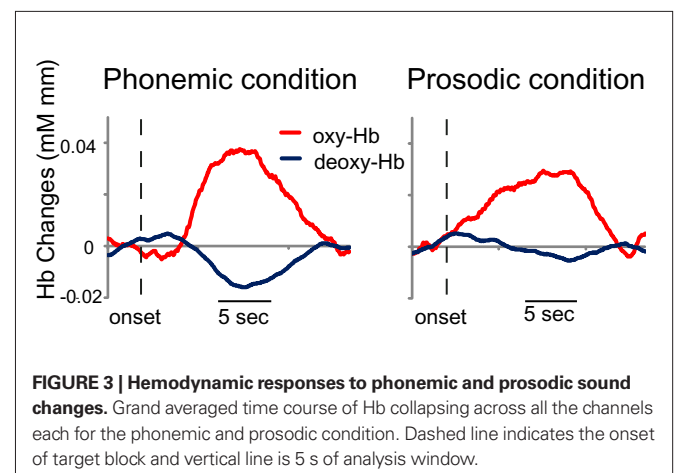
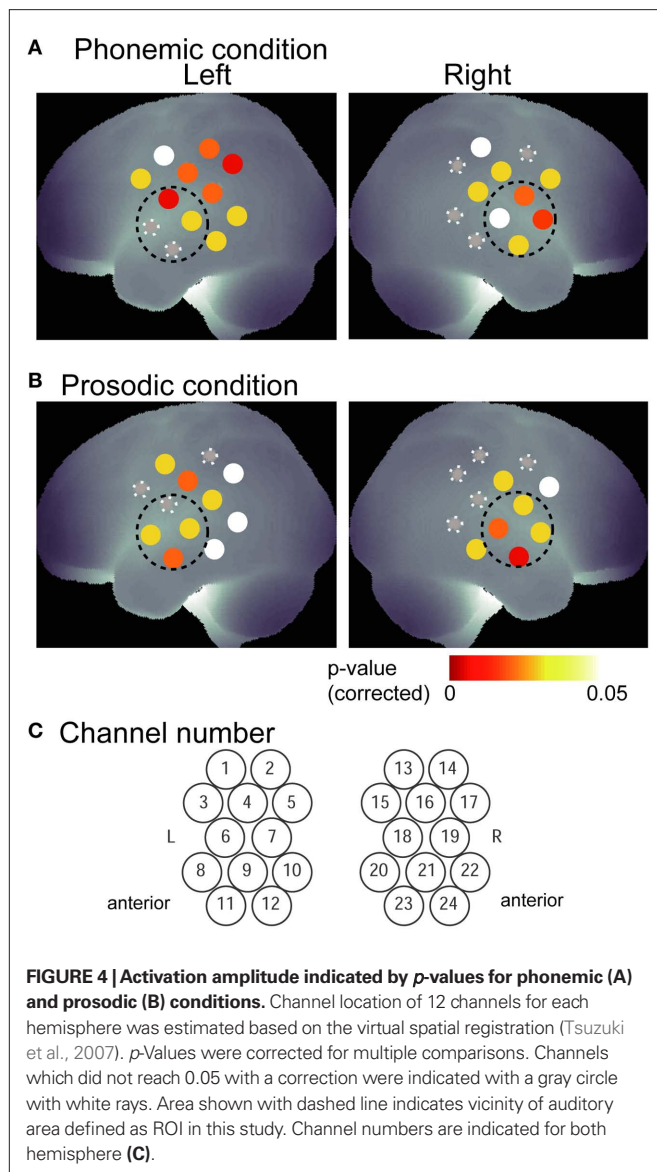


FIGURE 3 | Hemodynamic responses to phonemic and prosodic sound changes. Grand averaged time course of Hb collapsing across all the channels each for the phonemic and prosodic condition. Dashed line indicates the onset of target block and vertical line is 5 s of analysis window.



analytic techniques used in those studies to assess the laterality of auditory areas. Laterality indices (calculated for each participant) are plotted in **Figure 6** for each of the two experimental conditions. Consistent with the results obtained by the ROI analysis, only the prosodic condition showed a significant asymmetry effect. The LI scores for the prosodic condition were significantly lower than zero ($t = 3.07$, $p < 0.01$), indicating rightward dominance. t -Test also showed a significant difference between LI scores for phonemic and prosodic conditions ($t = 2.24$, $p = 0.016$).

DISCUSSION

To explore the early neural bases underlying segmental and suprasegmental processing, the present study measured hemodynamic responses to phonemic and prosodic contrasts in neonates. Results showed a large and significant activation in response to the prosodic change that was located in right temporal region. This suggests a functional specialization for suprasegmental properties

in neonates. By contrast, the phonemic (vowel) contrast showed symmetrical Hb changes in auditory areas; however, it is noteworthy that this contrast also elicited a strong leftward response in the inferior parietal region. Here we discuss these results in light of developmental hemispheric specialization of the temporal area for phonemic and prosodic processing by comparing the results from the previous infant studies.

As indicated in the introduction, previous NIRS studies that have used identical stimulus contrasts reported finding an absence of functional specialization of the auditory area for two different phonetic contrasts in 7- to 8- and 9- to 10-month-olds (e.g., Sato et al., 2003). But the latter research also presented evidence of a tendency for right hemispheric dominance with prosodic contrasts. The present study used neonates as participants and it produced a clearer outcome. Neonates' NIRS responses revealed significant right-dominance around auditory area in response to the prosodic change, suggesting that a specialized function of the right hemisphere for prosody processing is present at birth in human infants. The focus of this activation ranged over four channels in the right auditory region and appeared to involve the STS and mid temporal gyrus.

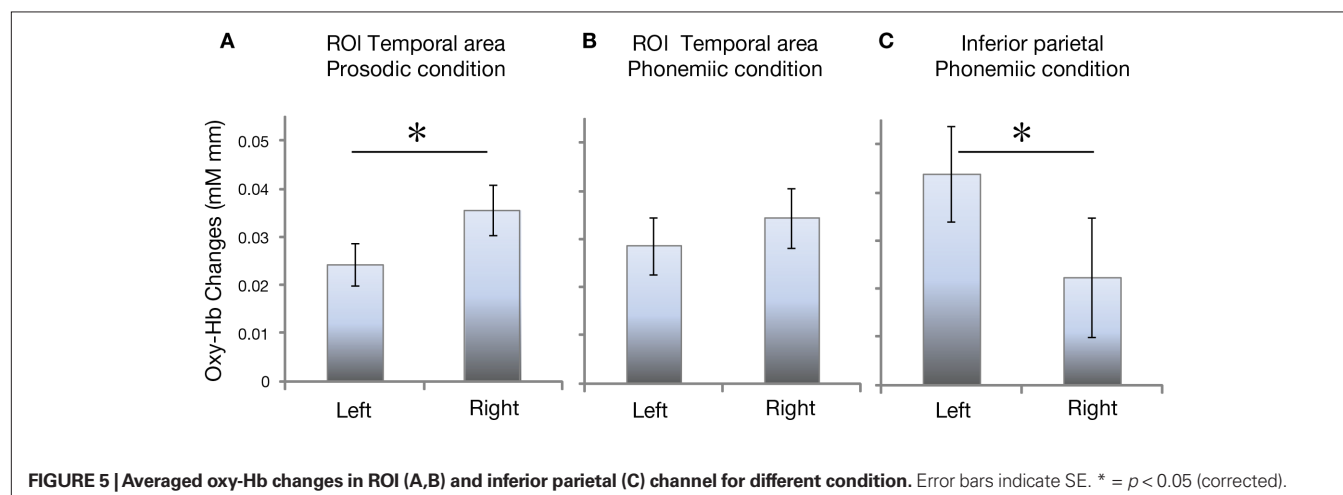
What kind of cognitive function is reflected in the brain activities in this area of the right hemisphere? This will depend upon a listeners age. It is difficult to associate activation in neonate response to a prosodic manipulation if this processing is interpreted to mean a high level of acquired language skills (e.g., distinguishing implied affirmations versus interrogation). Clearly, newborns will lack such skills. Rather, it is more reasonable to assume that this activity reflects a lower cognitive processing, one that involves differentiation of acoustic contours of those spectral components that change with a prosodic manipulation. There is further evidence to support this interpretation. For instance, neuroimaging data of adults showed a cerebral laterality that reflected differential responding to both fast versus slow band-noise stimuli (Boemio et al., 2005) and to temporal versus spectral modulated stimuli (Zatorre and Belin, 2001). Spectrally rich stimuli elicit activations in the anterior superior temporal gyrus as well as the right STS, and these activations increase with the richness of spectral variations (Zatorre and Belin, 2001). Although our prosodic stimulus is not long, it has pitch modulations with richer spectral changes than the other contrast of /itta/ and /itte/ which only has two spectrum differences. In the present study, it is assumed that contrasts between stimuli ending in a rising contour versus those with an unchanging pitch contour are chiefly processed around the right STS in neonates.

Other evidence speaks more directly to developmental issues. Telkemeyer et al. (2009) presented neonates with a subset of stimuli from Boemio et al. (2005) and showed a significant response near the right temporo-parietal area to "slow" stimuli, although its effect is not so powerful. Homae et al. (2006) presented 3-month-old infants with sentential speech prosody and reported dominant activations of the right temporo-parietal region. Although these activated regions are not in brain areas identical to those active regions found in our investigation, there is nonetheless, a rightward superiority in processing prosodic information in neonates or young infants that is in agreement with our findings. Furthermore, these data together with those gathered in the present study suggest the operation of a neuronal network involving the temporo-parietal region and STS/MTG which is partially active from birth.

Table 1 | Statistic results of significant channels.

Phonemic condition				Prosodic condition			
CH	t-Value	Uncorrected <i>p</i>	Corrected <i>p</i> <	CH	t-Value	Uncorrected <i>p</i>	Corrected <i>p</i> <
1	2.267	0.021	0.05	1	2.786	0.007	0.03
2	3.501	0.002	0.02	4	3.534	0.002	0.02
3	2.582	0.011	0.03	7	2.683	0.009	0.03
4	3.216	0.003	0.02	8	2.975	0.005	0.03
5	4.465	0.0003	0.005	9	2.882	0.006	0.03
6	4.951	0.0001	0.005	10	2.201	0.022	0.05
7	3.248	0.003	0.02	11	3.536	0.002	0.02
9	2.691	0.009	0.03	12	2.227	0.021	0.05
10	2.571	0.012	0.03	16	2.431	0.015	0.03
12	2.540	0.012	0.03	19	2.501	0.013	0.03
13	2.018	0.032	0.05	21	3.242	0.003	0.02
16	2.574	0.012	0.03	22	2.636	0.010	0.03
17	2.503	0.013	0.03	23	2.708	0.009	0.03
18	2.752	0.008	0.03	24	4.528	0.0002	0.005
19	3.243	0.003	0.02				
21	2.225	0.023	0.05				
22	3.985	0.001	0.01				
24	2.481	0.014	0.03				

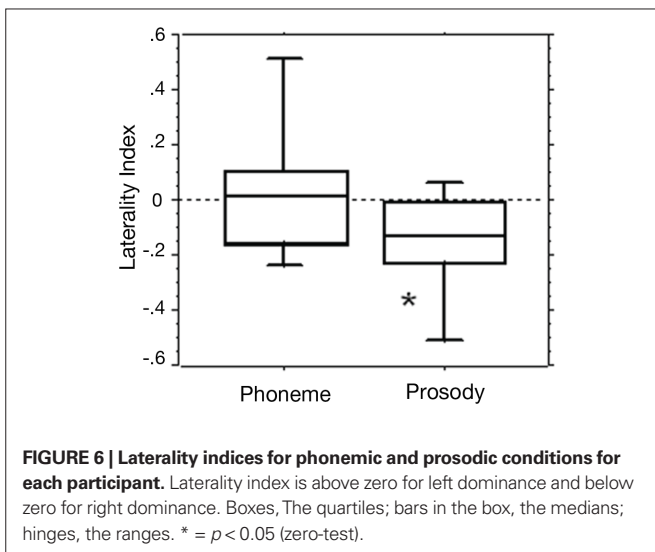
CH, channel number.

**FIGURE 5 |** Averaged oxy-Hb changes in ROI (A,B) and inferior parietal (C) channel for different condition. Error bars indicate SE. * = $p < 0.05$ (corrected).

Other evidence appears to conflict with these findings. Minagawa-Kawai et al. (2011b) presented stimuli used by Zatorre and Belin (2001) to neonates with aim of examining signal-driven mechanisms in early infancy. In this study, they used the contrast temporal versus spectral variations where speed of tone alternation or spectral richness was manipulated. They did not find clear hemispheric specialization associated with signal properties (temporal versus spectral), although intensity of signal change (relative entropy) did correlated with the amplitude of Hb changes. These conflicting results on the lateralization in young infants' brain may be associated with the difference between speech and non-speech stimuli, because it seems that speech elicited clearer lateralization (Peña

et al., 2003; Gervain et al., 2008) than non-speech did (Minagawa-Kawai et al., 2011b; Telkemeyer et al., 2009) in neonates suggesting a specific role of human vocalization. Further research is required to explain these discrepancies. This discrepancy underscores the need for greater attention to clarifying the acoustic definition of signal-driven system with regard to critical details of acoustic features that may be determining these conflicting outcomes. Thus the influence of the signal-driven system on the cerebral responses during processing speech is still tentative conclusion.

With respect to the phonemic vowel contrast of /itta/ versus /itte/, a cross-sectional study (Sato et al., 2003) showed that their youngest groups (7–8, 9–10 months of age) evoked activations



equally in bilateral temporal areas. It was only when infants approached 11 months of age that they showed a lateralization of the vowel difference in the form of leftward dominance. Our results provide additional evidence that the auditory region functions as an innate starting for the development of auditory processing. Taken together, both sets of findings permit the inference that bilateral engagement for processing vowel contrast continues from birth to ages of 7 months. Although we lack data for 2- to 6-month-olds, previous results for different vowel types showing a bilateral temporal response in 3- to 4- and 6- to 7-month-olds (Minagawa-Kawai et al., 2007) supports the idea of a continuous developmental trajectory through these intervening age levels.

Several interpretations may account for emergence of symmetrical auditory responses to the vowel changes. First, as in similar the prosodic condition, a signal-driven mechanism can explain these bilateral activations. As indicated earlier, a hypothetical signal-driven mechanism may determine bilateral responses in the temporal cortex rather than a leftward one in reaction to vowel changes. In general, vowels have been reported to be less lateralized than consonants (e.g., Haggard and Parkinson, 1971), and this may be attributable to the fact that vowels contain spectral components that change more slowly than consonants which exhibit quite rapid dynamic changes. Some neuroimaging studies also support this idea by showing greater involvement of the left planum temporale in processing CV than when a tone or a vowel is presented in isolation (Jancke et al., 2002). Admittedly, laterality is not entirely based upon signal factors, but at least the present study indicates their primary impact on neonates. In this sense, the present findings with both phonetic and prosodic contrasts can be explained by a signal-driven mechanism. In fact, a model incorporating this idea has been proposed by Minagawa-Kawai et al. (2011a). It describes developmental hemispheric specialization associated with language acquisition. Basically, this model assumes that lateralization for language emerges out of the interaction between pre-existing left-right biases in generic auditory processing (signal-driven mechanism), and a left-hemisphere predominance of particular learning mechanisms.

A second approach to this issue focuses upon the immaturity of the nervous system in neonates. Functional cerebral lateralization is typically assumed to reflect a mature neural network ranging over both hemispheres, but in resting states it has been shown that neonates have less connectivity across hemispheres than do 3 month-olds (Homae et al., 2010). However, this interpretation does not specifically take into account the right hemispheric dominance for prosodic contrasts. In any case, what appears clear is that bilateral activities for vowel processing eventually become functionally lateralized to the left auditory area as infants learn vowel categories of their native language. Thus, as an infants' brain matures physiologically it does so in conjunction with a reorganization of synaptic connections.

To this point our discussion of lateralization has been confined to the vicinity of auditory areas. However, another rather unexpected finding was discovered in this study: dominant activations were observed in the left parietal region during vowel discrimination. These activations seem to be in the supra marginal gyrus (SMG) according to the probabilistic spatial estimation (Tsuzuki et al., 2007). The neuroimaging literature often refers to SMG in relation to speech perception. An MRI study of lesions in aphasic adult patients by Caplan et al. (1995) indicates that the left SMG is the principal site of phonemic processing; thus, patients with lesions in this area typically fail to discriminate and identify phonemes. Further, Zatorre et al. (1992) also showed that discrimination of consonant types in CVC syllable, activated the left SMG. It seems left SMG is also involved in tasks requiring verbal or auditory short-term memory (Paulesu et al., 1993). Although the neonates did not engage in any particular task in the present study, auditory short-term memory is a likely candidate for explaining SMG activations. That is, cognitive process of discrimination during the target block may underlie in the activities of left SMG even in sleeping neonates. Specifically, in contrast to the baseline trial block in which the infants received the same word repetitively, in phonemic target trials, infants had to discriminate between two temporally separated words (/iita/ and /itte/) that differ in vowels. It seems likely that this would place demands on short-term memory. Thus, activity observed in the Left SMG may a type of memory processes that is required for phoneme detection/discrimination but not for prosodic discrimination. If this interpretation is correct, these data provide indirect evidence that a neuronal substrate implicated in short-term memory may also be functional in newborns. But a caveat is warranted regarding whether or not the left SMG activation is language-specific/phoneme-specific. Future studies using non-speech analogs of the present stimuli should clarify this issue.

The involvement of various cerebral mechanism and their role in laterality during phonetic processing in infants has been examined by EEG studies as well as dichotic listening studies. But evidence has been limited with regard to phonetic processing in neonates. What evidence exists shows that newborns exhibit discriminative reactions to vowel differences (Cheour-Luhtanen et al., 1995; Dehaene-Lambertz and Pena, 2001) and that their auditory areas are sensitive to categorical voicing difference (Simos and Moltes, 1997). Very young infants also tend to show delayed latency of mismatch negativity to phonemic difference as compared with that of adults suggesting infants' premature processing system (Dehaene-Lambertz and Gliga, 2004). However, laterality differences in infants, based upon early EEG studies, have provided rather diverse results showing left dominance (Dehaene-Lambertz and Baillet, 1998),

right dominance (Molfese and Molfese, 1988; Novak et al., 1989), and bilateral activation (Simos and Molfese, 1997). The diversity of such outcomes is probably due to the limitation of spatial resolution of EEG. But, dichotic listening, test using the sucking procedure for infants, has also revealed a complicated picture regarding the laterality (Bertoncini et al., 1989). Furthermore, many of these EEG studies did not precisely reveal the activation focus or brain region involved. However, a few studies employed high-density ERP and/or sophisticated dipole modeling and these should provide better spatial resolution. For instance, Dehaene-Lambertz et al. (2004) tested 3-weeks-old infants with sylvian infarct on the left hemisphere and they found a discriminative response to vowel differences that implied a right-hemisphere contribution to vowel perception at this age. A recent study on 2-months-old infants detected ERP source locations on the left hemisphere for vowel processing (Bristow et al., 2009). These locations involved the inferior frontal gyrus and superior temporal gyrus and temporal sulcus. Activation in the superior gyrus is consistent with our results but not with those of other studies. We had strong activations in the parietal regions but not for the inferior frontal gyrus. Although the diversity of these findings may be due to variations in testing instruments, stimulus presentation and infants' age, the co-registration of ERP and NIRS may further provide detailed evidence with respect to time course and brain region of phonemic processing in young infants' brain.

The NIRS methodology offer more reliable cortical localizations than EEG techniques, but studies using the former methodology have not investigated vowel processing in neonates. Nevertheless, some of these studies address aspects of neonates' speech perception that are relevant to the discussion of the NIRS data presented here. Neonates showed left-dominant Hb responses from the temporal area during listening to forward (normal) speech in contrast to bilateral response to backward speech (Peña et al., 2003). Such asymmetrical activations are also observed in response to repetition sequences of syllables against random controls (Gervain et al., 2008) suggesting neonates' ability to find out a certain type of language structure. These results imply that it is not only signal properties that modulate the laterality of neonates, because acoustic features of target and control stimuli are similar in these studies. Thus, as with adults, laterality in neonates may also be driven by cognitive activity elicited by the specific type of stimuli and/or the presentation method. Finally, the present study found activation focus in the temporal area and SMG, but prefrontal measurement with another probe pad would reveal other activation

focus reflecting novelty detection. Similar activity has been observed in the prefrontal region in 2–3 month-olds (Nakano et al., 2009). As discussed here, infant NIRS studies have enabled us to discuss localized brain function in relation to language development. Another possibly relevant parameter of NIRS that is implicated in this study is the latency or response shape of Hb time course. Although there was no statistically significant difference in latency between the conditions, prosodic condition with different response shape elicited rather slower Hb response than that for the phonemic condition. This could derive from difference of processing speed, and intonation contour may require higher spectral resolution.

CONCLUSION

In summary, by presenting segmental versus suprasegmental (phoneme versus prosody) contrasts to newborn infants, the present study revealed a functional lateralization to right temporal area for prosody processing and bilateral engagement of the auditory areas for vowel contrast. Overall, these results were explained by the signal properties of the acoustic stimuli which differentially activated distinct regions in the temporal cortex. This is the first evidence showing that neonates exhibit localized cerebral responses to phonemic contrasts of vowel and prosody. We further showed a left-dominant activation in neonates around inferior parietal region suggesting an early neuronal basis for auditory-verbal short-term memory. This study suggests that a brain mechanism for a certain form of signal-driven system in the speech stimulus context is present at birth and that it possibly operates in coordination with a domain driven system. This raises several important issues that merit further exploration in the development of infants' neurocognitive system, including differential impact of speech and non-speech on the lateralization of neonates' brain.

ACKNOWLEDGMENTS

The authors thank K. Kosaki, Y. Matsuzaki, M. Miwa, E. Okishio, and all the staffs of neonatal unit of Keio University Hospital for help with the study, T. Imaizumi for kindly providing us the sound stimuli, K. Maekawa for his generous comments and S. Ishii and A. Matsuzaki for help with conducting the experiment. This work was supported by Grant-in-Aid for Scientific Research (A) (KAKENHI, project No. 21682002), Academic Frontier Project supported by Ministry of Education, Culture, Sports, Science and Technology (MEXT) and the Global COE (Center of Excellence) program Keio University.

REFERENCES

- Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Series B* 57, 289–300.
- Bertoncini, J., Morais, J., Bijeljac-Babic, R., McAdams, S., Peretz, I., and Mehler, J. (1989). Dichotic perception and laterality in neonates. *Brain Lang.* 37, 591–605.
- Boemio, A., Fromm, S., Braun, A., and Poeppel, D. (2005). Hierarchical and asymmetric temporal sensitivity in human auditory cortices. *Nat. Neurosci.* 8, 389–395.
- Bristow, D., Dehaene-Lambertz, G., Mattout, J., Soares, C., Gliga, T., Baillet, S., and Mangin, J. F. (2009). Hearing faces: how the infant brain matches the face it sees with the speech it hears. *J. Cogn. Neurosci.* 21, 905–921.
- Caplan, D., Gow, D., and Makris, N. (1995). Analysis of lesions by MRI in stroke patients with acoustic-phonetic processing deficits. *Neurology* 45, 293–298.
- Cheour, M., Ceponiene, R., Lehtokoski, A., Luuk, A., Allik, J., Alho, K., and Naatanen, R. (1998). Development of language-specific phoneme representations in the infant brain. *Nat. Neurosci.* 1, 351–353.
- Cheour-Luhtanen, M., Alho, K., Kujala, T., Sainio, K., Reinikainen, K., Renlund, M., Aaltonen, O., Eerola, O., and Naatanen, R. (1995). Mismatch negativity indicates vowel discrimination in newborns. *Hear. Res.* 82, 53–58.
- Dehaene-Lambertz, G., and Baillet, S. (1998). A phonological representation in the infant brain. *Neuroreport* 9, 1885–1888.
- Dehaene-Lambertz, G., and Gliga, T. (2004). Common neural basis for phoneme processing in infants and adults. *J. Cogn. Neurosci.* 16, 1375–1387.
- Dehaene-Lambertz, G., Pallier, C., Serniclaes, W., Sprenger-Charolles, L., Jobert, A., and Dehaene, S. (2005). Neural correlates of switching from auditory to speech perception. *Neuroimage* 24, 21–33.
- Dehaene-Lambertz, G., and Pena, M. (2001). Electrophysiological evidence for automatic phonetic processing in neonates. *Neuroreport* 12, 3155–3158.
- Dehaene-Lambertz, G., Pena, M., Christophe, A., and Landrieu, P. (2004). Phoneme perception in a neonate with a left sylvian infarct. *Brain Lang.* 88, 26–38.
- Friederici, A. D., and Alter, K. (2004). Lateralization of auditory language functions: a dynamic dual pathway model. *Brain Lang.* 89, 267–276.
- Furuya, I., and Mori, K. (2003). Cerebral lateralization in spoken language processing measured by multi-channel near-infrared spectroscopy (NIRS). *Brain Nerve* 55, 226–231.

- Gandour, J., Wong, D., Hsieh, L., Weinzapfel, B., Van Lancker, D., and Hutchins, G. D. (2000). A crosslinguistic PET study of tone perception. *J. Cogn. Neurosci.* 12, 207–222.
- Gervain, J., Macagno, E., Cogoi, S., Pena, M., and Mehler, J. (2008). The neonate brain detects speech structure. *Proc. Natl. Acad. Sci. U.S.A.* 105, 14222–14227.
- Goldsmith, J. A. (1990). *Autosegmental and Metrical Phonology*. Cambridge: Basil Blackwell.
- Haggard, M. P., and Parkinson, A. M. (1971). Stimulus and task factors as determinants of ear advantages. *Q. J. Exp. Psychol.* 23, 168–177.
- Homae, F., Watanabe, H., Nakano, T., Asakawa, K., and Taga, G. (2006). The right hemisphere of sleeping infant perceives sentential prosody. *Neurosci. Res.* 54, 276–280.
- Homae, F., Watanabe, H., Otake, T., Nakano, T., Go, T., Konishi, Y., and Taga, G. (2010). Development of global cortical networks in early infancy. *J. Neurosci.* 30, 4877–4882.
- Imazumi, S., Mori, K., Kiritani, S., Hosoi, H., and Tonoike, M. (1998). Task-dependent laterality for cue decoding during spoken language processing. *Neuroreport* 9, 899–903.
- Jacquemot, C., Pallier, C., LeBihan, D., Dehaene, S., and Dupoux, E. (2003). Phonological grammar shapes the auditory cortex: a functional magnetic resonance imaging study. *J. Neurosci.* 23, 9541–9546.
- Jamison, H. L., Watkins, K. E., Bishop, D. V., and Matthews, P. M. (2006). Hemispheric specialization for processing auditory nonspeech stimuli. *Cereb. Cortex* 16, 1266–1275.
- Jancke, L., Wustenberg, T., Scheich, H., and Heinze, H. J. (2002). Phonetic perception and the temporal cortex. *Neuroimage* 15, 733–746.
- Jusczyk, P. W., Houston, D. M., and Newsome, M. (1999). The beginnings of word segmentation in english-learning infants. *Cogn. Psychol.* 39, 159–207.
- Kuhl, P. K. (2004). Early language acquisition: cracking the speech code. *Nat. Rev. Neurosci.* 5, 831–843.
- Mattock, K., and Burnham, D. (2006). Chinese and English infants' tone perception: evidence for perceptual reorganization. *Infancy* 10, 241–265.
- Mattock, K., Molnar, M., Polka, L., and Burnham, D. (2008). The developmental course of lexical tone perception in the first year of life. *Cognition* 106, 1367–1381.
- Mehler, J., Jusczyk, P., Lambertz, G., Halsted, N., Bertoncini, J., and Amiel-Tison, C. (1988). A precursor of language acquisition in young infants. *Cognition* 29, 143–178.
- Minagawa-Kawai, Y., Cristia, A., and Dupoux, E. (2011a). Cerebral lateralization and early speech acquisition: a developmental scenario. *J. Dev. Cogn. Neurosci.* 1, 217–232.
- Minagawa-Kawai, Y., Cristia, A., Vendelin, I., Cabrol, D., and Dupoux, E. (2011b). Assessing signal-driven mechanisms in neonates: brain responses to temporally and spectrally different sounds. *Front. Psychology* 2:135. doi: 10.3389/fpsyg.2011.00135
- Minagawa-Kawai, Y., Mori, K., Naoi, N., and Kojima, S. (2007). Neural attunement processes in infants during the acquisition of a language-specific phonemic contrast. *J. Neurosci.* 27, 315–321.
- Minagawa-Kawai, Y., Mori, K., and Sato, Y. (2005). Different brain strategies underlie the categorical perception of foreign and native phonemes. *J. Cogn. Neurosci.* 17, 1376–1385.
- Minagawa-Kawai, Y., Naoi, N., Kikuchi, N., Yamamoto, J., Nakamura, K., and Kojima, S. (2009). Cerebral laterality for phonemic and prosodic cue decoding in children with autism. *Neuroreport* 20, 1219–1224.
- Minagawa-Kawai, Y., van der Lely, H., Ramus, F., Sato, Y., Mazuka, R., and Dupoux, E. (2011c). Optical brain imaging reveals general auditory and language-specific processing in early infant development. *Cereb. Cortex* 21, 254–261.
- Molfese, D. L., and Molfese, V. J. (1988). Right-hemisphere responses from preschool children to temporal cues to speech and nonspeech materials: electrophysiological correlates. *Brain Lang.* 33, 245–259.
- Mottonen, R., Calvert, G. A., Jaaskelainen, I. P., Matthews, P. M., Thesen, T., Tuomainen, J., and Sams, M. (2006). Perceiving identical sounds as speech or non-speech modulates activity in the left posterior superior temporal sulcus. *Neuroimage* 30, 563–569.
- Mugitani, R. (2009). Developmental process of speech perception in native language during infancy. *Baby Sci.* 8, 38–49.
- Näätänen, R., Lehtokoski, A., Lennes, M., Cheour, M., Huotilainen, M., Iivonen, A., Vainio, M., Alku, P., Ilmoniemi, R. J., Luuk, A., Allik, J., Sinkkonen, J., and Alho, K. (1997). Language-specific phoneme representations revealed by electric and magnetic brain responses. *Nature* 385, 432–434.
- Nakano, T., Watanabe, H., Homae, F., and Taga, G. (2009). Prefrontal cortical involvement in young infants' analysis of novelty. *Cereb. Cortex* 19, 455–463.
- Nazzi, T., Jusczyk, P. W., and Johnson, E. K. (2000). Language discrimination by English-learning 5-month-olds: effects of rhythm and familiarity. *J. Mem. Lang.* 43, 1–19.
- Novak, G. P., Kurtzberg, D., Kreuzer, J. A., and Vaughan, H. G. Jr. (1989). Cortical responses to speech sounds and their formants in normal infants: maturational sequence and spatiotemporal analysis. *Electroencephalogr. Clin. Neurophysiol.* 73, 295–305.
- Paulesu, E., Frith, C. D., and Frackowiak, R. S. (1993). The neural correlates of the verbal component of working memory. *Nature* 362, 342–345.
- Peña, M., Maki, A., Kovacic, D., Dehaene-Lambertz, G., Koizumi, H., Bouquet, F., and Mehler, J. (2003). Sounds and silence: an optical topography study of language recognition at birth. *Proc. Natl. Acad. Sci. U.S.A.* 100, 11702–11705.
- Poeppel, D. (2003). The analysis of speech in different temporal integration windows: cerebral lateralization as “asymmetric sampling in time”. *Speech Commun.* 41, 245–255.
- Poeppel, D., Idsardi, W. J., and van Wassenhove, V. (2008). Speech perception at the interface of neurobiology and linguistics. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 363, 1071–1086.
- Sato, Y., Mori, K., Furuya, I., Hayashi, R., Minagawa-Kawai, Y., and Koizumi, T. (2003). Developmental changes in cerebral lateralization during speech processing measured by near infrared spectroscopy. *Jpn. J. Logoped. Phoniatr.* 44, 165–171.
- Sato, Y., Sogabe, Y., and Mazuka, R. (2007). Brain responses in the processing of lexical pitch-accent by Japanese speakers. *Neuroreport* 18, 2001–2004.
- Sato, Y., Sogabe, Y., and Mazuka, R. (2010). Development of hemispheric specialization for lexical pitch-accent in Japanese infants. *J. Cogn. Neurosci.* 22, 2503–2513.
- Schönwiesner, M., Rubsamen, R., and von Cramon, D. Y. (2005). Hemispheric asymmetry for spectral and temporal processing in the human antero-lateral auditory belt cortex. *Eur. J. Neurosci.* 22, 1521–1528.
- Shankweiler, D., and Studdert-Kennedy, M. (1967). Identification of consonants and vowels presented to left and right ears. *Quart. J. Exp. Psychol.* 19, 59–63.
- Shtyrov, Y., Kujala, T., Palva, S., Ilmoniemi, R. J., and Näätänen, R. (2000). Discrimination of speech and of complex nonspeech sounds of different temporal structure in the left and right cerebral hemispheres. *Neuroimage* 12, 657–663.
- Simos, P. G., and Molfese, D. L. (1997). Electrophysiological responses from a temporal order continuum in the newborn infant. *Neuropsychologia* 35, 89–98.
- Singh, A. K., and Dan, I. (2006). Exploring the false discovery rate in multichannel NIRS. *Neuroimage* 33, 542–549.
- Telkemeyer, S., Rossi, S., Koch, S. P., Nierhaus, T., Steinbrink, J., Poeppel, D., Obrig, H., and Wartenburger, I. (2009). Sensitivity of newborn auditory cortex to the temporal structure of sounds. *J. Neurosci.* 29, 14726–14733.
- Tsuzuki, D., Jurcak, V., Singh, A. K., Okamoto, M., Watanabe, E., and Dan, I. (2007). Virtual spatial registration of stand-alone fNIRS data to MNI space. *Neuroimage* 34, 1506–1518.
- Watanabe, H., Homae, F., and Taga, G. (2010). General to specific development of functional activation in the cerebral cortexes of 2- to 3-month-old infants. *Neuroimage* 50, 1536–1544.
- Weiss, M., and House, A. (1973). Perception of dichotically presented vowels. *J. Acoust. Soc. Am.* 53, 51–58.
- Yamashita, Y., Maki, A., and Koizumi, H. (1996). Near-infrared topographic measurement system: imaging of absorbers localized in a scattering medium. *Rev. Sci. Instrum.* 67, 730–732.
- Zatorre, R. J., and Belin, P. (2001). Spectral and temporal processing in human auditory cortex. *Cereb. Cortex* 11, 946–953.
- Zatorre, R. J., Evans, A. C., Meyer, E., and Gjedde, A. (1992). Lateralization of phonetic and pitch discrimination in speech processing. *Science* 256, 846–849.
- Zatorre, R. J., and Gandour, J. T. (2008). Neural specializations for speech and pitch: moving beyond the dichotomies. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 363, 1087–1104.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 22 December 2010; accepted: 09 August 2011; published online: 15 September 2011.

Citation: Arimitsu T, Uchida-Ota M, Yagihashi T, Kojima S, Watanabe S, Hokuto I, Ikeda K, Takahashi T and Minagawa-Kawai Y (2011) Functional hemispheric specialization in processing phonemic and prosodic auditory changes in neonates. *Front. Psychology* 2:202. doi: 10.3389/fpsyg.2011.00202

This article was submitted to *Frontiers in Language Sciences*, a specialty of *Frontiers in Psychology*.

Copyright © 2011 Arimitsu, Uchida-Ota, Yagihashi, Kojima, Watanabe, Hokuto, Ikeda, Takahashi and Minagawa-Kawai. This is an open-access article subject to a non-exclusive license between the authors and Frontiers Media SA, which permits use, distribution and reproduction in other forums, provided the original authors and source are credited and other Frontiers conditions are complied with.



Functional lateralization of speech processing in adults and children who stutter

Yutaka Sato^{1,2*}, Koichi Mori¹, Toshizo Koizumi^{1,3}, Yasuyo Minagawa-Kawai^{1,4}, Akihiro Tanaka^{1,5}, Emi Ozawa⁶, Yoko Wakaba⁷ and Reiko Mazuka^{2,8}

¹ Research Institute, National Rehabilitation Center for Persons with Disabilities, Tokorozawa, Saitama, Japan

² Laboratory for Language Development, Brain Science Institute, RIKEN, Wako, Saitama, Japan

³ Department of Otorhinolaryngology, Nara Medical University, Kashihara, Nara, Japan

⁴ Graduate School of Human Relations, Keio University, Tokyo, Japan

⁵ Waseda Institute for Advanced Study, Waseda University, Tokyo, Japan

⁶ Hospital of National Rehabilitation Center for Persons with Disabilities, Tokorozawa, Saitama, Japan

⁷ Master's Course, Child Study Program, Graduate School of Social Welfare, Tokyo, Japan

⁸ Department of Psychology and Neuroscience, Duke University, Durham, NC, USA

Edited by:

Judit Gervain, Université Paris
Descartes, France

Reviewed by:

Mohinsh Shukla, University of
Rochester, USA

Deryk Scott Beal, Boston University,
USA

*Correspondence:

Yutaka Sato, Laboratory for Language
Development, Brain Science Institute,
RIKEN, 2-1 Hirosawa, Wako, Saitama
351-0198, Japan.
e-mail: satoyu@brain.riken.jp

Developmental stuttering is a speech disorder in fluency characterized by repetitions, prolongations, and silent blocks, especially in the initial parts of utterances. Although their symptoms are motor related, people who stutter show abnormal patterns of cerebral hemispheric dominance in both anterior and posterior language areas. It is unknown whether the abnormal functional lateralization in the posterior language area starts during childhood or emerges as a consequence of many years of stuttering. In order to address this issue, we measured the lateralization of hemodynamic responses in the auditory cortex during auditory speech processing in adults and children who stutter, including preschoolers, with near-infrared spectroscopy. We used the analysis–resynthesis technique to prepare two types of stimuli: (i) a phonemic contrast embedded in Japanese spoken words (/itta/ vs. /itte/) and (ii) a prosodic contrast (/itta/ vs. /itta?/). In the baseline blocks, only /itta/ tokens were presented. In phonemic contrast blocks, /itta/ and /itte/ tokens were presented pseudo-randomly, and /itta/ and /itta?/ tokens in prosodic contrast blocks. In adults and children who do not stutter, there was a clear left-hemispheric advantage for the phonemic contrast compared to the prosodic contrast. Adults and children who stutter, however, showed no significant difference between the two stimulus conditions. A subject-by-subject analysis revealed that not a single subject who stutters showed a left advantage in the phonemic contrast over the prosodic contrast condition. These results indicate that the functional lateralization for auditory speech processing is in disarray among those who stutter, even at preschool age. These results shed light on the neural pathophysiology of developmental stuttering.

Keywords: developmental stuttering, language development, cerebral lateralization

INTRODUCTION

Developmental stuttering is a disorder of speech fluency characterized by involuntary repetitions, prolongations, and silent blocks, especially in the initial parts of utterances. It typically starts between 2 and 6 years of age, occurring in 4–5% of all preschool children (Bloodstein, 1995; Yairi and Ambrose, 1999). Although 70–80% of these children recover spontaneously, the stuttering persists after puberty in approximately 1% of the general population, more often in males than in females (Bloodstein, 1995; Yairi and Ambrose, 1999). Despite the previous physiological research including brain imaging techniques, the pathophysiology and the neural basis underlying developmental stuttering remains poorly understood (Brown et al., 2005; Watkins et al., 2008).

Because stuttering manifests as a motor dysfunction in speech, it has been argued that the symptoms represent breakdowns in the control, timing, and coordination of speech musculature (Brady, 1969; Bruce and Adams, 1978; Jayaram, 1984; Smith, 1995; Ludlow and Loucks, 2003). Moreover, it has been reported that persons

who stutter show disrupted motor activity in the articulatory, laryngeal, and respiratory systems during speech (Zimmerman, 1980; Conture et al., 1986; Peters and Boves, 1988; Zocchi et al., 1990; Denny and Smith, 1992; Smith et al., 1993; McClean and Runyan, 2000). There is evidence, however, that linguistic demands and changes in auditory inputs affect stuttering frequency. The disfluency tends to increase when the planned utterances are long and the speech rate is altered (Jayaram, 1984; Zackheim and Conture, 2003; Blomgren and Goberman, 2008; Sawyer et al., 2008) or the syntax is complex (Gordon et al., 1986; Melnick and Conture, 2000). It has been also reported that the phonological complexity affects the speech motor dynamics in adults who stutter (Smith et al., 2010). Better fluency can be induced with changes in the auditory input, such as delayed or frequency-altered auditory feedback of speech, choral reading (in unison with other speakers), masking by white noise, and external rhythmic cues (e.g., metronome) (Johnson and Rosen, 1937; Cherry and Sayers, 1956; van Riper, 1971; Trotter and Silverman, 1974;

Hargrave et al., 1994; Bloodstein, 1995). These findings suggest that stuttering is not simply an impairment in the motor system. As Hampton and Weber-Fox (2008) argue, many current models of stuttering incorporate other factors: atypical neurophysiology, genetic factors, environment, personality, learning ability, auditory processing, and the ability to produce speech and language (Bloodstein, 1995; Lawrence and Barclay, 1998; Guitar, 2006). This indicates that not only a motor disability but many factors can influence stuttering.

Neuroimaging studies during the last 15 years have reported that adults who stutter demonstrate both structural and functional abnormalities compared to people who do not stutter (Wu et al., 1995; Fox et al., 1996, 2000; Braun et al., 1997; Salmelin et al., 1998, 2000; De Nil et al., 2000, 2008; Ingham et al., 2000, 2004; Foundas et al., 2001, 2003; Ingham, 2001; Sommer et al., 2002; Preibisch et al., 2003; Jäncke et al., 2004; Weber-Fox et al., 2004, 2008; Biermann-Ruben et al., 2005; Brown et al., 2005; Corbera et al., 2005; Cykowski et al., 2008; Giraud et al., 2008; Watkins et al., 2008; Weber-Fox and Hampton, 2008; Chang et al., 2009; Kell et al., 2009; Sakai et al., 2009; Beal et al., 2010, 2011; Kaganovich et al., 2010; Liotti et al., 2010; Lu et al., 2010; Kikuchi et al., 2011). Among various theories of the pathophysiology of stuttering (Bloodstein, 1995; Lawrence and Barclay, 1998; Guitar, 2006), abnormal patterns of cerebral hemispheric dominance for speech processing have been consistently demonstrated. During speech production, people who stutter (PWS) show anomalous patterns characterized by overactivation, particularly in the right hemisphere, in speech-motor related brain areas, and by reduced activation in the left superior temporal, fronto-temporal, and temporo-parietal areas compared with fluency inducing conditions or fluent speakers (Wu et al., 1995; Fox et al., 1996; Braun et al., 1997; De Nil et al., 2000; Ingham et al., 2000). It has been reported that there are differences in brain responses to auditory processing during speech or passive listening tasks, and non-linguistic auditory processing between adults or children who do and do not stutter (Hampton and Weber-Fox, 2008; Beal et al., 2010, 2011; Kaganovich et al., 2010; Kikuchi et al., 2011).

Although these imaging data generally suggest that the structural and functional lateralization in language-related brain regions for speech perception or production differs between adults who do and do not stutter, the crucial question remains unresolved with regard to the possible causal relationship between stuttering and brain lateralization. Since previous studies have mainly examined adults, it has been a matter of debate whether the anatomical and functional increases in the right hemisphere in adults who stutter are the results of compensatory mechanisms used over a lifetime of stuttering (Braun et al., 1997; Preibisch et al., 2003; Chang et al., 2008; Lu et al., 2010). Alternatively, it is possible that the abnormal functional lateralization is observed in children who stutter if it is related to the onset of the stuttering. Data from children who stutter may shed light on this issue, but few such studies have been conducted (Özge et al., 2004; Chang et al., 2008; Weber-Fox et al., 2008; Kaganovich et al., 2010; Beal et al., 2011). Since stuttering typically starts prior to 6 years of age, a method that enables studying functional lateralization in this age group is necessary for elucidating the pathophysiology of developmental stuttering.

Conventional brain imaging techniques, such as functional magnetic resonance imaging (fMRI) and positron emission tomography (PET) are not well-suited for young subjects, due to safety concerns and/or the requirement for rigorous restraint. Magnetoencephalography (MEG) must be applicable for young infants for its safety, but may have difficulties to measure children younger than around 5 years. Although electroencephalography (EEG) can be applicable to young subjects, its ability to localize the focus of activity is generally poor (Minagawa-Kawai et al., 2008). In contrast, near-infrared spectroscopy (NIRS) can non-invasively measure human brain function under a variety of conditions with little restraint on young subjects, even neonates (Chen et al., 2002; Kennan et al., 2002; Peña et al., 2003; Homae et al., 2006, 2007; Minagawa-Kawai et al., 2007, 2011; Sato et al., 2010; Zaramella et al., 2001). NIRS has a reasonable resolution for exploring functional lateralization due to the limited spread of near-infrared light in the brain (Yamashita et al., 1996; Yamamoto et al. 2002; Fukui et al., 2003), unlike evoked potentials. Consequently, NIRS was our logical method of choice for examining the functional lateralization in children and adults who stutter.

We focused on cortical auditory speech processing in this study, based on its suggested involvement in stuttering (Hall and Jerger, 1978; Toscher and Rupp, 1978). Disfluency in PWS can be ameliorated by manipulating the auditory input (Johnson and Rosen, 1937; Cherry and Sayers, 1956; van Riper, 1971; Trotter and Silverman, 1974; Hargrave et al., 1994; Bloodstein, 1995), and choral reading has been shown to reverse the deactivation in the cerebral auditory regions during speech in PWS (Fox et al., 1996). While these are related to the auditory functions for self-monitoring of one's own speech, behavioral studies using a dichotic listening paradigm have demonstrated abnormal linguistic processing in PWS in terms of cerebral hemispheric dominance (Curry and Gregory, 1969; Brady and Berson, 1975; Sommers et al., 1975; Blood, 1985). Although these paradigms would be useful for examining the cerebral differences between people who do and do not stutter, young children may fail to accomplish tasks requiring overt and prompt speech responses. Similarly, tasks requiring intensive attention, like a dichotic listening test, may not be reliably performed. Thus, we used a simple listening task that is applicable to even very young subjects, and can measure functional lateralization for language processing in the absence of overt speech planning or production (Sato et al., 2003; Minagawa-Kawai et al., 2007). We used words with phonemic and prosodic contrasts as auditory stimuli. Using the same stimuli, functional brain mapping with NIRS has demonstrated a left-side advantage for the phonemic contrast compared to the prosodic contrast in normal adults, school-age, and preschool children, as well as infants older than 1 year (Furuya et al., 2001; Furuya and Mori, 2003; Sato et al., 2003).

MATERIALS AND METHODS

PARTICIPANTS

Ten adults (10 males, age range 18–44 years), seven school-age children (two females and five males, age range 6–12 years), and six preschool-age children (one female and five males, age range 3–5 years) who stutter participated in the present study. All subjects were native Japanese speakers with no reported history of hearing impairments. They did not show impairments in speech

understanding. They were all right-handed, as assessed by the Edinburgh Handedness Inventory (Oldfield, 1971). Participants were recruited from a hospital or a self-help group for stuttering. An additional seven subjects (two adults, four school-age children, and one preschooler) were tested, but excluded from the analysis due to non-right-handedness. Stuttering severity was assessed with a 7-scale rating system from 1 (very mild) to 7 (very severe; Akahoshi et al., 1981; Johnson et al., 1963). The stuttering severity ranged from 2 to 7 in adults (2, $n = 2$; 4, $n = 4$; 5, $n = 2$; 6, $n = 1$; 7, $n = 1$), from 2 to 5 in school-age children (2, $n = 2$; 3, $n = 2$; 4, $n = 2$; 5, $n = 1$), and from 2 to 4 in preschoolers (2, $n = 2$; 3, $n = 3$; 4, $n = 1$). As data of age-matched normal controls for each stuttering group, we referred to data measured in previous studies (10 normal adults, 10 males, age range 20–32 years; 10 normal school-age children, 3 females and 7 males, age range 6–10 years; 8 preschool-age children, 4 females and 4 males, age range 3–5 years; Furuya et al., 2001; Furuya and Mori, 2003; Sato et al., 2003). The participants or their parents gave written informed consent before the experiment. This study was approved by the Ethical Committees of National Rehabilitation Center for Persons with Disabilities (NRCD).

STIMULI

Three different inflected forms of a Japanese verb /iku/ (meaning “to go”) were produced with a synthesis-by-analysis system (ASL, Kay Elemetrics Corp., USA) based on a speech signal recorded by a male adult (Imaizumi et al., 1998). By changing the formant frequencies and the vocal pitch contour, (A) past declarative /itta/ (“went”), (B) imperative /itte/ (“Go away”), and (C) interrogative /itta?/ (“went?”) forms of the verb were synthesized. These words consisted of a common initial /i/ vowel with a length of 80 ms, a 200-ms silent interval, and then the final syllable with a length of 80 ms. Only the final syllable was changed in the two derived words. The phonemic contrasting pair /itta/ and /itte/ had different final vowels due to the manipulation of the frequencies of their formants 1 and 2 but has an identical falling pitch pattern. The prosodic contrasting pair /itta/ and /itta?/ were only different in the pitch contours of the same final vowel.

NIRS RECORDING

Recordings of the changes in hemoglobin (Hb) concentrations in the bilateral temporal areas were made with multi-channel NIRS (ETG-100, Hitachi Medical Co., Japan; OMM-2001, Shimadzu, Japan), using near-infrared lasers at two (780 and 830 nm; ETG-100) or three wavelengths (780, 800, and 830 nm; OMM-2001). OMM-2001 was used only for the measurement of the school-age control group. OMM-2001 used the additional middle wavelength (800 nm), which should have the almost identical path-length *in vivo* as the other wavelengths. The recording channels resided in the optical path in the brain between the nearest pairs of incident and detection probes, which were separated by 3 cm on the scalp surface (Nakajima et al., 1993; Fukui et al., 2003). Five incident and four detection probes arranged in a 3 × 3 square lattice were placed on each lateral side of the head, which made the total number of recording channels 12 on either side because each pair of adjacent incident and detection probes constituted a single measurement channel (Figure 1). After the optical measurement, the positions of the optical probes were recorded with a three-dimensional (3-D) digitizer (Polhemus,

Colchester, VT, USA). The 3-D coordinates were superimposed onto T1-weighted MR brain images for each adult subject to identify the centers of recording sites. T1-weighted anatomical images were acquired in 80 contiguous axial slices with a thickness of 2.0 mm using a 1.5-T scanner [Excelart, Toshiba Medical, Japan; repetition time/echo time (TR/TE) 15/3.4 ms, flip angle (FA) 20°, matrix 256 × 192, field-of-view (FOV) 22 cm × 22 cm]. The channels close to the lateral end of the border between the transverse temporal gyrus and the planum temporale (PT), as projected onto a parasagittal MRI, should be in the auditory area and were referred to as auditory channels (Minagawa-Kawai et al., 2002; Furuya and Mori, 2003). This procedure selected the recording channels whose centers were within a 1.5-cm radius of the above-mentioned border. Thus, the channels should include the signals in the auditory cortex due to the spread of the laser in the brain tissue (Yamashita et al., 1996). Since it was difficult to acquire MR brain images of some young subjects, the positions of optical probes were recorded either with a 3-D digitizer or a digital camera for identification of approximate recording locations. Because the primary auditory cortex is located approximately 6 cm perpendicularly above the plane containing the bilateral preauricular point (PA) and the nasion, the channels at and around the height above, and the anteroposterior position at, the PA were presumed to be in or close to the auditory area, and referred to as auditory channels.

PROCEDURE

The experiments were carried out in a sound-attenuated room. Stimuli were presented at a comfortable level (60–70 dB SPL) via insert earphones (EAR TONE 3A) for adults and a loudspeaker (i15, TANNOY) for children in accordance with the previous studies for the control subjects (Furuya et al., 2001; Furuya and Mori, 2003; Sato et al., 2003). Because it was difficult to confirm that ear plugs were tightly pushed into ears in the measurements of children, the speaker was used for them. Each participant was tested in two conditions in respective sessions with a block design paradigm. In the phonemic condition, the baseline block contained only /itta/ which was repeated

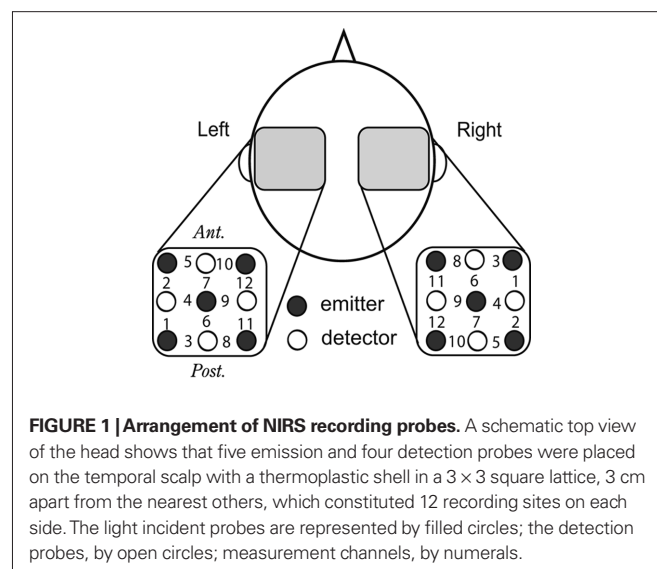


FIGURE 1 | Arrangement of NIRS recording probes. A schematic top view of the head shows that five emission and four detection probes were placed on the temporal scalp with a thermoplastic shell in a 3 × 3 square lattice, 3 cm apart from the nearest others, which constituted 12 recording sites on each side. The light incident probes are represented by filled circles; the detection probes, by open circles; measurement channels, by numerals.

approximately once every second (0.9–1.1 s), whereas the test block consisted of /itta/ and /itte/ presented in a pseudo-random order with equal probabilities at the same rate as in the baseline block. Baseline and test blocks each lasted for 20 s, and they were presented alternately at least five times. The prosodic condition was similar to the phonemic condition, except for the presentation of the /itta/ and /itta?/ combination in the test block. The presentation order of these two conditions was counterbalanced across subjects within each group.

DATA PROCESSING

Changes in the concentrations of Hb were calculated from the attenuation data of the laser beams. The total Hb responses during the test blocks in each condition were averaged synchronously and were smoothed with a 5-s moving average after manually excluding the blocks with gross motion artifacts. The maximal positive total Hb change was evaluated against the 10 s pre-test baseline period for each condition and for each auditory channel. To choose one of the auditory channels for statistical analysis, we first averaged the positive total Hb response across the two conditions and then selected the channel that exhibited the maximum value on each side (Minagawa-Kawai et al., 2007).

Because the measurement of Hb concentration obtained with continuous wave lasers lacks a reference to optical path-length, we cannot determine absolute values in principle. Consequently, the comparison or integration of data between different channels or across subjects may be difficult to justify. However, a recent study demonstrated that the optical path-lengths are similar among nearby channels and between homologous regions of left and right hemispheres within a subject (Katagiri et al., 2010). On the basis of these findings, the maximal values of total Hb changes in the left and right auditory channels were subjected to a two-way analysis of variance (ANOVA), with conditions (phoneme and prosody) and sides (left and right) as within-subject factors in each group. It should be noted that Katagiri et al.'s (2010) findings do not extend to a comparison across different subjects. Moreover, the two different NIRS systems that differ in the number of lasers used for calculating Hb concentrations (2 vs. 3) were used for the school-age children (ETG-100 for school-age children who stutter and OMM-2001 for school-age children who do not stutter). Since it is possible that the analyses using the Hb values from these machines have different sensitivities, the data were subjected to within-subject ANOVAs for each group, in which a single system was used.

In order to assess cerebral lateralization, a laterality index, $LI = (L - R)/(L + R)$, was calculated from the peaks of the averaged maximal total Hb responses in the left (L) and the right (R) auditory channels. Note that LI is not affected by the possible sensitivity difference of the recording systems. LI could range from -1 to 1, with a positive value indicating left dominance. We compared LI values between the two conditions (Wilcoxon signed-rank test) in each subject group. Subject-by-subject analysis was also performed: without averaging over repeated blocks, the left and right peaks of total Hb changes were collected from individual test blocks, for which respective LIs were calculated for comparison between the two contrast conditions within each subject (Mann–Whitney *U*-test).

In addition, the response peaks during the test blocks were compared with 0 (the average of the 10-s pre-test baseline) in each condition, side and group (one-sample *t*-test with false discovery rate (FDR) correction at $q < 0.05$ for each group). This was done to determine whether or not the phonemic or prosodic changes in the test blocks elicited significantly larger total Hb changes than the baseline blocks.

RESULTS

HEMODYNAMIC RESPONSES OF PEOPLE WHO DO AND DO NOT STUTTER

Figure 2 shows NIRS responses in adult, school-aged, and preschool-aged PWS, in addition to typical responses of a non-stuttering adult (Furuya and Mori, 2003). This figure indicates the averaged time courses of the total Hb during test blocks in each condition for a non-stuttering adult and all stuttering participants. Whereas the adult who did not stutter showed a larger response on the left side under the phonemic condition and a larger response on the right side under the prosodic condition, few of the stuttering participants showed such left–right reversal between the two conditions.

Figure 3 shows the averaged peak values of the total Hb responses in the left and right sides under the two conditions. All groups showed significant responses during the test blocks against zero baselines on the left and the right sides under each condition (one-sample *t*-test, $p < 0.05$ respectively).

The lateralization pattern across the phonemic and prosodic conditions differed between the stuttering and control groups. Results of two-way ANOVA, with condition (phoneme and prosody) and side (left and right) as within-subject factors in each adult group showed that the PWS exhibited no significant interaction or main effects for the two factors [$F(1,9) = 0.30$, $p > 0.10$; $F(1,9) = 0.02$, $p > 0.10$; $F(1,9) = 3.33$, $p > 0.10$, interaction, condition, side respectively], whereas the control (Ctrl) group exhibited a significant interaction [$F(1,9) = 12.07$, $p < 0.01$] but no effects for condition [$F(1,9) = 0.32$, $p > 0.10$] or side [$F(1,9) = 0.05$, $p > 0.10$], suggesting that the activations in the left and the right differed between the conditions only in the control group (**Figure 3A**). Simple effect tests showed that the left-side response was significantly larger than the right-side response under the phonemic condition (Holm correction, $p < 0.05$).

Similar results were observed in the school-age groups (**Figure 3B**). The school-age children who stutter exhibited no significant interaction or main effects for the two factors [$F(1,6) = 0.81$, $p > 0.10$; $F(1,6) = 0.06$, $p > 0.10$; $F(1,6) = 0.11$, $p > 0.10$, interaction, condition, side respectively]. In contrast, the school-age control group exhibited a significant interaction [$F(1,9) = 12.29$, $p < 0.01$] but no effects for condition [$F(1,9) = 0.00$, $p > 0.10$] or side [$F(1,9) = 1.98$, $p > 0.10$]. Simple effect tests showed the larger left-side response (compared to right side) under the phonemic condition ($p < 0.05$) and the larger right-side response (compared to left side) under the prosodic condition ($p < 0.05$).

The preschool children who stutter showed a different response pattern all together. They showed no significant interaction or main effect for the condition factor [$F(1,5) = 0.87$, $p > 0.10$; $F(1,5) = 0.00$, $p > 0.10$, interaction, condition respectively], but showed a main effect for the side factor [$F(1,5) = 69.55$, $p < 0.01$], suggesting that right-side activation was predominant under both conditions in this

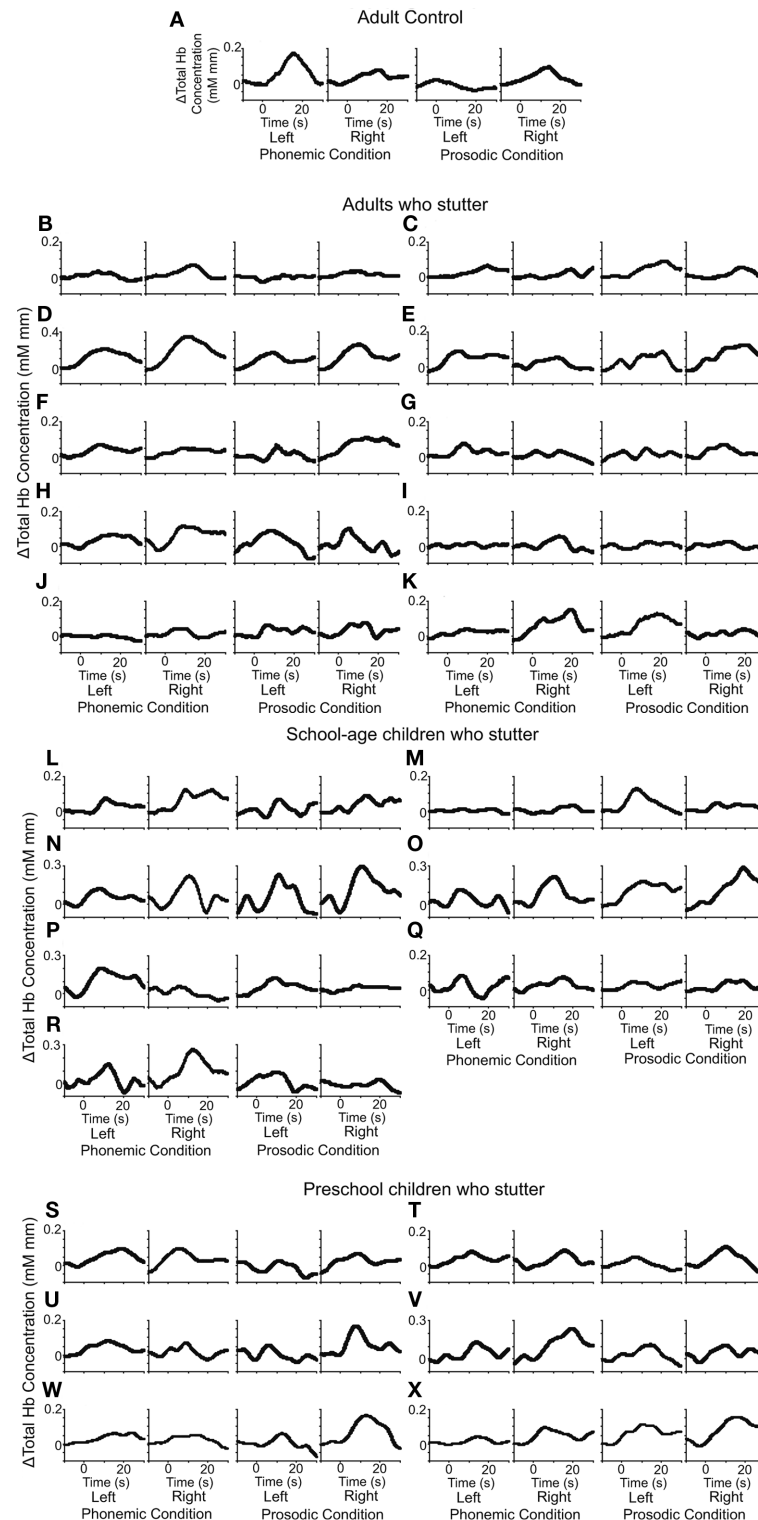
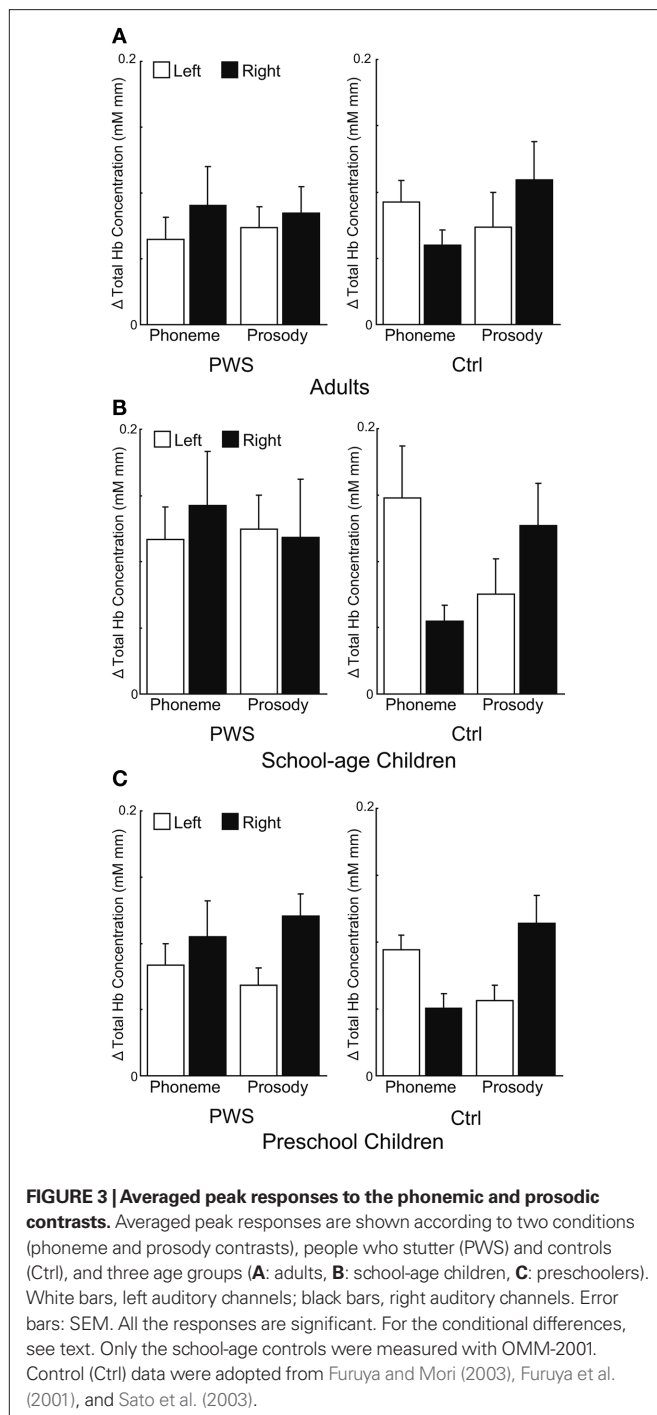


FIGURE 2 | Hemoglobin responses evoked by the phonemic and prosodic contrasts. Time courses of NIRS responses of total Hb in the temporal areas (left and right) are shown for individual subjects: the responses of a non-stuttering adult (adult control; data adopted from Furuya and Mori, 2003) (A), and the responses of people who stutter (PWS): adults (B–K), school-age children (L–R), and preschool children (S–X). The abscissas

indicate time and the ordinates indicate total Hb concentration changes. The vertical lines at 0 and 20 s show the beginning and the end of the test blocks, respectively. In (A), the larger responses are seen in the left side under the phonemic condition and in the right side under the prosodic condition. Few of the stuttering subjects showed this normal response pattern under either condition.



group (Figure 3C). In contrast, the preschool control group exhibited similar results as the adult and school-age control groups: they showed a significant interaction [$F(1,7) = 13.32, p < 0.01$] but no effects for condition [$F(1,7) = 0.80, p > 0.10$] or side [$F(1,7) = 0.38, p > 0.10$]. Simple effect tests showed the larger left-side response (compared to right side) under the phonemic condition ($p < 0.05$). Note that all three control groups showed the significant interaction of condition \times side factors despite of the use of the different measurement system for the school-age control group. This must

indicate that a system did not show a different sensitivity compared with the other one at least in this analysis. Moreover, similar results were observed between adults and children in either people who do or do not stutter group, suggesting that the different way of stimuli presentation, namely, by the earphones or the speaker, did not seriously impact on the current results.

ANALYSES OF LATERALITY INDEX

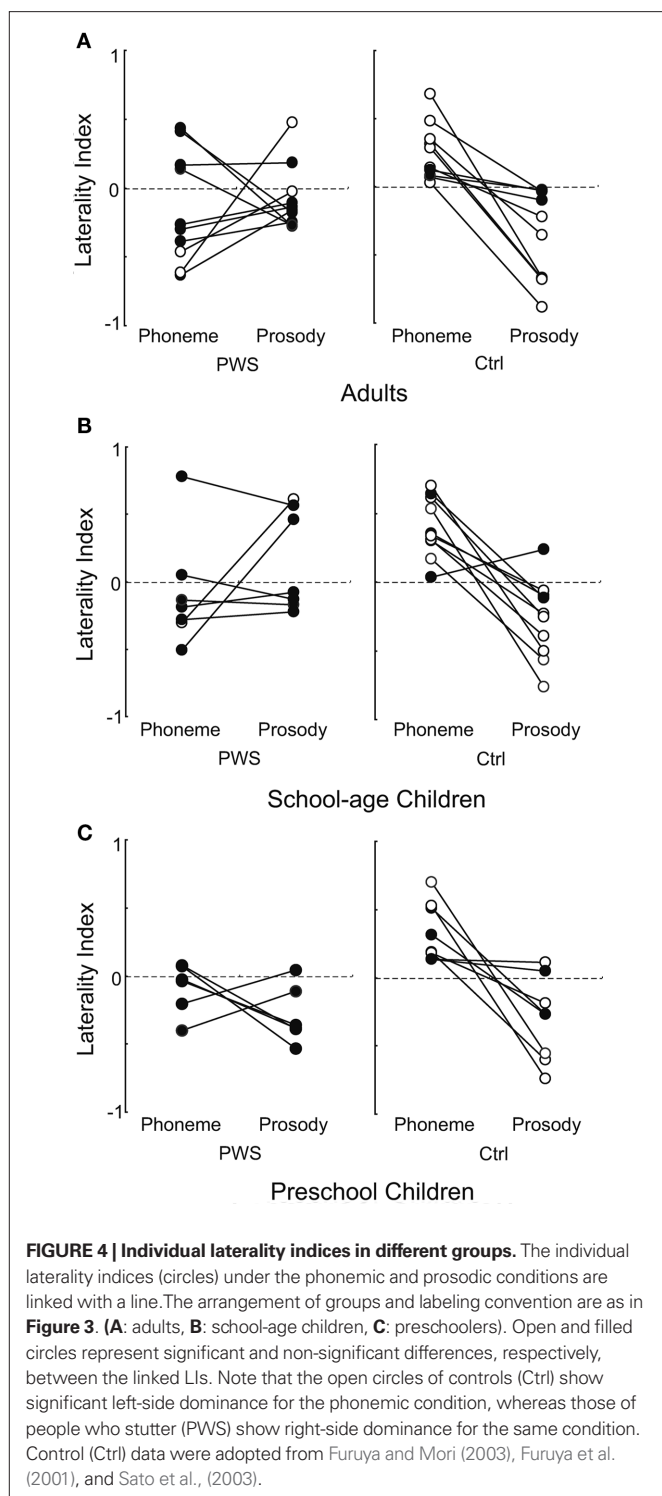
Figure 4 shows individual laterality indices under the phonemic and prosodic conditions in the three age groups with respective controls. First, we analyzed the differences in LIs between the two conditions in each subject group. The group analyses showed that each stuttering group failed to show significant differences in LIs between the two contrast conditions (Wilcoxon signed-rank test, adult, $p > 0.10$; school-age, $p > 0.10$; preschool, $p > 0.10$). On the other hand, all control groups showed significant differences in LIs between the two contrast conditions (adult control, $p < 0.01$; school-age control, $p < 0.01$; preschool control, $p < 0.05$).

SUBJECT-BY-SUBJECT ANALYSIS OF LATERALITY INDEX FOR THE TWO CONDITIONS

Subject-by-subject analysis revealed that two adults (20%) and one school-age child (14%) who stutter showed significant differences between their respective phonemic and prosodic LIs (white circles in Figure 4). The remaining subjects who stutter (adults 80%; school-age 86%; preschool 100%) showed no significant difference in LI between the two conditions (filled circles in Figure 4). Note that the significant differences of the three subjects who stutter were due to the rightward LIs for the phonemic contrast in comparison to that for the other condition. This was opposite the normal control subjects: seven of the adults (70%), seven of the school-aged (70%), and five of the preschool (63%) children who do not stutter showed significant leftward LIs for the phonemic contrast. The remaining control subjects (adults 30%; school-age 30%; preschool 37%) showed no significant difference in the LI between the two conditions. The ratios of subjects in the response laterality patterns differed significantly between the people who do and do not stutter in each age group (Fisher's exact test, adult, $p < 0.01$; school-age, $p < 0.01$; preschool, $p < 0.05$).

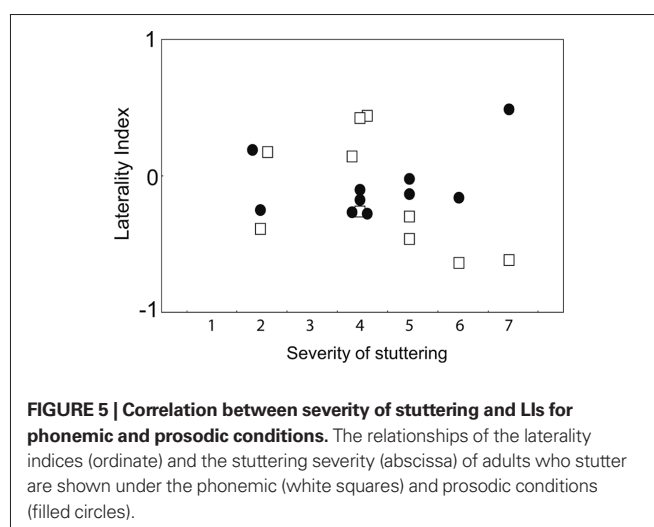
CORRELATION BETWEEN LATERALITY INDEX AND STUTTERING SEVERITY

Correlation analyses indicated that stuttering severity was negatively correlated with the LI under the phonemic condition, but not significantly correlated with the LI under the prosodic condition in adults who stutter (Spearman's $\rho = -0.65, p < 0.05$; Spearman's $\rho = 0.36, p > 0.05$, phonemic and prosodic conditions, respectively; Figure 5). In contrast, neither school-age nor preschool-age children who stutter showed significant correlations between severity and LI under either condition (school-age: phonemic condition, Spearman's $\rho = 0.26, p > 0.05$; prosodic condition, Spearman's $\rho = 0.53, p > 0.05$; preschool-age: phonemic condition, Spearman's $\rho = -0.31, p > 0.05$; prosodic condition, Spearman's $\rho = 0.19, p > 0.05$). Although we analyzed the combined data of school-age and preschool children, no significant correlation was observed between severity and LI under either condition (phonemic condition, Spearman's $\rho = 0.03, p > 0.05$; prosodic condition, Spearman's $\rho = 0.45, p > 0.05$).



DISCUSSION

To examine whether abnormal functional lateralization is associated with the onset or the result of stuttering, we used NIRS to measure brain responses of PWS during the auditory processing of phonemic and prosodic contrasts, and compared these data with age-matched control data. The Hb data analyses revealed that all control groups showed differential left–right activation patterns



between the two contrast conditions. That is, normal controls showed left-side and right-side dominance for the phonemic and the prosodic contrasts, respectively, whereas the adult and school-age stuttering groups showed no lateralized responses to either contrast. The preschool children who stutter, on the other hand, showed right-side dominance under both conditions. The group analysis of LIs found the abnormal lateralization in PWS: all control groups showed left-dominant responses to the phonemic contrast compared to the prosodic contrast, but no stuttering group showed differential lateralized responses between the two contrasts. This abnormality in functional lateralization during auditory processing is in line with previous studies using a dichotic listening paradigm (Curry and Gregory, 1969; Brady and Berson, 1975; Sommers et al., 1975; Blood, 1985) and with one using MEG (Salmelin et al., 1998). The current results in children who stutter are compatible with the previous studies showing that brain responses to auditory processing differ between children who do and do not stutter (Kaganovich et al., 2010; Beal et al., 2011). The present study confirms the abnormal auditory functional lateralization seen in PWS, and extends it from adults to down to the preschool age (i.e., shortly after the onset of stuttering). This seems to imply that it is relevant to the onset of stuttering rather than simply a consequence of long-term stuttering.

Previous studies using the same paradigm as in the current study have shown that infants around 1 year old already show a significant leftward shift of responses to phonemic contrasts compared with those to non-phonemic contrasts (Sato et al., 2003; Minagawa-Kawai et al., 2007). The present result of the atypical functional lateralization of children who stutter at 3–5 years of age still leaves room for further research to investigate how the abnormality is established. Two alternatives are possible: children who are at risk for stuttering might reset the normal lateralization they once had at around 1 year of age, or they never develop the normal pattern of lateralization before they start stuttering. Longitudinal studies of younger children of 1–3 years old are needed in order to examine the causal relationship between the abnormality and the onset of stuttering.

Adults who stutter show reduced asymmetry in PT (Foundas et al., 2001) and increased white matter volumes in the right hemisphere, including superior temporal gyrus (Jäncke et al., 2004). It

has been also reported that PWS with atypical PT asymmetry (i. e., rightward PT asymmetry) show more disfluency than PWS with typical PT asymmetry ($L > R$), and PWS with atypical PT asymmetry improved disfluency more with delayed auditory feedback (DAF) than those who stutter with typical PT asymmetry (Foundas et al., 2004). As alteration in the auditory signal by DAF might correct an auditory perceptual defect, it is possible that anomalous auditory processing is related to structural abnormalities in PWS. If this is true, not only anatomical PT asymmetry but anomalous functional lateralization found in this study might be clinically applicable to a prediction of improvement of disfluency by DAF. We also have to consider that brain anatomy can be changed by behavior or function. Since alterations in behavior such as music training, handedness, and language proficiency can produce changes in brain anatomy (Gaser and Schlaug, 2003; Buchel et al., 2004; Mechelli et al., 2004), most of the anatomic anomalies in PWS could be the result of either the long-term dysfunctions in auditory processing or the insufficient connection between the anterior and posterior language areas and the left motor cortex (Sommer et al., 2002; Watkins et al., 2008). In fact, no increases in volume were found in the right-hemisphere speech regions in children who stutter, and no reduction in asymmetry in the PT were observed (Chang et al., 2008). Our findings of abnormal functional lateralization as early as 3–5 years thus suggest that the anatomical abnormalities in the posterior language cortices develop as the result of more than 15 years of anomalous functional lateralization in the auditory area.

In the present study, all three age groups of PWS demonstrated similar results in terms of the lack of the stimulus dependence of LI: non-significant differences of LI were found between the phonemic and prosodic contrast conditions. Nevertheless, the Hb data revealed a difference in the preschool children who stutter, in contrast with the other two stuttering groups: only the preschoolers showed larger right-side responses to both phonemic and prosodic contrasts. It is possible that the affected arcuate fasciculus (Sommer et al., 2002; Chang et al., 2008) with smaller Broca's area (Chang et al., 2008) exerts a disruptive influence on the left auditory cortex and on the interaction between the left anterior and posterior language areas in the rapidly developing immature brain, thus shifting the responses to both stimuli to the right. Later, compensatory connections (yet to be specified) may restore the activation in the left auditory area to some extent especially by prosody. Alternatively, if the prosodic processing is intact in early childhood, it should be handled in the right side starting at 1 year of age (Sato et al., 2003). If the phonemic processing has to be handled by the right side due to inefficient connections between the left auditory and the anterior language areas, both phonemic and prosodic functions would be predominantly handled in the right auditory area, of which the

prosodic processing would later shift to the left to make more room for the increasing demand of phonemic processing in the right. The former is possibly less demanding for timing, so that it could be handled even with the defective arcuate fasciculus, or through the intact ventral pathway to the frontal area. The shift to the left of the prosodic processing may be related to poor perception in PWS of linguistic stimuli with prosodic information (Blood, 1996). Clearly, more longitudinal studies are needed to clarify these points.

The correlation analyses in the current study revealed that in the adults who stutter, stuttering severity was negatively correlated with the LI for the phonemic contrast, whereas neither school-age nor preschool stuttering groups showed significant correlations. This indicates that more severe stuttering symptoms are correlated with more abnormal lateralization patterns for the phonemic contrasts. This finding confirms the close relationship between the auditory deficit and stuttering. The results of the children, however, should not be directly compared to that of the adults due to the narrower distribution of stuttering severity.

To conclude, the NIRS method is practical not only for elucidating neural correlates of stuttering in children and adults alike, but also for evaluating stuttering individually. Due to the paucity of childhood functional studies, there has been long-standing controversy whether the ambiguous lateralization for cerebral linguistic processing is a cause or a result of stuttering. The current study using NIRS shows the progress of abnormal lateralization of receptive speech processing in the posterior language area, and tracks the anomaly down to the preschool age, soon after the onset of stuttering, thus narrowing down the search period for causality to between 1 year and the onset of stuttering (3–5 years). Since the NIRS method is well-suited for studying children and infants, future longitudinal studies using NIRS should greatly contribute to elucidating the pathophysiology of developmental stuttering. Subject-by-subject analysis in this study revealed that not a single subject who stutters showed a left-hemisphere advantage in the phonemic contrast processing over the prosodic contrast processing. Together with the significant correlation between abnormal lateralization in the phonemic condition and the severity of stuttering, these results may provide useful information for the clinical prognosis and treatment of stuttering.

ACKNOWLEDGMENTS

The authors would like to thank Satoshi Imaizumi for kindly providing the sound stimuli and Izumi Furuya, Ryoko Hayashi and Yoshimasa Sakata for their technical advice and contributions. This work was supported by grants from the Ministry of Health, Labor and Welfare of Japan to the second author (Koichi Mori; H14-Kokoro-001, 15130801, and H16-Shogai-001).

REFERENCES

- Akahoshi, S., Ozawa, E., Kunishima, K., Suzuki, N., Doi, A., Fukawa, A., and Moriama, H. (1981). Tentative assessment procedure for stuttering. *Jpn. J. Logoped. Phoniatr.* 22, 194–208.
- Beal, D. S., Cheyne, D. O., Gracco, V. L., Quraan, M. A., Taylor, M. J., and De Nil, L. F. (2010). Auditory evoked fields to vocalization during passive listening and active generation in adults who stutter. *Neuroimage* 52, 1645–1653.
- Beal, D. S., Quraan, M. A., Cheyne, D. O., Taylor, M. J., Gracco, V. L., and De Nil, L. F. (2011). Speech-induced suppression of evoked auditory fields in children who stutter. *Neuroimage* 54, 2994–3003.
- Biermann-Ruben, K., Salmelin, R., and Schnitzler, A. (2005). Right rolandic activation during speech perception in stutterers: a MEG study. *Neuroimage* 25, 793–801.
- Blomgren, M., and Goberman, A. M. (2008). Revisiting speech rate and utterance length manipulations in stuttering speakers. *J. Commun. Disord.* 41, 159–178.
- Blood, G. W. (1985). Laterality differences in child stutterers: heterogeneity levels, and statistical treatments. *J. Speech Hear. Res.* 50, 66–72.
- Blood, I. M. (1996). Disruptions in auditory and temporal processing in adults who stutter. *Percept. Mot. Skills* 82, 272–274.
- Bloodstein, O. (1995). *A Handbook on Stuttering*. San Diego: Singular.
- Brady, J. P. (1969). Studies on the metronome effect on stuttering. *Behav. Res. Ther.* 7, 197–204.
- Brady, J., and Berson, J. (1975). Stuttering, dichotic listening, and cerebral

- dominance. *Arch. Gen. Psychiatry* 32, 1449–1452.
- Braun, A. R., Varga, M., Stager, S., Schulz, G., Selbie, S., Maisog, J. M., Carson, R. E., and Ludlow, C. L. (1997). Altered patterns of cerebral activity during speech and language production in developmental stuttering. *An H2(15) O positron emission tomography study. Brain* 120, 761–784.
- Brown, S., Ingham, R. J., Ingham, J. C., Laird, A. R., and Fox, P. T. (2005). Stuttered and fluent speech production: an ALE meta-analysis of functional neuroimaging studies. *Hum. Brain Mapp.* 25, 105–117.
- Bruce, M. C., and Adams, M. R. (1978). Effects of two types of motor practice on stuttering adaptation. *J. Speech Hear. Res.* 21, 421–428.
- Buchel, C., Raedler, T., Sommer, M., Sach, M., Weiller, C., and Koch, M. A. (2004). White matter asymmetry in the human brain: a diffusion tensor MRI study. *Cereb. Cortex* 14, 945–951.
- Chang, S. E., Erickson, K. I., Ambrose, N. G., Hasegawa-Johnson, M. A., and Ludlow, C. L. (2008). Brain anatomy differences in childhood stuttering. *Neuroimage* 39, 1333–1344.
- Chang, S. E., Kenney, M. K., Loucks, T. M., and Ludlow, C. L. (2009). Brain activation abnormalities during speech and non-speech in stuttering speakers. *Neuroimage* 46, 201–212.
- Chen, S., Sakatani, K., Lichty, W., Ning, P., Zhao, S., and Zuo, H. (2002). Auditory-evoked cerebral oxygenation changes in hypoxic-ischemic encephalopathy of newborn infants monitored by near infrared spectroscopy. *Early Hum. Dev.* 67, 113–121.
- Cherry, C., and Sayers, B. (1956). Experiments upon the total inhibition of stammering by external control, and some clinical results. *J. Psychosom. Res.* 1, 233–246.
- Conture, E. G., Rothenberg, M., and Molitor, R. D. (1986). Electroglossographic observations of young stutterers' fluency. *J. Speech Hear. Res.* 29, 384–393.
- Corbera, S., Corral, M. J., Escera, C., and Idiazábal, M. A. (2005). Abnormal speech sound representation in persistent developmental stuttering. *Neurology* 65, 1246–1252.
- Curry, F. K. W., and Gregory, H. H. (1969). The performance of stutterers on dichotic listening tasks thought to reflect cerebral dominance. *J. Speech Hear. Res.* 12, 73–82.
- Cykowski, M., Kochunov, P., Ingham, R. J., Ingham, J. C., Mangin, J. F., Rivière, D., Lancaster, J. L., and Fox, P. T. (2008). Perisylvian sulcal morphology and cerebral asymmetry patterns in adults who stutter. *Cereb. Cortex* 18, 571–583.
- De Nil, L. F., Beal, D. S., Lafaille, S. J., Kroll, R. M., Crawley, A. P., and Gracco, V. L. (2008). The effects of simulated stuttering and prolonged speech on the neural activation patterns of stuttering and nonstuttering adults. *Brain Lang.* 107, 114–123.
- De Nil, L. F., Kroll, R. M., Kapur, S., and Houle, S. (2000). A positron emission tomography study of silent and oral single word reading in stuttering and nonstuttering adults. *J. Speech Hear. Res.* 43, 1038–1053.
- Denny, M., and Smith, A. (1992). Gradations in a pattern of neuromuscular activity associated with stuttering. *J. Speech Hear. Res.* 35, 1216–1229.
- Foundas, A. L., Bollich, A. M., Corey, D. M., Hurley, M., and Heilman, K. M. (2001). Anomalous anatomy of speech-language areas in adults with persistent developmental stuttering. *Neurology* 57, 207–215.
- Foundas, A. L., Bollich, A. M., Feldman, J., Corey, D. M., Hurley, M., Lemen, L. C., and Heilman, K. M. (2004). Aberrant auditory processing and atypical planum temporale in developmental stuttering. *Neurology* 63, 1640–1646.
- Foundas, A. L., Corey, D. M., Angeles, V., Bollich, A. M., Crabtree-Hartman, E., and Heilman, K. M. (2003). Atypical cerebral laterality in adults with persistent developmental stuttering. *Neurology* 61, 1378–1385.
- Fox, P. T., Ingham, R. J., Ingham, J. C., Hirsch, T. B., Downs, J. H., Martin, C., Jerabek, P., Glass, T., and Lancaster, J. L. (1996). A PET study of the neural systems of stuttering. *Nature* 382, 158–161.
- Fox, P. T., Ingham, R. J., Ingham, J. C., Zamarripa, F., Xiong, J. H., and Lancaster, J. L. (2000). Brain correlates of stuttering and syllable production. A PET performance-correlation analysis. *Brain* 123, 1985–2004.
- Fukui, Y., Ajichi, Y., and Okada, E. (2003). Monte Carlo prediction of near-infrared light propagation in realistic adult and neonatal head models. *Appl. Opt.* 42, 2881–2887.
- Furuya, I., and Mori, K. (2003). Cerebral lateralization in spoken language processing measured by multi-channel near-infrared spectroscopy (NIRS). *Brain Nerve* 55, 226–231.
- Furuya, I., Mori, K., Minagawa, Y., and Hayashi, R. (2001). Cerebral lateralization of speech processing in infants measured by near-infrared spectroscopy. *Neurosci. Res.* S25, S125.
- Gaser, C., and Schlaug, G. (2003). Gray matter differences between musicians and nonmusicians. *Ann. N. Y. Acad. Sci.* 999, 514–517.
- Giraud, A. L., Neumann, K., Bachoud-Levi, A. C., von Gudenberg, A. W., Euler, H. A., Lanfermann, H., and Preibisch, C. (2008). Severity of dysfluency correlates with basal ganglia activity in persistent developmental stuttering. *Brain Lang.* 104, 190–199.
- Gordon, P., Luper, H. L., and Peterson, H. A. (1986). The effects of syntactic complexity on the occurrence of disfluencies in 5 year old nonstutterers. *J. Fluency Disord.* 11, 151–164.
- Guitar, B. (2006). *Stuttering: An Integrated Approach to Its Nature and Treatment*. Philadelphia: Lippincott Williams & Wilkins.
- Hall, J. W., and Jerger, J. (1978). Central auditory function in stutterers. *J. Speech Hear. Res.* 21, 324–337.
- Hampton, A., and Weber-Fox, C. (2008). Non-linguistic auditory processing in stuttering: evidence from behavior and event-related brain potentials. *J. Fluency Disord.* 33, 253–273.
- Hargrave, S., Kalinowski, J., Stuart, A., Armon, J., and Jones, K. (1994). Effect of frequency-altered feedback on stuttering frequency at normal and fast speech rates. *J. Speech Hear. Res.* 37, 1313–1319.
- Homae, F., Watanabe, H., Nakano, T., Asakawa, K., and Taga, G. (2006). The right hemisphere of sleeping infant perceives sentential prosody. *Neurosci. Res.* 54, 276–280.
- Homae, F., Watanabe, H., Nakano, T., and Taga, G. (2007). Prosodic processing in the developing brain. *Neurosci. Res.* 59, 29–39.
- Imaizumi, S., Mori, K., Kiritani, S., Hosoi, H., and Tonoike, M. (1998). Task-dependent laterality for cue decoding during spoken language processing. *Neuroreport* 9, 899–903.
- Ingham, R. J. (2001). Brain imaging studies of developmental stuttering. *J. Commun. Disord.* 34, 493–516.
- Ingham, R. J., Fox, P. T., Ingham, J. C., Xiong, J. H., Zamarripa, F., Hardies, L. J., and Lancaster, J. L. (2004). Brain correlates of stuttering and syllable production: gender comparison and replication. *J. Speech Hear. Res.* 47, 321–341.
- Ingham, R. J., Fox, P. T., Ingham, J. C., and Zamarripa, F. (2000). Is overt stuttered speech a prerequisite for the neural activations associated with chronic developmental stuttering? *Brain Lang.* 75, 163–194.
- Jäncke, L., Hänggi, J., and Steinmetz, H. (2004). Morphological brain differences between adult stutterers and non-stutterers. *BMC Neurol.* 4, 23. doi: 10.1186/1471-2377-4-23
- Jayaram, M. (1984). Distribution of stuttering in sentences: relationship to sentence length and clause position. *J. Speech Hear. Res.* 27, 338–341.
- Johnson, W., Darley, F. L., and Spiersbach, D. C. (1963). *Diagnostic Methods in Speech Pathology*. New York: Harper & Row.
- Johnson, W., and Rosen, L. (1937). Studies in the psychology of stuttering: VII. Effect of certain changes in speech pattern upon frequency of stuttering. *J. Speech Disord.* 2, 105–109.
- Kaganovich, N., Wray, A. H., and Weber-Fox, C. (2010). Non-linguistic auditory processing and working memory update in pre-school children who stutter: an electrophysiological study. *Dev. Neuropsychol.* 35, 712–736.
- Katagiri, A., Dan, I., Tuzuki, D., Okamoto, M., Yokose, N., Igarashi, K., Hoshino, T., Fujiwara, T., Katayama, Y., Yamaguchi, Y., and Sakatani, K. (2010). Mapping of optical pathlength of human adult head at multi-wavelengths in near infrared spectroscopy. *Adv. Exp. Med. Biol.* 662, 205–212.
- Kell, C. A., Neumann, K., von Kriegstein, K., Posenenske, C., von Gudenberg, A. W., Euler, H., and Giraud, A. L. (2009). How the brain repairs stuttering. *Brain* 132, 2747–2760.
- Kennan, R. P., Kim, D., Maki, A., Koizumi, H., and Constable, R. T. (2002). Non-invasive assessment of language lateralization by transcranial near infrared optical topography and functional MRI. *Hum. Brain Mapp.* 16, 183–189.
- Kikuchi, Y., Ogata, K., Umesaki, T., Yoshiura, T., Kenjo, M., Hirano, Y., Okamoto, T., Komune, S., and Tobimatsu, S. (2011). Spatiotemporal signatures of an abnormal auditory system in stuttering. *Neuroimage* 55, 891–899.
- Lawrence, M., and Barclay, D. M. III. (1998). Stuttering: a brief review. *Am. Fam. Physician* 57, 2175–2178.
- Liotti, M., Ingham, J. C., Takai, O., Paskos, D. K., Perez, R., and Ingham, R. J. (2010). Spatiotemporal dynamics of speech sound perception in chronic developmental stuttering. *Brain Lang.* 115, 141–147.
- Lu, C., Peng, D., Chen, C., Ning, N., Ding, G., Li, K., Yang, Y., and Lin, C. (2010). Altered effective connectivity and anomalous anatomy in the basal ganglia-thalamocortical circuit of stuttering speakers. *Cortex* 46, 49–67.
- Ludlow, C. L., and Loucks, T. (2003). Stuttering: a dynamic motor control disorder. *J. Fluency Disord.* 28, 273–295.
- McClean, M. D., and Runyan, C. M. (2000). Variations in the relative speeds of orofacial structures with stuttering severity. *J. Speech Hear. Res.* 43, 1524–1531.
- Mechelli, A., Crinion, J. T., Noppeney, U., O'Doherty, J., Ashburner, J., Frackowiak, R. S., and Price, C. J. (2004). Neurolinguistics: structural plasticity in the bilingual brain. *Nature* 431, 757.

- Melnick, K. S., and Conture, E. G. (2000). Relationship of length and grammatical complexity to the systematic and nonsystematic speech errors and stuttering of children who stutter. *J. Fluency Disord.* 25, 21–45.
- Minagawa-Kawai, Y., Mori, K., Furuya, I., Hayashi, R., and Sato, Y. (2002). Assessing cerebral representations of short and long vowel categories by NIRS. *Neuroreport* 13, 581–584.
- Minagawa-Kawai, Y., Mori, K., Hebden, J. C., and Dupoux, E. (2008). Optical imaging of infants' neurocognitive development: recent advances and perspectives. *Dev. Neurobiol.* 68, 712–728.
- Minagawa-Kawai, Y., Mori, K., Naoi, N., and Kojima, S. (2007). Neural attunement processes in infants during the acquisition of a language-specific phonemic contrast. *J. Neurosci.* 27, 315–321.
- Minagawa-Kawai, Y., van der Lely, H., Ramus, F., Sato, Y., Mazuka, R., and Dupoux, E. (2011). Optical brain imaging reveals general auditory and language-specific processing in early infant development. *Cereb. Cortex* 21, 254–261.
- Nakajima, T., Miki, A., Kuwabara, T., Watanabe, H., Shimizu, K., Takeo, K., Iwamoto, S., Fujita, M., Naruse, S., and Yuasa, T. (1993). Functional mapping of the human visual system with near-infrared spectroscopy and BOLD functional MRI. *J. Appl. Physiol.* 75, 1842–1846.
- Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia* 9, 97–113.
- Özge, A., Toros, F., and Cömeleko lu, U. (2004). The role of hemispherical asymmetry and regional activity of quantitative EEG in children with stuttering. *Child Psychiatry Hum. Dev.* 34, 269–280.
- Peña, M., Maki, A., Kovacic, D., Dehaene-Lambertz, G., Koizumi, H., Bouquet, F., and Mehler, J. (2003). Sounds and silence: an optical topography study of language recognition at birth. *Proc. Natl. Acad. Sci. U.S.A.* 103, 11702–11705.
- Peters, H. F. M., and Boves, L. (1988). Coordination of aerodynamic and phonatory processes in fluent speech of stutterers. *J. Speech Hear. Res.* 31, 352–361.
- Preibisch, C., Neumann, K., Raab, P., Euler, H. A., von Gudenberg, A. W., Lanfermann, H., and Giraud, A. L. (2003). Evidence for compensation for stuttering by the right frontal operculum. *Neuroimage* 20, 1356–1364.
- Sakai, N., Masuda, S., Shimotomai, T., and Mori, K. (2009). Brain activation in adults who stutter under delayed auditory feedback: an fMRI study. *Int. J. Speech Lang. Pathol.* 11, 2–11.
- Salmelin, R., Schnitzler, A., Schmitz, F., and Freund, H. J. (2000). Single word reading in developmental stutterers and fluent speakers. *Brain* 123, 1184–1202.
- Salmelin, R., Schnitzler, A., Schmitz, F., Jäncke, L., Witte, O. W., and Freund, H. J. (1998). Functional organization of the auditory cortex is different in stutterers and fluent speakers. *Neuroreport* 9, 2225–2229.
- Sato, Y., Mori, K., Furuya, I., Hayashi, R., Minagawa-Kawai, Y., and Koizumi, T. (2003). Developmental changes in cerebral lateralization to spoken language in infants: measured by near-infrared spectroscopy. *Jpn. J. Logoped. Phoniatr.* 44, 165–171.
- Sato, Y., Sogabe, Y., and Mazuka, R. (2010). Development of hemispheric specialization for lexical pitch–accent in Japanese infants. *J. Cogn. Neurosci.* 22, 2503–2513.
- Sawyer, J., Chon, H., and Ambrose, N. G. (2008). Influences of rate, length, and complexity on speech disfluency in a single-speech sample in preschool children who stutter. *J. Fluency Disord.* 33, 220–240.
- Smith, A. (1995). “Muscle activity in stuttering,” in *Proceedings of the 1st World Congress of Fluency Disorders*, eds C. W. Starkweather and H. F. M. Peters (Nijmegen: Nijmegen University Press), 39–42.
- Smith, A., Luschei, E., Denny, M., Wood, J., Hirano, M., and Badylak, S. (1993). Spectral analyses of laryngeal and orofacial muscles in stutterers. *J. Neurol. Neurosurg. Psychol.* 56, 1301–1311.
- Smith, A., Sadagopan, N., Walsh, B., and Weber-Fox, C. (2010). Increasing phonological complexity reveals heightened instability in inter-articulatory coordination in adults who stutter. *J. Fluency Disord.* 35, 1–18.
- Sommer, M., Koch, M. A., Paulus, W., Weiller, C., and Büchel, C. (2002). Disconnection of speech-relevant brain areas in persistent developmental stuttering. *Lancet* 360, 380–383.
- Sommers, R. K., Brady, W. A., and Moore, W. J. Jr. (1975). Dichotic ear preferences of stuttering children and adults. *Percept. Mot. Skills* 41, 931–938.
- Toscher, M. M., and Rupp, R. R. (1978). A study of the central auditory processes in stutterers using the Synthetic Sentence Identification (SSI) Test battery. *J. Speech Hear. Res.* 21, 779–792.
- Trotter, W. D., and Silverman, F. H. (1974). Does the effect of pacing speech with a miniature metronome on stuttering wear off? *Percept. Mot. Skills* 39, 429–430.
- van Riper, C. (1971). *The Nature of Stuttering*. Englewood Cliffs, NJ: Prentice-Hall, Inc.
- Watkins, K. E., Smith, S. M., Davis, S., and Howell, P. (2008). Structural and functional abnormalities of the motor system in developmental stuttering. *Brain* 131, 50–59.
- Weber-Fox, C., and Hampton, A. (2008). Stuttering and natural speech processing of semantic and syntactic constraints on verbs. *J. Speech Hear. Res.* 51, 1058–1071.
- Weber-Fox, C., Spencer, R., Spruill, J. E. III, and Smith, A. (2004). Phonological processing in adults who stutter: electrophysiological and behavioral evidence. *J. Speech Lang. Hear. Res.* 47, 1244–1258.
- Weber-Fox, C., Spruill, J. E. III, Spencer, R., and Smith, A. (2008). Atypical neural functions underlying phonological processing and silent rehearsal in children who stutter. *Dev. Sci.* 11, 321–337.
- Wu, J. C., Maguire, G., Riley, G., Fallon, J., LaCasse, L., Chin, S., Klein, E., Tang, C., Cadwell, S., and Litzenberg, S. (1995). A positron emission tomography [18F] deoxyglucose study of developmental stuttering. *Neuroreport* 6, 501–505.
- Yairi, E., and Ambrose, N. G. (1999). Early childhood stuttering I: persistency and recovery rates. *J. Speech Lang. Hear. Res.* 42, 1097–1112.
- Yamamoto, T., Maki, A., Kadoya, T., Tanikawa, Y., Yamad, Y., Okada, E., and Koizumi, H. (2002). Arranging optical fibres for the spatial resolution improvement of topographical images. *Phys. Med. Biol.* 47, 3429–3440.
- Yamashita, Y., Maki, A., and Koizumi, H. (1996). Near-infrared topographic measurement system: imaging of absorbers localized in a scattering medium. *Rev. Sci. Instrum.* 67, 730–732.
- Zackheim, C. T., and Conture, E. G. (2003). Childhood stuttering and speech disfluencies in relation to children's mean length of utterance: a preliminary study. *J. Fluency Disord.* 28, 115–141.
- Zaramella, P., Freato, F., Amigoni, A., Salvadori, S., Marangoni, P., Suppiej, A., Schiavo, B., and Chiandetti, L. (2001). Brain auditory activation measured by near-infrared spectroscopy (NIRS) in neonates. *Pediatr. Res.* 49, 213–219.
- Zimmerman, G. N. (1980). Articulatory behaviors associated with stuttering: a cineradiographic analysis. *J. Speech Lang. Hear. Res.* 23, 108–121.
- Zocchi, L., Estenne, M., Johnston, S., Del Ferro, L., Ward, M. E., and Macklem, P. T. (1990). Respiratory muscle incoordination in stuttering speech. *Am. Rev. Respir. Dis.* 141, 1510–1515.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 14 December 2010; paper pending published: 22 January 2011; accepted: 04 April 2011; published online: 27 April 2011.
Citation: Sato Y, Mori K, Koizumi T, Minagawa-Kawai Y, Tanaka A, Ozawa E, Wakaba Y and Mazuka R (2011) Functional lateralization of speech processing in adults and children who stutter. *Front. Psychology* 2:70. doi: 10.3389/fpsyg.2011.00070
This article was submitted to *Frontiers in Language Sciences*, a specialty of *Frontiers in Psychology*.
Copyright © 2011 Sato, Mori, Koizumi, Minagawa-Kawai, Tanaka, Ozawa, Wakaba and Mazuka. This is an open-access article subject to a non-exclusive license between the authors and Frontiers Media SA, which permits use, distribution and reproduction in other forums, provided the original authors and source are credited and other Frontiers conditions are complied with.



Large-scale brain networks underlying language acquisition in early infancy

Fumitaka Homae^{1*}, Hama Watanabe², Tamami Nakano³ and Gentaro Taga^{2,4}

¹ Department of Language Sciences, Tokyo Metropolitan University, Tokyo, Japan

² Graduate School of Education, University of Tokyo, Tokyo, Japan

³ Department of Neurophysiology, Juntendo University, Tokyo, Japan

⁴ Core Research for Evolutional Science and Technology, Japan Science and Technology Agency, Kawaguchi, Japan

Edited by:

Judit Gervain, CNRS – Université Paris Descartes, France

Reviewed by:

Richard N. Aslin, University of Rochester, USA

Marco Ferrari, University of L'Aquila, Italy

*Correspondence:

Fumitaka Homae, Department of Language Sciences, Tokyo Metropolitan University, 1-1 Minami Osawa, Hachioji-shi, Tokyo 192-0397, Japan.
e-mail: fhomae@tmu.ac.jp

A critical issue in human development is that of whether the language-related areas in the left frontal and temporal regions work as a functional network in preverbal infants. Here, we used 94-channel near-infrared spectroscopy to reveal the functional networks in the brains of sleeping 3-month-old infants with and without presenting speech sounds. During the first 3 min, we measured spontaneous brain activation (period 1). After period 1, we provided stimuli by playing Japanese sentences for 3 min (period 2). Finally, we measured brain activation for 3 min without providing the stimulus (period 3), as in period 1. We found that not only the bilateral temporal and temporoparietal regions but also the prefrontal and occipital regions showed oxygenated hemoglobin signal increases and deoxygenated hemoglobin signal decreases when speech sounds were presented to infants. By calculating time-lagged cross-correlations and coherences of oxy-Hb signals between channels, we tested the functional connectivity for the three periods. The oxy-Hb signals in neighboring channels, as well as their homologous channels in the contralateral hemisphere, showed high correlation coefficients in period 1. Similar correlations were observed in period 2; however, the number of channels showing high correlations was higher in the ipsilateral hemisphere, especially in the anterior–posterior direction. The functional connectivity in period 3 showed a close relationship between the frontal and temporal regions, which was less prominent in period 1, indicating that these regions form the functional networks and work as a hysteresis system that has memory of the previous inputs. We propose a hypothesis that the spatiotemporally large-scale brain networks, including the frontal and temporal regions, underlie speech processing in infants and they might play important roles in language acquisition during infancy.

Keywords: coherence, development, functional connectivity, infant, neuroimaging, NIRS, speech perception, temporal correlation

INTRODUCTION

When we listen to speech sounds of our native language, we automatically convert a train of phonological/prosodic information into meaning in our brains. We know that multiple steps lie behind this process. Although, in most cases, we do not consciously notice every step, the first step is the detection and analysis of acoustic information (Price, 2010). The acoustic information is processed in the bilateral temporal regions, and the role of each hemisphere is considered to be different. One model proposes that the left and right hemispheres are related to the processing of temporal and spectral information, respectively (Zatorre et al., 2002; Zatorre and Gandour, 2008). Another model suggests that both hemispheres have a short window in temporal integration (about 20–80 ms) and that the right hemisphere has a relatively long window (about 150–300 ms; Poeppel et al., 2008). The two models support the many studies that showed bilateral activation in speech and music perception and that reported hemispheric differences between the processes. Infant studies also reported bilateral activation in the temporal regions (Kotilahti et al., 2005) and lateralized activation (Telkemeyer et al., 2009), in which the right hemisphere showed

responses to slow acoustic modulations. If, however, we assume the time window to be shorter than 300 ms, the left-lateralization in brain activation during speech processing would not be fully explained, because phonological information and word-level/sentence-level prosody are converged during word recognition and sentence comprehension over a longer time scale. The assumption of a longer time window in the brain, beyond the time span of the acoustic processing, will enable us to give a fuller picture of speech processing and sentence comprehension.

The spoken language processing progresses from acoustic analysis to detect words and construct sentences, and further evolves to understand the meaning of the sentences. These succeeding processes require longer time than acoustic and phonological processing, which takes up to several hundred milliseconds. Syntax, which is the core of a language faculty, interplays other cores of phonology and meanings (semantics). This grammatical rule system works to construct the hierarchical and recursive structures unique to the human language (Hauser et al., 2002). The processing of a hierarchically structured sequence has been shown to cause activation in the left inferior frontal region (Brodmann's area, BA; 44/45), as

well as in the left frontal operculum (BA 6/FO) and the temporal regions (Friederici et al., 2006). The tractography using the diffusion tensor imaging (DTI) depicted the connection between the left BA 44/45 and the posterior and middle portion of the superior temporal region via the superior longitudinal fasciculus (Catani et al., 2002; Friederici et al., 2006). It was argued that these regions form a syntactic network, in which the function of BA 44/45 is to support the hierarchical reconstruction of the syntactic structure from the sequential input (Friederici, 2006). This network would need to function in a longer time window than several hundred milliseconds. The critical issue in both developmental and language sciences is whether the frontal and temporal regions of preverbal infants work as a functional network, which could then be considered a candidate for the neural foundation of language acquisition.

The functional architecture of the infant brain is beginning to be examined by resting-state functional connectivity magnetic resonance imaging (fcMRI; Fransson et al., 2007, 2011; Gao et al., 2009; Smyser et al., 2010). These fcMRI studies reported the putative precursors of the default-mode networks and their development. We recently found the development of global cortical networks from neonates to 3- and 6-month-old infants by using multi-channel near-infrared spectroscopy (NIRS; Homae et al., 2010): The temporal, parietal, and occipital regions show increases in homologous connectivity connecting the left and right hemispheres, and the fronto-occipital connectivity show U-shaped changes in the course of development. The resting-state measurements successfully provide information about the organization of infant brains, whereas the functional networks related to perceptual and cognitive processing are not fully examined. Supporting evidence for these networks in infants is revealed by anatomical and functional imaging studies. DTI studies on infants visualized the superior longitudinal fasciculus connecting these regions (Zhang et al., 2007; Dubois et al., 2009). Our previous studies (Homae et al., 2007; Nakano et al., 2009), Gervain et al. (2008), and Imada et al. (2006) reported co-activation in the frontal and temporal regions, suggesting that these regions collaborate to process speech sounds. Changeux and his collaborators proposed a “global neuronal workspace (GNW)” model, in which a long-distance temporofrontal GNW circuit of infants is activated by speech stimuli (Dehaene et al., 1998; Lagercrantz and Changeux, 2009). Based on these observations, we attempt to clarify the state-dependent functional networks in the infant brain specifically involved in speech processing. We expect that such a long-distance connectivity would have a longer time window and show high correlations in a lower frequency domain.

In the present study, we measured the brain activation and functional connectivity of 3-month-old infants by using 94-channel NIRS (Figures 1A,B). We prepared three periods for measurements (Figure 1C). During the first 3 min, we measured spontaneous fluctuation of activity in the brain (period 1). After period 1, we provided stimuli by playing Japanese sentences for 3 min (period 2). Finally, we measured brain activation for 3 min without providing the stimulus (period 3), as for period 1. We evaluated the changes in oxygenated and deoxygenated hemoglobin (oxy-Hb and deoxy-Hb) signals for each measurement channel. In addition, we mapped the functional connectivity in each period by calculating the time-lagged cross-correlations of oxy-Hb, deoxy-Hb, and total-Hb (summation of oxy-Hb and deoxy-Hb) signals between

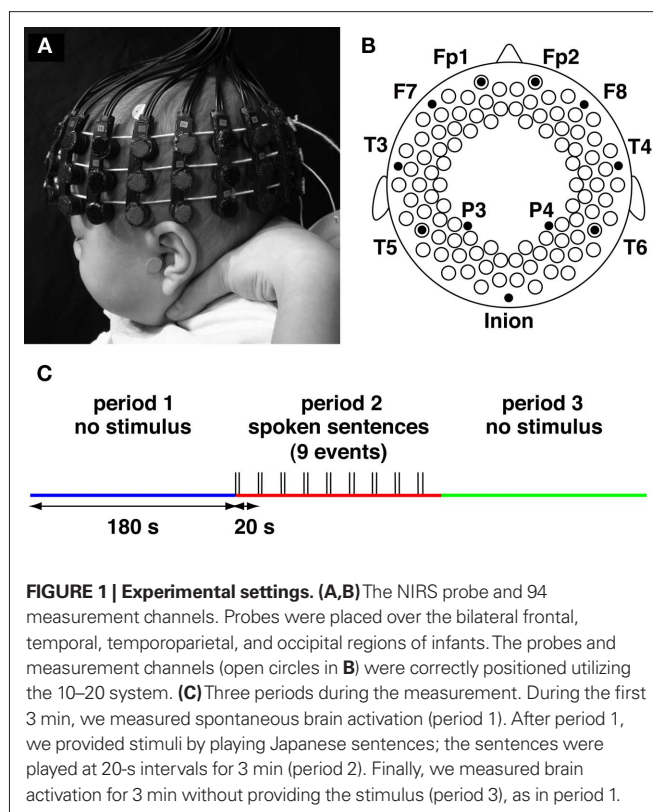


FIGURE 1 | Experimental settings. (A,B) The NIRS probe and 94 measurement channels. Probes were placed over the bilateral frontal, temporal, temporoparietal, and occipital regions of infants. The probes and measurement channels (open circles in B) were correctly positioned utilizing the 10–20 system. (C) Three periods during the measurement. During the first 3 min, we measured spontaneous brain activation (period 1). After period 1, we provided stimuli by playing Japanese sentences; the sentences were played at 20-s intervals for 3 min (period 2). Finally, we measured brain activation for 3 min without providing the stimulus (period 3), as in period 1.

channels. Our primary concern was to reveal the functional networks that are activated during the presentation of speech sounds. We further examined whether the functional networks show context-dependent changes by contrasting the networks under no-stimulus conditions both before and after the presentation of speech sounds.

MATERIALS AND METHODS

PARTICIPANTS

Twenty-one 3-month-old infants participated in the present study (11 girls and 10 boys; mean age: 111.6 days; range: 104–123 days). All infants were full-term healthy Japanese. They were sleeping quietly while they were studied. An additional 56 infants were studied, but they were excluded from the analysis due to either producing large head movements resulting in motion artifacts in the signals ($N = 13$) or probe obstruction by hair ($N = 7$). The measurement was stopped when infants awoke from their sleep during the experiments ($N = 36$). A success rate of sleeping infants was 51.2% (21/41). We checked the number of times the infants moved their heads, bodies, arms, and legs even if they were asleep. The number of movements by each infant at all periods (3 min each) was less than 2 (0: $N = 10$, 1: $N = 7$, and 2: $N = 4$). The total number of movements was 15 ($1 \times 7 + 2 \times 4 = 15$), and the occurrences of movements were equally distributed (five times during each period). The motion during period 2 was not concentrated to the first trial; we observed movements during the 1st, 2nd, 3rd, 5th, and 7th trials within the nine trials of period 2. This behavioral analysis suggests that the state of arousal in the infant group did not change during the three periods. The

measurement was stopped when infants awoke from their sleep during the experiments. Informed consent was obtained from the parents of the infants prior to the initiation of the experiments. The study was approved by the ethics committee of the Graduate School of Education, University of Tokyo.

STIMULI

The speech stimuli used in the present study were a subset of normal speech sounds described previously (Homae et al., 2006, 2007). The stimuli consisted of nine Japanese sentences recorded by a female Japanese speaker (16 bit, 22050 Hz). The mean duration of the sentences was 4.0 s.

STIMULUS PRESENTATION

All experiments were conducted in a sound-attenuated room (the background noise: less than 30 dB SPL). The infant was held in an experimenter's arms during the measurement of cortical activation. The infants were almost motionless and slept soundly throughout the experimental sessions. We previously reported that NIRS recordings from infants in daytime sleep provide long-duration and motion-free data with a sufficiently high signal-to-noise ratio, and the obtained data can be used to evaluate cortical responses to sounds (Homae et al., 2006, 2007, in press; Taga et al., 2007; Nakano et al., 2008, 2009). Stimulus-dependent hemodynamic responses in the brain to speech sounds have been reported in fMRI studies on sleeping adults and infants (Portas et al., 2000; Dehaene-Lambertz et al., 2002), and in NIRS studies on sleeping neonates and infants (Peña et al., 2003; Homae et al., 2006, 2007).

During the first 3 min, we measured spontaneous fluctuation of activity in the brain (period 1, **Figure 1C**). The fluctuation of brain activation and the functional connectivity have been reported previously using this data set (Homae et al., 2010). In period 2, we presented nine different sentences, each of which lasted 4 s followed by 16 s of silence. Speech sounds were presented at a maximum amplitude of 65 dB SPL using a BOSE MMS-1 speaker system placed in front of the infant. During the inter-stimulus interval, no sound was presented. Finally, we measured the brain activation for 3 min without providing the stimulus (period 3), as for period 1.

NIRS RECORDINGS

We used multi-channel NIRS instrument (ETG-7000, Hitachi Medical Corporation, Tokyo, Japan). The NIRS instrument exploits the optical properties of hemoglobin, which has oxygenated and deoxygenated forms with different absorption spectra in the near-infrared (NIR) wavelength region. By using two NIR wavelengths (785 and 830 nm in ETG-7000) and applying the data analyses based on the modified Lambert–Beer law, these instruments measure the relative changes in the concentrations of oxy-Hb and deoxy-Hb in the cerebral cortex at preset measurement points. Detailed descriptions of the principles underlying NIRS have been previously described (Jöbsis, 1977; Reynolds et al., 1988; Maki et al., 1995; Villringer and Chance, 1997; Obrig and Villringer, 2003; Hoshi, 2007; Wolf et al., 2007; Minagawa-Kawai et al., 2008; Lloyd-Fox et al., 2010; Gervain et al., 2011). NIRS has been successfully used to investigate cortical activation in infants in response to auditory stimuli (Peña et al., 2003; Kotilahti et al., 2005, 2010;

Homae et al., 2006, 2007, in press; Minagawa-Kawai et al., 2007; Taga and Asakawa, 2007; Taga et al., 2007; Nakano et al., 2008, 2009; Telkemeyer et al., 2009).

Near-infrared light was emitted from laser diodes through incident optical fibers. The maximum intensity of NIR light was set at 0.6 mW. The received light was detected by avalanche photodiodes through detection via optical fibers and separated into individual light sources, depending on each wavelength. We used two sets of 3×10 arrays composed of 15 incident and 15 detection fibers, which were mounted on a flexible cap over the frontal, temporal, temporo-parietal, and occipital areas of each hemisphere (**Figures 1A,B**). Each pair of adjacent incident and detection fibers defined a single measurement channel, which enabled us to simultaneously measure the time course of oxy-Hb and deoxy-Hb signals with a 0.1-s time resolution. The distance between incident and detection fibers was set at approximately 2 cm (Taga et al., 2007). The measurement channels were correctly positioned by reference to the international 10–20 system of electrode placement using landmarks of external auditory pores, vertex, andinion from each infant. Because few available atlases exist for the infant brain, we used previous studies on adults (Homan et al., 1987; Steinmetz et al., 1989; Herwig et al., 2003; Okamoto et al., 2004) to estimate the craniocerebral correlation for each measurement channel. A recent MRI study suggested that the cortical structure in infants is similar to adults in many aspects (Hill et al., 2010). We have reported functional mapping for audio–visual stimuli in infants, which is consistent with an estimated map from the craniocerebral correlation (Watanabe et al., 2008, 2010).

DATA ANALYSIS

We examined the variation in the oxy-Hb signals, which estimated changes in the regional cerebral blood oxygenation during brain activation. In addition, we also analyzed the deoxy-Hb signals. We evaluated relative changes in oxy-Hb and deoxy-Hb signals contingent on an arbitrarily assigned 0 baseline from the start of the measurement period, which was based on the modified Lambert–Beer law. Because the precise optical path length of the light traveling through brain tissue cannot be evaluated by continuous-wave NIRS, the units of oxy-Hb and deoxy-Hb signals were determined by multiplying the molar concentration by length (mM·mm).

For each individual data set, we used a band-pass filter from 0.009 to 0.08 Hz, which has been used in previous studies on adult participants to eliminate cardiac and respiratory rhythms (Fox et al., 2005; White et al., 2009), and extracted 3-min data (i.e., 1,800 time points) from the continuous time courses (periods 1, 2, and 3). The band-pass filter eliminated cardiac pulsation (about 2 Hz in infants) and smoothed signal drifts over long time scales and motion artifacts.

We initially extracted data blocks from the time course data of period 2. Each data block ranged from 0.5 s prior to stimulus onset to 19.5 s after stimulus onset. By detecting rapid changes in the summation of oxy-Hb and deoxy-Hb signals before applying the band-pass filter, we determined data blocks to be eliminated with a low signal-to-noise ratio due to obstruction by hair and those with movement artifacts. We then calculated the mean signal of the first 11 time points (i.e., from 0.5 s prior to stimulus onset, dur-

ing which no sounds was presented, to 0.5 s after stimulus onset) in each data block and used this value as the baseline for each block. By averaging the signal changes over data blocks for each subject in period 2, we obtained the hemodynamic responses at each measurement channel. To identify the activated regions in the period 2, we evaluated the individual data as random effects and performed a *t*-test, which was performed against 0 baseline (one-sample, two-tailed *t*-test), for each channel. Multiple comparisons among the measurement channels were considered by adopting an all-measurement-channels false discovery rate (FDR) correction at $Q < 0.05$ (Benjamini and Hochberg, 1995; Genovese et al., 2002; Singh and Dan, 2006).

For each infant data of oxy-Hb, deoxy-Hb, and total-Hb (summation of oxy-Hb and deoxy-Hb) signals, we calculated the time-lagged correlation coefficients (*r*) between the time course of a single channel and the time course from all other measurement channels (number of pairs: $(94 \times 93)/2 = 4,371$). The 3-min data of two channels determined a single *r* value. We settled 20-s time window (± 10 s) for lags, and calculated correlation sequences over the lag range (from -10 to 10 s). Because sampling rate was 10 Hz, we obtained 201 *r* values for each pair as a sequence. We adopted the maximum *r* value among the 201 values as the *r* value for the pair. In most cases, we used *r* values within 2-s lags in the present analysis (mean values of lags in oxy-Hb signals: -0.034 , -0.150 , and 0.124 s in periods 1, 2, and 3, respectively; mode values of lags in oxy-Hb signals: 0.0 s in all periods). We used all the filtered data to estimate the cross-correlations. Because we used correlation coefficients in our analyses, the different optical path lengths of the measurement channels did not affect our results. We considered both positive and negative *r* values (range: $-1 \leq r \leq 1$) and evaluated all the *r* values. To reveal the changes in connectivity between the periods, we analyzed the differences on a channel pair basis. We first converted the *r* values to *z* scores by Fischer's *z* transformation. We evaluated the individual *z* scores as random effects and performed paired *t*-tests (statistical threshold: $p < 0.005$). We applied two types of the comparison: (1) period 1 and period 2 and (2) period 1 and period 3. We recently found that the patterns of frequency-specific functional connectivity are different between oxy-Hb and deoxy-Hb signals (Sasai et al., 2011). Further, Mesquita et al. (2010) calculated resting-state functional connectivity using oxy-Hb, deoxy-Hb, and total-Hb signals. They reported a trend toward a higher correlation between homologous regions in total-Hb signals in comparison to oxy-Hb and deoxy-Hb signals; however, one cannot exclude qualitative and quantitative differences among the signals. Because oxy-Hb signals displayed a better signal-to-noise ratio than deoxy-Hb signals in our previous study (Homae et al., 2007), we focused on oxy-Hb signal changes to make direct comparisons between periods.

We further calculated the squared coherence of oxy-Hb signals between channel pairs. In this analysis, we used raw data without any filters. We applied Welch's averaged, modified periodogram method (using a 1024 point Fourier transform, Hanning window, and overlap of 512 points) to estimate the cross spectral density and the power spectral density. The magnitudes of squared coherence were calculated from them. The values for each frequency were averaged in each infant.

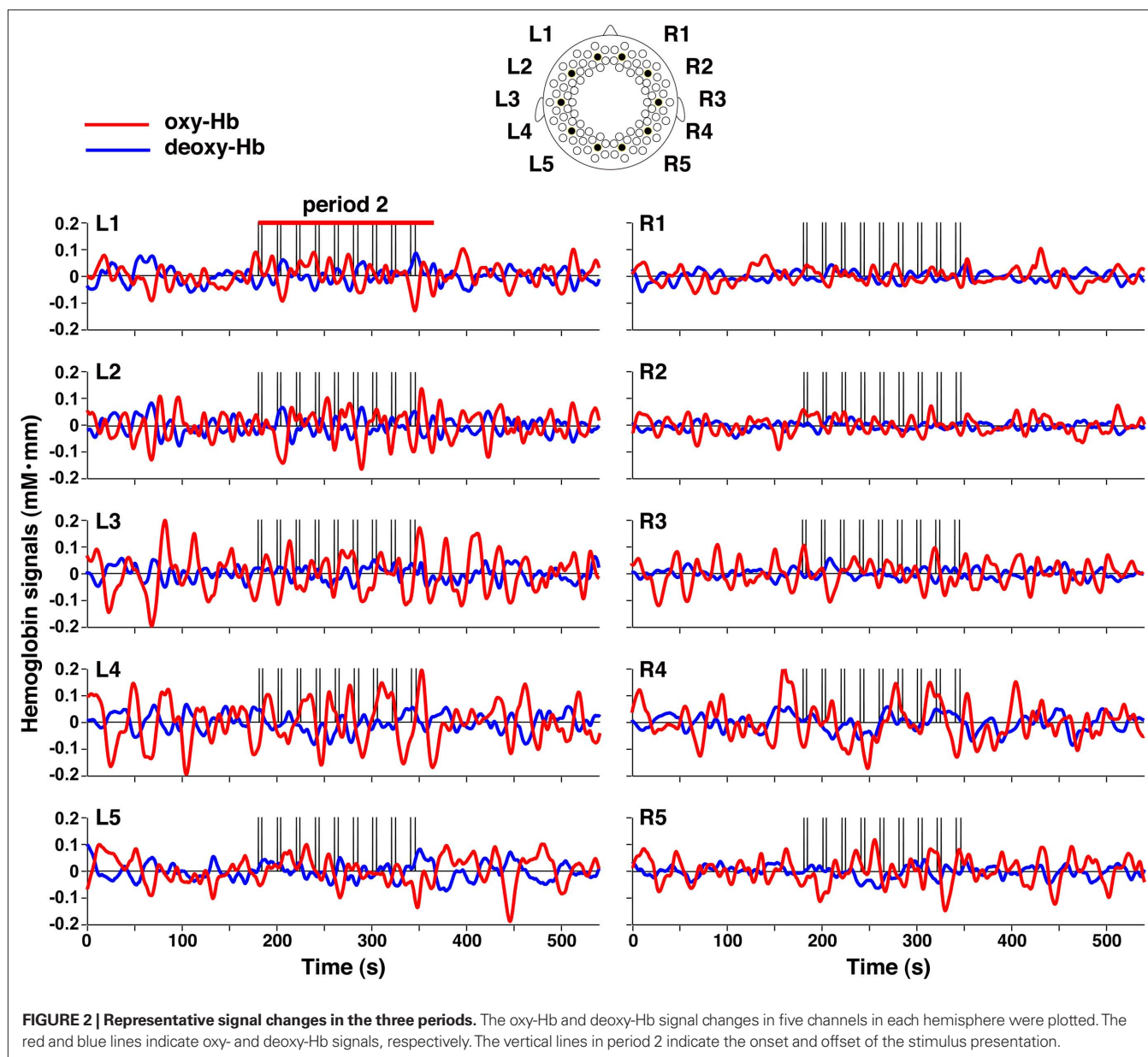
In the cluster analyses for oxy-Hb, deoxy-Hb, and total-Hb signals, we defined the distance between the two channels by calculating $1 - r$. We applied the Ward method for determining the distance and constructed a dendrogram. We showed four clusters for each period.

RESULTS

We found that 3-month-old infants showed spontaneous brain activity and hemodynamic responses to speech sounds in both the left and right hemispheres (**Figure 2**). The temporal, temporoparietal, and occipital regions showed remarkable signal changes from the baseline during period 2 (around the onset of the stimulus presentation, see Materials and Methods). These regions showed increases in oxy-Hb signals as well as decreases in deoxy-Hb signals. Moreover, the oxy-Hb signal changes were larger than the deoxy-Hb signal changes. To quantify the hemodynamic responses to speech sounds, we created an average time course of all the measurement channels in period 2 (**Figure 3**). The maximum change in the oxy-Hb and deoxy-Hb signals occurred at 8.8 and 10.6 s, respectively. For the statistical analyses using *t*-tests, we used the mean changes of the oxy-Hb signals in the time window from 6.9 to 10.7 s, and the deoxy-Hb signals in the time window from 8.9 to 12.2 s after the onset of the stimulus presentation, in which the signal changes were greater than 2 SDs for all time points of the averaged time course of oxy-Hb and deoxy-Hb signals.

BILATERAL ACTIVATION IN RESPONSE TO SPEECH SOUNDS

By conducting statistical analyses of the mean changes in oxy-Hb signals, we examined cortical activation in response to the speech sounds (period 2). We found that the bilateral prefrontal, temporal, and temporoparietal regions of the 3-month-old infants showed significant activation (**Figure 4**). The activation patterns in the bilateral regions of infants were consistent with our previous results using speech sounds (Homae et al., 2006; Taga and Asakawa, 2007; Nakano et al., 2008, 2009). We further found significant signal changes in the bilateral occipital regions. These bilateral changes were also observed in deoxy-Hb signals. The increase in oxy-Hb signals and the decrease in deoxy-Hb signals in the occipital regions indicated that these regions of 3-month-old infants were activated when speech sounds were presented to the infants. We calculated the time of maximal changes in oxy-Hb and deoxy-Hb signals. The timings in the oxy-Hb signals in the temporal channels of the left and right hemispheres were 8.5 and 8.8 s, respectively (**Figure 5A**). The timings in the oxy-Hb signals in the occipital channels of the left and right hemispheres were 9.4 and 9.1 s, respectively (**Figure 5B**). The timing of deoxy-Hb signal in the temporal channels of both the left and right hemispheres was 8.5 s (**Figure 5C**), whereas in the occipital channels of the left and right hemispheres, the timings were 10.8 and 10.9 s, respectively (**Figure 5D**). The rates of increases or decreases during the first 5 s were also different between the temporal and occipital regions. The difference between the regions were not specific to the above four channels (see the inset in **Figure 5**). We calculated the mean peak time in oxy-Hb signals in each column from temporal regions to the occipital regions (10 columns in total). The mean timing in the temporal channels (three channels



in each hemisphere, column 1), and in the occipital channels (two channels in each hemisphere, column 10) was 8.47 and 9.20 s, respectively (**Figure 5E**). These analyses demonstrate that the temporal profiles of signal changes were region specific. Based on these findings, we suggest that the activation seen in the extensive cortical regions are not global systemic effects, but reflect region-specific cortical activation.

To test whether the signal changes in the left and right hemispheres showed differences, we applied direct comparisons between the signal changes in homologous channels (47 pairs). These comparisons were based on the assumption that the optical path lengths in homologous regions were equivalent to each other. No regions showed significant differences between the homologous channels. This result demonstrated that both hemispheres were involved in the processing of speech sounds.

FUNCTIONAL CONNECTIVITY IN THE THREE PERIODS

Next, we calculated temporal correlations between all the pairs of measurement channels during each period ($r > 0.5$, **Figures 6 and 7**). In period 1, during which no stimuli were presented, the prominent connectivity caused by the spontaneous brain activation was observed in homologous channels, as reported in our previous study (Homae et al., 2010). When speech sounds were presented (period 2), the number of channel pairs that exhibited high correlations increased throughout the brain. Not only the homologous connectivity, but also fronto-posterior connectivity in the ipsilateral hemisphere appeared. After the stimulus presentation (period 3), we found that the fronto-posterior connectivity in the ipsilateral hemisphere still showed high correlations. This tendency remained even if we calculated correlations using the last 150-s data. In the following analyses,

we focused on two direct comparisons of these correlations: period 1 vs. period 2, and period 1 vs. period 3. The former comparison will clarify cortical networks related to the processing of speech sounds. The latter will reveal the aftereffects of the perceiving of speech sounds. If the processing occurs only in the midst of speech-stimulus presentation, we will obtain null results in the latter comparison; otherwise, the functional networks will show state-dependent changes.

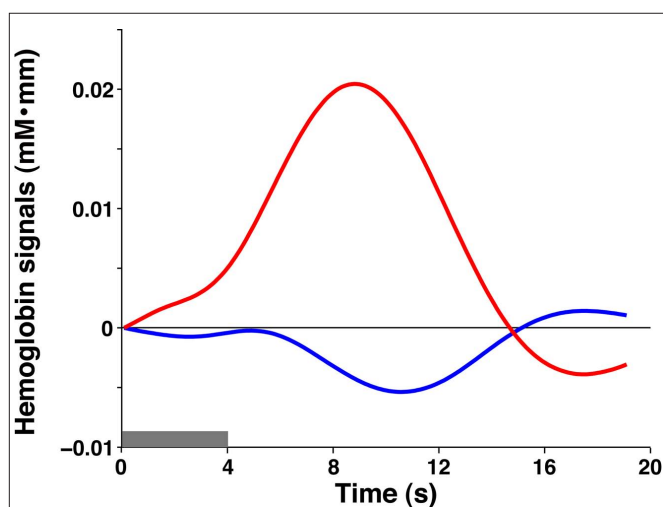


FIGURE 3 | The averaged time courses of oxy-Hb and deoxy-Hb signal changes in all the measurement channels during period 2. The gray bar indicates the period during which the speech sounds were presented (mean duration: 4.0 s). The red and blue lines indicate oxy-Hb and deoxy-Hb signals, respectively.

FUNCTIONAL CONNECTIVITY IN PERIOD 2

The direct comparison between period 1 and period 2 revealed the increases in connectivity in period 2 (**Figure 8A**). Both the homologous connectivity and fronto-posterior connectivity in period 2 was higher than those in period 1. The higher fronto-posterior connectivity in the ipsilateral hemisphere was found in the left and the right hemispheres (**Figure 8B**). The most prominent changes were observed in prefrontal and posterior temporal/occipital connectivity in the right hemisphere ($p = 0.0001$, a bold line in **Figure 8B**), and the left homologous connectivity also showed large changes (**Figure 8C**). When we ceased to present speech sounds (period 3), correlations between these regions returned to the level of period 1, suggesting that these networks functioned during the processing of speech sounds.

In addition, hemispheric differences were observed in this comparison. For one example, correlations between the anterior regions of the prefrontal cortex and the occipital regions in the left hemisphere were higher than those in the right hemisphere. Another example is that the right temporoparietal region, which is related to the processing of pitch information (Homae et al., 2006, 2007, in press), showed an increase in connectivity with the right prefrontal region (thick circles in **Figure 8B**), whereas the left homologous pair did not exhibit such changes ($p > 0.1$). These differences were not clarified when the intensity of signal changes was tested. The analyses of functional connectivity provided novel information about the processing of speech sounds in the infant brain.

FUNCTIONAL CONNECTIVITY IN PERIOD 3

Both periods 1 and 3 were no-stimulus conditions, but the context of period 3 differed from that of period 1. If the presentation of speech sounds affects the brain networks for a long time, functional connectivity in period 3 would be expected to differ from that of period 1. The direct comparison between the two periods

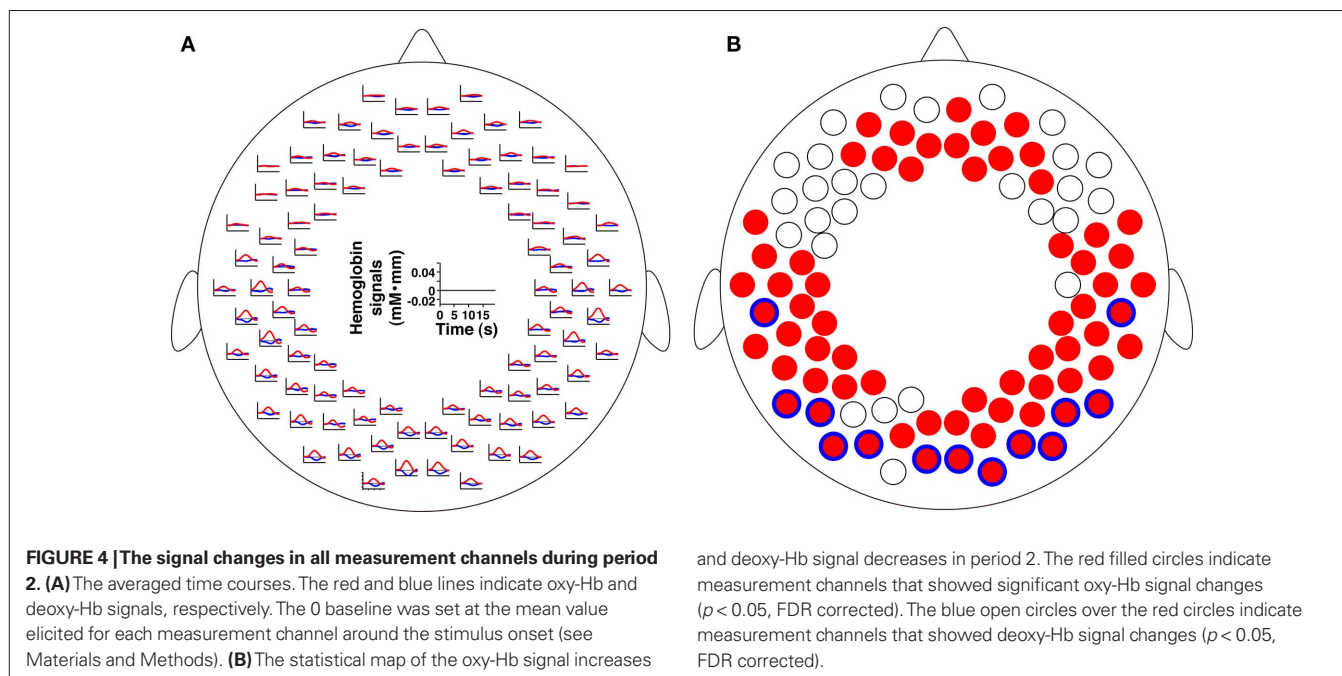


FIGURE 4 | The signal changes in all measurement channels during period 2. **(A)** The averaged time courses. The red and blue lines indicate oxy-Hb and deoxy-Hb signals, respectively. The 0 baseline was set at the mean value elicited for each measurement channel around the stimulus onset (see Materials and Methods). **(B)** The statistical map of the oxy-Hb signal increases

and deoxy-Hb signal decreases in period 2. The red filled circles indicate measurement channels that showed significant oxy-Hb signal changes ($p < 0.05$, FDR corrected). The blue open circles over the red circles indicate measurement channels that showed deoxy-Hb signal changes ($p < 0.05$, FDR corrected).

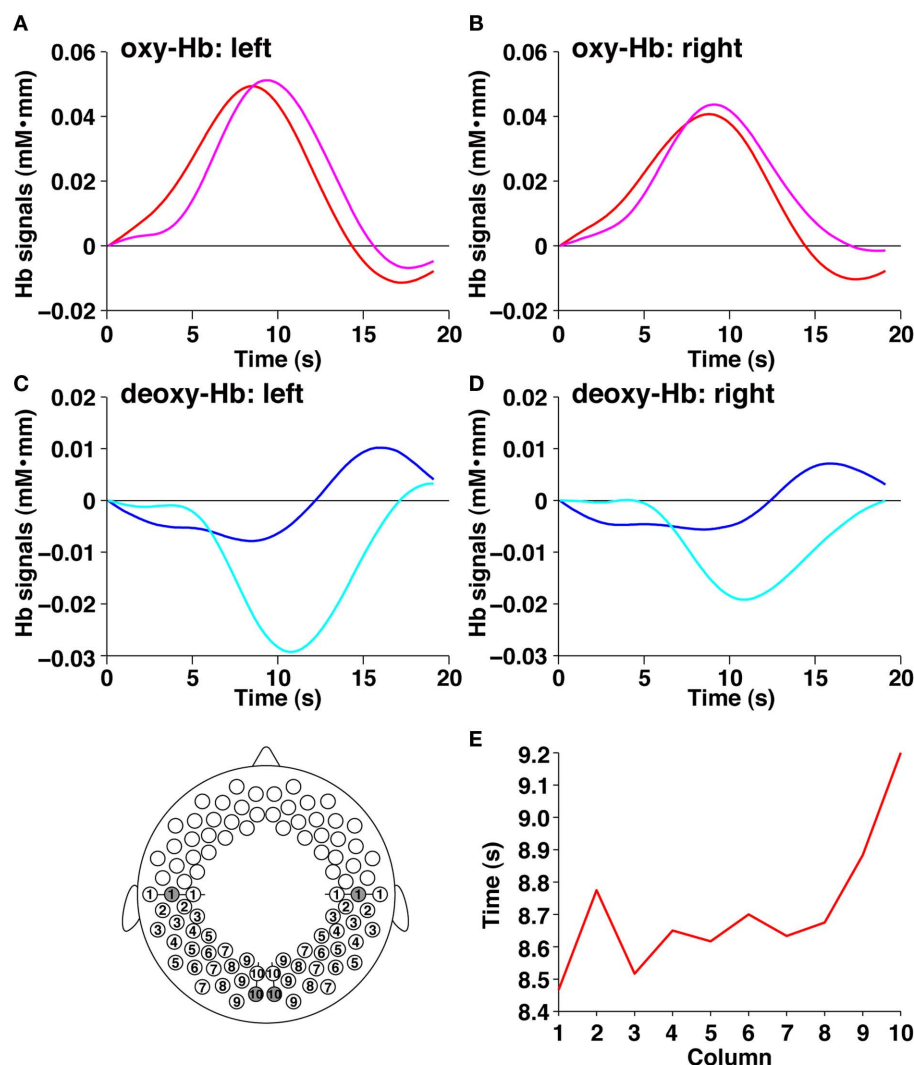


FIGURE 5 | The signal changes in the temporal and occipital channels during period 2. (A) The average time course of oxy-Hb signals in the left hemisphere. The red and magenta lines indicate the temporal and occipital channels, respectively (filled circles in the bottom panel). **(B)** The average time course of oxy-Hb signals in the right hemisphere. The red and magenta lines indicate the temporal and occipital channels, respectively. **(C)** The average time course of deoxy-Hb signals in the left hemisphere. The blue and cyan lines indicate the temporal and occipital channels, respectively (filled circles in the bottom panel). **(D)** The average time course of deoxy-Hb signals in the right

hemisphere. The blue and cyan lines indicate the temporal and occipital channels, respectively. **(E)** The average peak time of the oxy-Hb signals in 10 columns from temporal to occipital columns. The abscissa indicates the columns shown in the bottom panel. The average time courses were calculated from oxy-Hb signals of 6 channels (columns 1, 3, 5, 7, and 9 in the left and right hemispheres; three channels in each hemisphere) or four channels (columns 2, 4, 6, 8, and 10 in the left and right hemispheres; two channels in each hemisphere), and the peak time in each time course was investigated. The ordinate shows the peak time from the onset of the stimulus presentation.

supported this possibility. We found that the fronto-posterior connectivity between the prefrontal regions and the temporal/temporoparietal regions in period 3 are higher than those in period 1 (**Figure 9A**). We used all four pairs in each hemisphere, which connect the frontal and posterior regions (bold lines in **Figure 9A**), to calculate mean z values for each subject (**Figure 9B**). The mean r values of the four pairs in period 3 were 0.572, 0.557, 0.554, and 0.523 in the left hemisphere and 0.472, 0.462, 0.458, and 0.447 in the right hemisphere. We confirmed that the differences between period 1 and period 3 were significant (both hemispheres:

$p < 0.0005$), indicating that the presentation of speech sounds affected cortical activity after the presentation period, especially in the long-distance brain networks.

Interestingly, the four pairs in the left hemisphere showed increases from period 2 to period 3 (sign test, $p < 0.05$), while such a difference was not observed in the right hemisphere ($p > 0.9$). This trend was remarkable in the low-frequency band in the coherence analysis (**Figure 9C**). The coherence in lesser than 0.04 Hz of the left hemisphere was highest during period 3, while one in the right hemisphere did not show such a difference during periods 2 and 3.

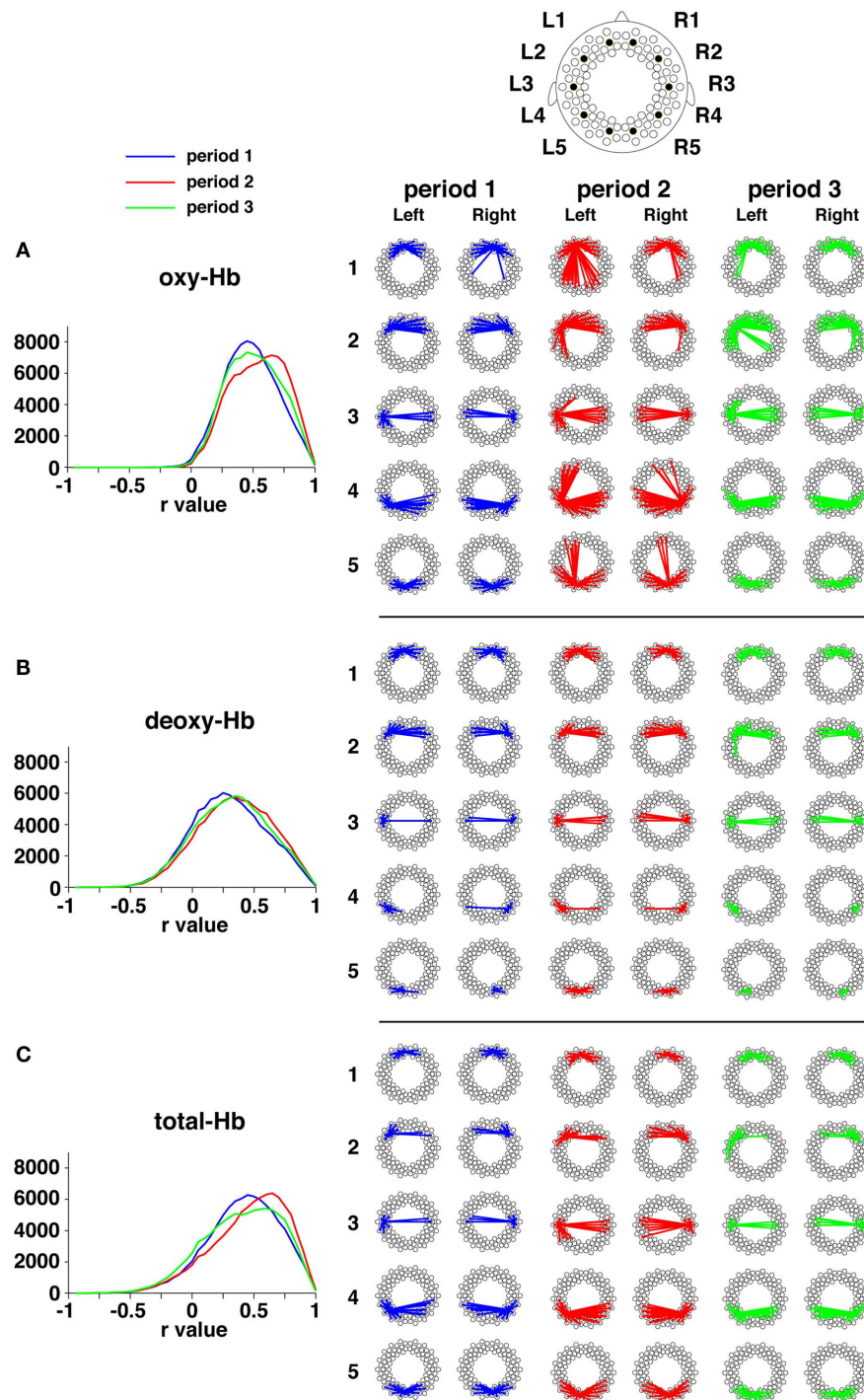


FIGURE 6 | Functional networks based on oxy-Hb signals, deoxy-Hb signals, and total-Hb signals. (A) Temporal correlations in oxy-Hb signals across the three periods. The distributions of correlation coefficients (*r* values) are shown by the three lines. The ordinates denote the total number of channel pairs across the infants. The representative correlations

during each period are exhibited. The five channels in each hemisphere are the same as those shown in **Figure 2**. The lines show correlations that are higher than 0.5 (averaged across all infants in each group). **(B)** Temporal correlations in deoxy-Hb signals. **(C)** Temporal correlations in total-Hb signals.

We conducted a two-by-two ANOVA (hemisphere, left and right; period, 2 and 3) using the values between 0.009 – 0.04 Hz, and found that the interaction was significant ($F(1,20) = 4.58, p < 0.05$).

The simple main effect showed that the coherence of period 3 in the left hemisphere was higher than that in the right hemisphere ($p < 0.01$). These results indicate that the increase in correlations

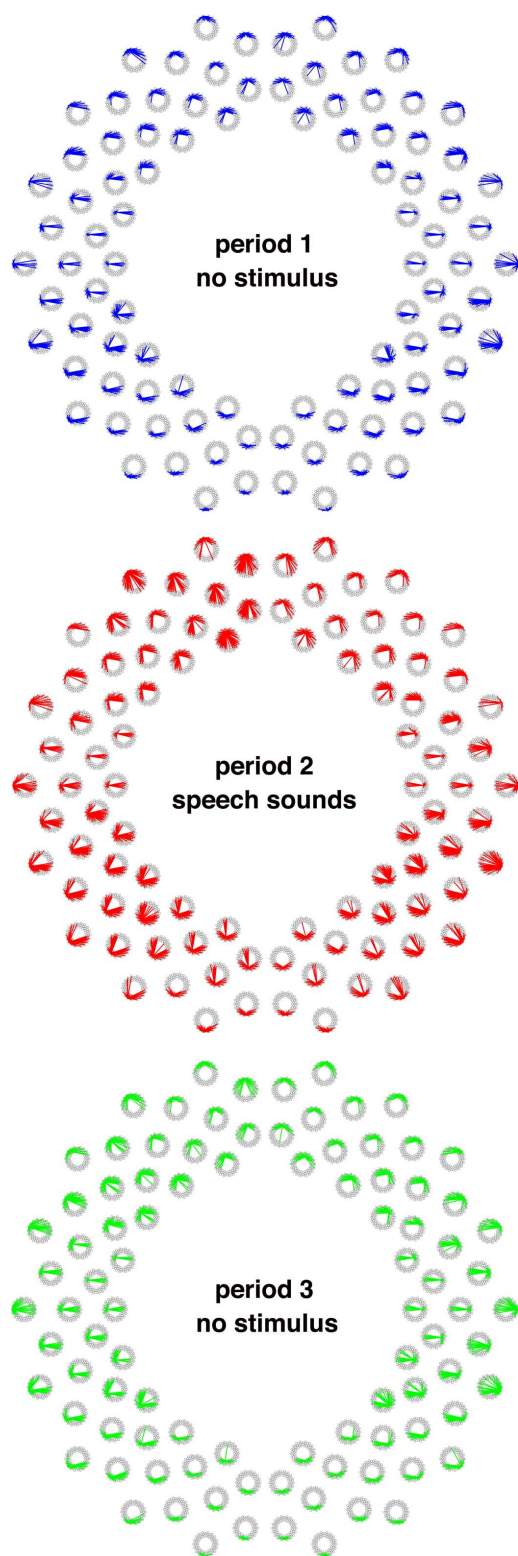


FIGURE 7 | All the connectivity in the three periods. The lines show correlations that were higher than 0.5 (averaged across all infants). The correlations between one of the 94 measurement channels and other channels are shown according to the arrangement of channel settings (Figure 1B).

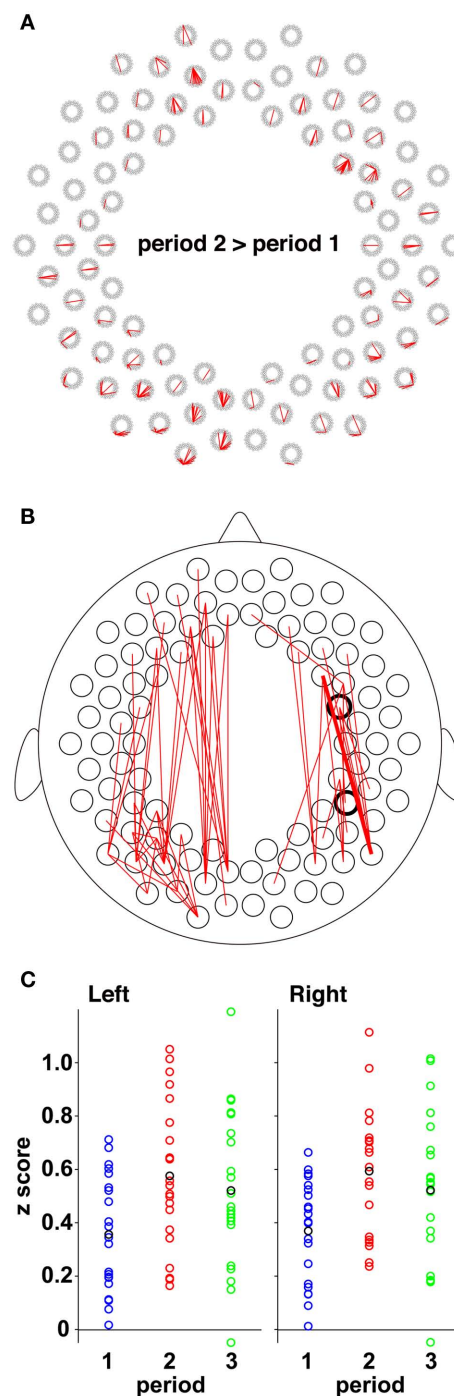


FIGURE 8 | The direct comparison between period 1 and period 2. (A) All the connectivity that showed higher correlations in period 2 in comparison with period 1 ($p < 0.005$). (B) The ipsilateral connectivity revealed by the direct comparison ($p < 0.005$). The bold line indicates the connectivity that showed the largest difference. The thick circles show the measurement channels of a right temporoparietal region and a right prefrontal region. (C) Individual data of temporal correlations normalized to z scores. The scatter plots show z scores in the pair indicated by the bold line in (B) and its left homologous pair. The blue, red, and green circles indicate individual data of z scores during periods 1, 2, and 3, respectively. The black circles indicate mean values in each period.

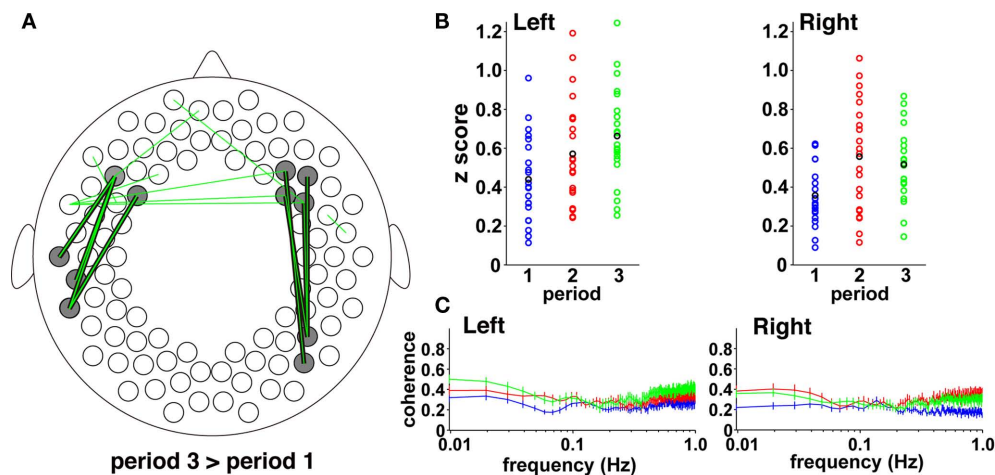


FIGURE 9 | The direct comparison between period 3 and period 1. (A) All the connectivity that showed higher correlations in period 3 in comparison with period 1 ($p < 0.005$). All the connectivity results were overlaid in a single figure. The bold lines indicate the connectivity that connects the frontal and temporal channels (four pairs in each hemisphere). **(B)** Individual data of temporal correlations normalized to z scores. The scatter plots show mean z scores in the four pairs indicated by the bold lines in each hemisphere in **(A)**. The blue, red, and green circles indicate individual data of z scores during periods 1, 2, and 3,

respectively. The black circles indicate mean values in each period. **(C)** The magnitudes of squared coherence between the frontal and temporal channels. We calculated the squared coherence of oxy-Hb signals between the four pairs in each hemisphere, shown in **(A)**. The four values for each frequency were averaged for each infant. The blue, red, and green lines indicate the mean values in periods 1, 2, and 3, respectively (error bar: the SE). Data below 1 Hz were presented in each period. The abscissa and ordinate show the frequency in a log scale and the magnitude of squared coherence, respectively.

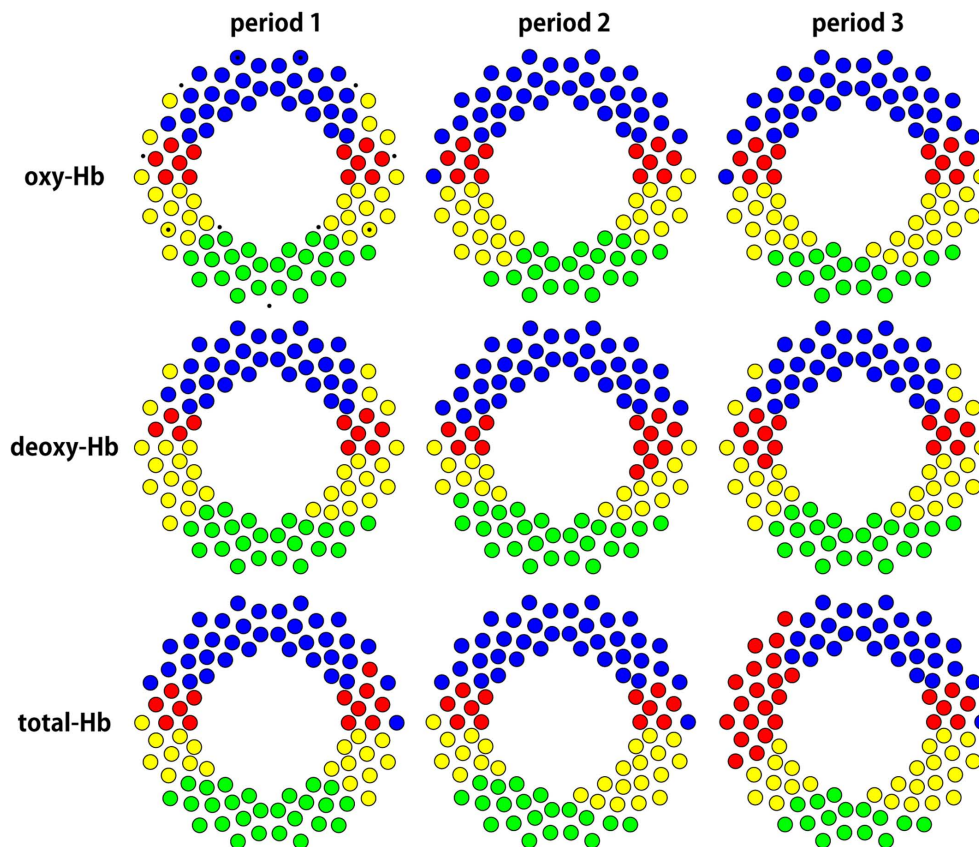


FIGURE 10 | Spatial configuration of the 4 clusters obtained by cluster analyses. We defined the distance between two channels by calculating $1 - r$, and conducted cluster analyses. The four clusters in each period can be identified by coloring. The small dots are the landmarks shown in **Figure 1B**.

from period 2 to 3 was observed in the lower frequency of activation in the left hemisphere. Further, the coherence analysis suggests that the highest correlation of the left hemisphere during period 3 does not depend on signal fluctuations immediately after the end point of the speech presentation (period 2), but reflects activation of the fronto-temporal network in the left hemisphere during period 3.

THE CLUSTER ANALYSES

To reveal the spatial formation of regions that had similar temporal correlations, we applied cluster analyses to the correlation coefficients calculated from oxy-Hb, deoxy-Hb, and total-Hb signals (**Figure 10**). In all periods, the four clusters showed bilateral patterns, in which clusters were formed across the midline and included homologous channels. The bilateral frontal, temporal, temporoparietal, and parieto-occipital regions formed clusters. The clusters above T3 and T4 (red cluster) based on oxy-Hb signals showed notable consistency in terms of spatial configuration in both the hemispheres across all three periods. However, when the cluster analysis was done based on total-Hb signals, the red cluster in the left hemisphere during period 3 extended to frontal and posterior temporal regions. The left–right asymmetry in the fronto-temporal direction was consistent with the result in the analyses above (**Figure 9**). Our findings are consistent with the idea that long-distance functional networks show lateralization in the infant brain.

DISCUSSION

The present study demonstrated that extensive cortical regions of 3-month-old infants formed large-scale brain networks that are related to the processing of speech sounds. We found that not only the bilateral temporal and temporoparietal regions, but also the prefrontal and occipital regions showed oxy-Hb signal increases and deoxy-Hb signal decreases when speech sounds were presented to the infants. While the bilateral regions were responsive to speech sounds, a hemispheric difference was marked in the fronto-temporal networks. Our correlation analyses revealed long-distance functional connectivity in the antero-posterior direction during and after the presentation of spoken sentences, suggesting that the large-scale brain networks of infants provide the foundation for speech processing.

THE LEFT AND RIGHT HEMISPHERES

The hemispheric difference in speech perception in infancy is one of the greatest questions in the developmental cognitive neurosciences. An NIRS study on neonates was reported by Kotilahti et al. (2005), in which the bilateral temporal regions were responsive to sinusoidal tones. This bilateral activation in response to auditory stimuli was reproduced by Telkemeyer et al. (2009). They showed not only bilateral activation, but also right lateralization in response to slow acoustic modulations. These studies suggest that speech sounds, including both rapid and slow changes, induce bilateral activation in the temporal regions. This suggestion has been supported by multiple NIRS studies of infants (Homae et al., 2006, 2007; Minagawa-Kawai et al., 2007; Taga and Asakawa, 2007; Nakano et al., 2008, 2009; Bortfeld et al., 2009; Kotilahti et al., 2010; Sato et al., 2010). It is informative to investigate activation in the bilateral auditory regions, when speech sounds and other sounds are presented, before testing hemispheric differences in signal intensity. We consider that if the bilateral activation in auditory regions is

not observed under conditions of auditory-stimulus presentation, we have to carefully examine the auditory/speech stimuli, tasks, and probe settings in NIRS measurements of the study.

In the present study, we found that the bilateral cortical regions were responsive to speech sounds, and there were no significant differences in the amount of signal changes between homologous regions. We presented only normal speech in period 2, and thus both segmental and suprasegmental information is included in all the stimuli. In such a case, infants would perceive and process both rapid and slow changes in acoustic/speech information, resulting in a similar degree of activation in the left and right hemispheres. A lateralized activation pattern would emerge if some features in speech sounds were highlighted by the presentation of multiple types of speech stimuli. In our previous studies, we presented normal and flattened speech sounds to infants and found the right-lateralized activation pattern in the temporoparietal region related to processing pitch information in speech sounds (Homae et al., 2006, 2007). In the present study, we observed significant activation in the temporoparietal regions in both hemispheres, and an increase in correlations between the right temporoparietal region and the right prefrontal region in period 2. The connectivity was right-lateralized and the increase was not observed in the left homologue. These findings and our previous studies suggest that the prosodic information in normal speech would be processed in the right-lateralized network.

ACTIVATION AND CONNECTIVITY IN THE EXTENSIVE CORTICAL REGIONS

The speech sounds we presented in period 2 induced bilateral activation in extensive cortical regions posterior to the temporal regions, including the occipital regions. These distributed activation patterns in sleeping infants have been observed in our previous studies, but the present result is the first demonstration of oxy-Hb increases and deoxy-Hb decreases in the occipital regions. The oxy-Hb and deoxy-Hb signal changes indicated the neural responses, in these regions, to the speech sounds. There are at least three possibilities that explain this broad activation. First, corticocortical connections convey the information from the temporal region to the posterior regions. In the perinatal period, the temporal lobe expands and the auditory cortex develops. Mature axons in the auditory cortex, which run parallel to the cortical surface for long distances, are present in the most superficial layer (Moore, 2002). These axons could be related to cortico-cortical connections. Second, direct connections between the auditory and visual cortices subserve the information transmission. Although the direct pathway has not been clarified in infants, neurophysiological and neuroimaging studies on animals and human adults provide multiple evidence (Bavelier and Neville, 2002; Falchier et al., 2002; Rockland and Ojima, 2003; Eckert et al., 2008). An fMRI study on adults reported that the visual cortices showed responses to the auditory stimuli and the peak times of the signal changes in the visual cortex were later than those in the auditory cortex (Martuzzi et al., 2007). This trend is observed in our study. If the speech information is conveyed to the auditory regions in the cerebral cortex firstly, the activation in the visual cortex would not precede to the activation in the auditory cortex. Third, the thalamus plays a role in the transmission from the auditory cortices to the visual cortices. Audio-visual interaction via the thalamus has been reported in the early stages of audio-visual processing (Baier et al., 2006; Noesselt et al., 2010). It is possible that activation in the medial geniculate body

induces activation in the lateral geniculate body and the visual cortices in the sleeping infants. In awake infants of 2- to 4-months old, in contrast, activation in the perisylvian regions and deactivation in the occipital region were reported in an fMRI study (Dehaene-Lambertz et al., 2006). Thus, infants' wakefulness would determine the activation pattern in the occipital regions. Excitatory and inhibitory balance in the visual regions of infants might depend on awake/sleep states or stimulus differences. Future studies are needed to clarify the mechanisms of extensive activation in the posterior regions and their dependency on infants' wakefulness.

The correlation between the left occipital regions and the left frontal regions were high in period 2. We have reported that the correlation between these regions in the resting state exhibited U-shaped changes: It was lowest in 3-month-old infants in comparison with neonates and 6-month-old infants (Homae et al., 2010). A recent resting-state fMRI study reported consistent results on neonates. The right dorsolateral prefrontal cortex had functional connectivity with the left homologous region and the bilateral precuneus (Fransson et al., 2011). The present results demonstrate that the left fronto-occipital network in 3-month-old infants could be activated if speech sounds were presented, although physiological confounds, including respiratory, cardiovascular, and blood pressure oscillations, should be carefully considered. One anatomical basis connecting these regions might be the occipitofrontal fascicle, which was traced in a DTI study on adults (Makris et al., 2007). The potential role of this fascicle in adults is visual processing, but its role in early infancy is not yet known. We observed significant activation in the occipital regions and the frontal regions only in period 2 and connectivity between these regions in the period. Because speech sounds were presented in period 2, speech sounds would modulate activity in the occipital and frontal regions and the fronto-occipital connectivity. We propose a possibility that the fronto-occipital network might make some contribution to the perception of speech sounds in sleeping infants.

LARGE-SCALE FUNCTIONAL NETWORKS AND HYSTERESIS IN THE INFANT BRAIN

To date, functional studies on infant brain networks mainly dealt with no-stimulus conditions and reported the precursors of default-mode networks (Fransson et al., 2007; Gao et al., 2009). Recently, we found the development of global cortical networks in the infant brain (Homae et al., 2010): The resting-state connectivity between homologous regions in the temporal and posterior cortices, and the connectivity between the left temporal and parietal regions, are strengthened from neonates to 3- and 6-month-old infants. However, the relationship between the frontal and temporal regions was not evident in the resting state of infants, in contrast to connectivity between the left frontal and temporal regions in the resting state of adults (Zhang et al., 2010). Our previous studies (Homae et al., 2007; Nakano et al., 2009) and those of others (Imada et al., 2006; Gervain et al., 2008) show co-occurrence of activation in the temporal/temporoparietal and frontal regions when speech sounds are presented, leading us to hypothesize that the fronto-temporal networks begin to function during early infancy. The presentation of speech sounds would change the state of networks including these regions. In the present study, we clarified that the connectivity between the prefrontal and posterior portions of the temporal regions was high in both hemispheres under the stimulus presentation condition (period 2). The anatomical connection between the prefrontal region and the

temporal/temporoparietal region has been reported in primate studies (Petrides and Pandya, 1988; Romanski et al., 1999) and in human studies using DTI (Makris et al., 2005; Tomassini et al., 2007). In the infants, a small tract of the superior longitudinal fasciculus and the arcuate fasciculus connecting these regions was identified and reconstructed by tractography of DTI (Zhang et al., 2007; Dubois et al., 2009). These studies suggest that these anatomical connections, at least partly, exist in infancy. Our results are consistent with the hypothesis that the fronto-temporal networks begin to function in infancy, and may play a role in the processing of speech sounds.

Speech sounds affect brain activation of infants even after the presentation has ceased. The comparison between the functional connectivity in a no-stimulus condition before and after the speech-stimulus presentation revealed resultant modulation of speech sounds in the bilateral fronto-temporal connectivity. Pairs in the left hemisphere that showed high correlations in period 3 did not exhibit significantly high correlations in the period 2 (Figures 8A and 9A). This implies that high correlations in period 3 were not remainders of activation in the period 2, but the reflection of co-activation of these regions during period 3. The findings of this study demonstrate that the functional network of the infant brain is state-dependent, and further, it may suggest that the functional network works as a hysteresis system that has a memory of the previous inputs. Hysteresis in the neural mechanisms is, in principle, related to learning and development. A recent fMRI study reported that motor learning by 11-min visuomotor training can modulate subsequent activity within resting-state networks (Albert et al., 2009). The hysteresis system that we found in the infant brain would be involved in the learning of speech perception and language acquisition.

The possibility that the left and right functional networks play distinct roles is suggested by the hemispheric difference in the networks between periods 2 and 3. The left fronto-temporal network showed higher correlations in period 3 in comparison with those in period 2, whereas such a difference was not observed in the right network. The candidate roles are, for example, that the right network is involved in retaining speech information and the left network produces some representation that is relevant to speech information. The coherence analysis revealed that the high correlations in the left networks mostly depended on synchronization in the lower frequency. Our findings allow us to propose the hypothesis that the large-scale (long-distance in spatial and long-range in temporal) connectivity between the frontal and temporal regions underlies the speech processing of infants. The large-scale brain networks would enable the processing of internal and external information hierarchically and recursively, which might constrain language acquisition in infancy. These networks are supposed to be dynamic and not static, and thus, their roles would vary during the course of development. The accumulation of knowledge of the functional networks in the developing infant brain, as well as data on functional brain mapping, opens up a new field – the developmental brain science of language – and will further clarify how and why infants acquire their native languages.

ACKNOWLEDGMENTS

This work was partly supported by Grants-in-Aid for Scientific Research 21613006 (Fumitaka Homae) and 20670001 (Gentaro Taga). The authors thank Kayo Asakawa for technical and administrative assistance.

REFERENCES

- Albert, N. B., Robertson, E. M., and Miall, R. C. (2009). The resting human brain and motor learning. *Curr. Biol.* 19, 1023–1027.
- Baier, B., Kleinschmidt, A., and Müller, N. G. (2006). Cross-modal processing in early visual and auditory cortices depends on expected statistical relationship of multisensory information. *J. Neurosci.* 26, 12260–12265.
- Bavelier, D., and Neville, H. J. (2002). Cross-modal plasticity: where and how? *Nat. Rev. Neurosci.* 3, 443–452.
- Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate – a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B Methodol.* 57, 289–300.
- Bortfeld, H., Fava, E., and Boas, D. A. (2009). Identifying cortical lateralization of speech processing in infants using near-infrared spectroscopy. *Dev. Neuropsychol.* 34, 52–65.
- Catani, M., Howard, R. J., Pajevic, S., and Jones, D. K. (2002). Virtual in vivo interactive dissection of white matter fasciculi in the human brain. *Neuroimage* 17, 77–94.
- Dehaene, S., Kerszberg, M., and Changeux, J. P. (1998). A neuronal model of a global workspace in effortful cognitive tasks. *Proc. Natl. Acad. Sci. U.S.A.* 95, 14529–14534.
- Dehaene-Lambertz, G., Dehaene, S., and Hertz-Pannier, L. (2002). Functional neuroimaging of speech perception in infants. *Science* 298, 2013–2015.
- Dehaene-Lambertz, G., Hertz-Pannier, L., Dubois, J., Meriaux, S., Roche, A., Sigman, M., and Dehaene, S. (2006). Functional organization of perisylvian activation during presentation of sentences in preverbal infants. *Proc. Natl. Acad. Sci. U.S.A.* 103, 14240–14245.
- Dubois, J., Hertz-Pannier, L., Cachia, A., Mangin, J. F., Le Bihan, D., and Dehaene-Lambertz, G. (2009). Structural asymmetries in the infant language and sensorimotor networks. *Cereb. Cortex* 19, 414–423.
- Eckert, M. A., Kamdar, N. V., Chang, C. E., Beckmann, C. F., Greicius, M. D., and Menon, V. (2008). A cross-modal system linking primary auditory and visual cortices: evidence from intrinsic fMRI connectivity analysis. *Hum. Brain Mapp.* 29, 848–857.
- Falchier, A., Clavagnier, S., Barone, P., and Kennedy, H. (2002). Anatomical evidence of multimodal integration in primate striate cortex. *J. Neurosci.* 22, 5749–5759.
- Fox, M. D., Snyder, A. Z., Vincent, J. L., Corbetta, M., Van Essen, D. C., and Raichle, M. E. (2005). The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *Proc. Natl. Acad. Sci. U.S.A.* 102, 9673–9678.
- Fransson, P., Aden, U., Blennow, M., and Lagercrantz, H. (2011). The functional architecture of the infant brain as revealed by resting-state fMRI. *Cereb. Cortex* 21, 145–154.
- Fransson, P., Skiold, B., Horsch, S., Nordell, A., Blennow, M., Lagercrantz, H., and Aden, U. (2007). Resting-state networks in the infant brain. *Proc. Natl. Acad. Sci. U.S.A.* 104, 15531–15536.
- Friederici, A. D. (2006). Broca's area and the ventral premotor cortex in language: functional differentiation and specificity. *Cortex* 42, 472–475.
- Friederici, A. D., Bahlmann, J., Heim, S., Schubotz, R. I., and Anwander, A. (2006). The brain differentiates human and non-human grammars: functional localization and structural connectivity. *Proc. Natl. Acad. Sci. U.S.A.* 103, 2458–2463.
- Gao, W., Zhu, H., Giovanello, K. S., Smith, J. K., Shen, D., Gilmore, J. H., and Lin, W. (2009). Evidence on the emergence of the brain's default network from 2-week-old to 2-year-old healthy pediatric subjects. *Proc. Natl. Acad. Sci. U.S.A.* 106, 6790–6795.
- Genovese, C. R., Lazar, N. A., and Nichols, T. (2002). Thresholding of statistical maps in functional neuroimaging using the false discovery rate. *Neuroimage* 15, 870–878.
- Gervain, J., Macagno, F., Cogoi, S., Peña, M., and Mehler, J. (2008). The neonate brain detects speech structure. *Proc. Natl. Acad. Sci. U.S.A.* 105, 14222–14227.
- Gervain, J., Mehler, J., Werker, J. F., Nelson, C. A., Csibra, G., Lloyd-Fox, S., Shukla, M., and Aslin, R. N. (2011). Near-infrared spectroscopy: a report from the McDonnell infant methodology consortium. *Dev. Cogn. Neurosci.* 1, 22–46.
- Hauser, M. D., Chomsky, N., and Fitch, W. T. (2002). The faculty of language: what is it, who has it, and how did it evolve? *Science* 298, 1569–1579.
- Herwig, U., Satrapi, P., and Schonfeldt-Lecuona, C. (2003). Using the International 10–20 EEG system for positioning of transcranial magnetic stimulation. *Brain Topogr.* 16, 95–99.
- Hill, J., Dierker, D., Neil, J., Inder, T., Knutsen, A., Harwell, J., Coalson, T., and Van Essen, D. (2010). A surface-based analysis of hemispheric asymmetries and folding of cerebral cortex in term-born human infants. *J. Neurosci.* 30, 2268–2276.
- Homae, F., Watanabe, H., Nakano, T., Asakawa, K., and Taga, G. (2006). The right hemisphere of sleeping infant perceives sentential prosody. *Neurosci. Res.* 54, 276–280.
- Homae, F., Watanabe, H., Nakano, T., and Taga, G. (2007). Prosodic processing in the developing brain. *Neurosci. Res.* 59, 29–39.
- Homae, F., Watanabe, H., Nakano, T., and Taga, G. (in press). Functional development in the infant brain for auditory pitch processing. *Hum. Brain Mapp.*
- Homae, F., Watanabe, H., Otobe, T., Nakano, T., Go, T., Konishi, Y., and Taga, G. (2010). Development of global cortical networks in early infancy. *J. Neurosci.* 30, 4877–4882.
- Homan, R. W., Herman, J., and Purdy, P. (1987). Cerebral location of international 10–20 system electrode placement. *Electroencephalogr. Clin. Neurophysiol.* 66, 376–382.
- Hoshi, Y. (2007). Functional near-infrared spectroscopy: current status and future prospects. *J. Biomed. Opt.* 12, 062106.
- Imada, T., Zhang, Y., Cheour, M., Taulu, S., Ahonen, A., and Kuhl, P. K. (2006). Infant speech perception activates Broca's area: a developmental magnetoencephalography study. *Neuroreport* 17, 957–962.
- Jöbsis, F. F. (1977). Noninvasive, infrared monitoring of cerebral and myocardial oxygen sufficiency and circulatory parameters. *Science* 198, 1264–1267.
- Kotilahti, K., Nissilä, I., Huottilainen, M., Mäkelä, R., Gavrielides, N., Noponen, T., Björkman, P., Fellman, V., and Katila, T. (2005). Bilateral hemodynamic responses to auditory stimulation in newborn infants. *Neuroreport* 16, 1373–1377.
- Kotilahti, K., Nissilä, I., Näsi, T., Lipiäinen, L., Noponen, T., Meriläinen, P., Huottilainen, M., and Fellman, V. (2010). Hemodynamic responses to speech and music in newborn infants. *Hum. Brain Mapp.* 31, 595–603.
- Lagercrantz, H., and Changeux, J. P. (2009). The emergence of human consciousness: from fetal to neonatal life. *Pediatr. Res.* 65, 255–260.
- Lloyd-Fox, S., Blasi, A., and Elwell, C. E. (2010). Illuminating the developing brain: the past, present and future of functional near infrared spectroscopy. *Neurosci. Biobehav. Rev.* 34, 269–284.
- Maki, A., Yamashita, Y., Ito, Y., Watanabe, E., Mayanagi, Y., and Koizumi, H. (1995). Spatial and temporal analysis of human motor activity using non-invasive NIR topography. *Med. Phys.* 22, 1997–2005.
- Makris, N., Kennedy, D. N., McInerney, S., Sorensen, A. G., Wang, R., Caviness, V. S., and Pandya, D. N. (2005). Segmentation of subcomponents within the superior longitudinal fascicle in humans: a quantitative, in vivo, DT-MRI study. *Cereb. Cortex* 15, 854–869.
- Makris, N., Papadimitriou, G. M., Sorg, S., Kennedy, D. N., Caviness, V. S., and Pandya, D. N. (2007). The occipitofrontal fascicle in humans: a quantitative, in vivo, DT-MRI study. *Neuroimage* 37, 1100–1111.
- Martuzzi, R., Murray, M. M., Michel, C. M., Thiran, J. P., Maeder, P. P., Clarke, S., and Meuli, R. A. (2007). Multisensory interactions within human primary cortices revealed by BOLD dynamics. *Cereb. Cortex* 17, 1672–1679.
- Mesquita, R. C., Franceschini, M. A., and Boas, D. A. (2010). Resting state functional connectivity of the whole head with near-infrared spectroscopy. *Biomed. Opt. Express* 1, 324–336.
- Minagawa-Kawai, Y., Mori, K., Hebden, J. C., and Dupoux, E. (2008). Optical Imaging of infants' neurocognitive development: recent advances and perspectives. *Dev. Neurobiol.* 68, 712–728.
- Minagawa-Kawai, Y., Mori, K., Naoi, N., and Kojima, S. (2007). Neural attunement processes in infants during the acquisition of a language-specific phonemic contrast. *J. Neurosci.* 27, 315–321.
- Moore, J. K. (2002). Maturation of human auditory cortex: implications for speech perception. *Ann. Otol. Rhinol. Laryngol. Suppl.* 189, 7–10.
- Nakano, T., Homae, F., Watanabe, H., and Taga, G. (2008). Anticipatory cortical activation precedes auditory events in sleeping infants. *PLoS ONE* 3, e3912. doi: 10.1371/journal.pone.0003912
- Nakano, T., Watanabe, H., Homae, F., and Taga, G. (2009). Prefrontal cortical involvement in young infants' analysis of novelty. *Cereb. Cortex* 19, 455–463.
- Noesselt, T., Tyll, S., Boehler, C. N., Buderer, E., Heinze, H. J., and Driver, J. (2010). Sound-induced enhancement of low-intensity vision: multisensory influences on human sensory-specific cortices and thalamic bodies relate to perceptual enhancement of visual detection sensitivity. *J. Neurosci.* 30, 13609–13623.
- Obrig, H., and Villringer, A. (2003). Beyond the visible – imaging the human brain with light. *J. Cereb. Blood Flow Metab.* 23, 1–18.
- Okamoto, M., Dan, H., Sakamoto, K., Takeo, K., Shimizu, K., Kohno, S., Oda, I., Isobe, S., Suzuki, T., Kohyama, K., and Dan, I. (2004). Three-dimensional probabilistic anatomical cranio-cerebral correlation via the international 10–20 system oriented for transcranial functional brain mapping. *Neuroimage* 21, 99–111.

- Peña, M., Maki, A., Kovacic, D., Dehaene-Lambertz, G., Koizumi, H., Bouquet, F., and Mehler, J. (2003). Sounds and silence: an optical topography study of language recognition at birth. *Proc. Natl. Acad. Sci. U.S.A.* 100, 11702–11705.
- Petrides, M., and Pandya, D. N. (1988). Association fiber pathways to the frontal cortex from the superior temporal region in the rhesus monkey. *J. Comp. Neurol.* 273, 52–66.
- Poepfel, D., Idsardi, W. J., and van Wassenhove, V. (2008). Speech perception at the interface of neurobiology and linguistics. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 363, 1071–1086.
- Portas, C. M., Krakow, K., Allen, P., Josephs, O., Armony, J. L., and Frith, C. D. (2000). Auditory processing across the sleep-wake cycle: simultaneous EEG and fMRI monitoring in humans. *Neuron* 28, 991–999.
- Price, C. J. (2010). The anatomy of language: a review of 100 fMRI studies published in 2009. *Ann. N.Y. Acad. Sci.* 1191, 62–88.
- Reynolds, E. O. R., Wyatt, J. S., Azzopardi, D., Delpy, D. T., Cady, E. B., Cope, M., and Wray, S. (1988). New non-invasive methods for assessing brain oxygenation and hemodynamics. *Br. Med. Bull.* 44, 1052–1075.
- Rockland, K. S., and Ojima, H. (2003). Multisensory convergence in calcarine visual areas in macaque monkey. *Int. J. Psychophysiol.* 50, 19–26.
- Romanski, L. M., Tian, B., Fritz, J., Mishkin, M., Goldman-Rakic, P. S., and Rauschecker, J. P. (1999). Dual streams of auditory afferents target multiple domains in the primate prefrontal cortex. *Nat. Neurosci.* 2, 1131–1136.
- Sasai, S., Homae, F., Watanabe, H., and Taga, G. (2011). Frequency-specific functional connectivity in the brain during resting state revealed by NIRS. *Neuroimage* 56, 252–257.
- Sato, Y., Sogabe, Y., and Mazuka, R. (2010). Development of hemispheric specialization for lexical pitch-accent in Japanese infants. *J. Cogn. Neurosci.* 22, 2503–2513.
- Singh, A. K., and Dan, I. (2006). Exploring the false discovery rate in multichannel NIRS. *Neuroimage* 33, 542–549.
- Smyser, C. D., Inder, T. E., Shimony, J. S., Hill, J. E., Degnan, A. J., Snyder, A. Z., and Neil, J. J. (2010). Longitudinal analysis of neural network development in preterm infants. *Cereb. Cortex* 20, 2852–2862.
- Steinmetz, H., Fürst, G., and Meyer, B. U. (1989). Craniocerebral topography within the international 10–20 system. *Electroencephalogr. Clin. Neurophysiol.* 72, 499–506.
- Taga, G., and Asakawa, K. (2007). Selectivity and localization of cortical response to auditory and visual stimulation in awake infants aged 2 to 4 months. *Neuroimage* 36, 1246–1252.
- Taga, G., Homae, F., and Watanabe, H. (2007). Effects of source-detector distance of near infrared spectroscopy on the measurement of the cortical hemodynamic response in infants. *Neuroimage* 38, 452–460.
- Telkemeyer, S., Rossi, S., Koch, S. P., Nierhaus, T., Steinbrink, J., Poeppel, D., Obrig, H., and Wartenburger, I. (2009). Sensitivity of newborn auditory cortex to the temporal structure of sounds. *J. Neurosci.* 29, 14726–14733.
- Tomassini, V., Jbabdi, S., Klein, J. C., Behrens, T. E. J., Pozzilli, C., Matthews, P. M., Rushworth, M. F. S., and Johansen-Berg, H. (2007). Diffusion-weighted Imaging tractography-based parcellation of the human lateral premotor cortex identifies dorsal and ventral subregions with anatomical and functional specializations. *J. Neurosci.* 27, 10259–10269.
- Villringer, A., and Chance, B. (1997). Non-invasive optical spectroscopy and imaging of human brain function. *Trends Neurosci.* 20, 435–442.
- Watanabe, H., Homae, F., Nakano, T., and Taga, G. (2008). Functional activation in diverse regions of the developing brain of human infants. *Neuroimage* 43, 346–357.
- Watanabe, H., Homae, F., and Taga, G. (2010). General to specific development of functional activation in the cerebral cortexes of 2- to 3-month-old infants. *Neuroimage* 50, 1536–1544.
- White, B. R., Snyder, A. Z., Cohen, A. L., Petersen, S. E., Raichle, M. E., Schlaggar, B. L., and Culver, J. P. (2009). Resting-state functional connectivity in the human brain revealed with diffuse optical tomography. *Neuroimage* 47, 148–156.
- Wolf, M., Ferrari, M., and Quaresima, V. (2007). Progress of near-infrared spectroscopy and topography for brain and muscle clinical applications. *J. Biomed. Opt.* 12, 062104.
- Zatorre, R. J., Belin, P., and Penhune, V. B. (2002). Structure and function of auditory cortex: music and speech. *Trends Cogn. Sci.* 6, 37–46.
- Zatorre, R. J., and Gandour, J. T. (2008). Neural specializations for speech and pitch: moving beyond the dichotomies. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 363, 1087–1104.
- Zhang, J. Y., Evans, A., Hermoye, L., Lee, S. K., Wakana, S., Zhang, W. H., Donohue, P., Miller, M. I., Huang, H., Wang, X. Q., van Zijl, P. C. M., and Mori, S. (2007). Evidence of slow maturation of the superior longitudinal fasciculus in early childhood by diffusion tensor imaging. *Neuroimage* 38, 239–247.
- Zhang, Y. J., Lu, C. M., Biswal, B. B., Zang, Y. F., Peng, D. L., and Zhu, C. Z. (2010). Detecting resting-state functional connectivity in the language system using functional near-infrared spectroscopy. *J. Biomed. Opt.* 15, 047003.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 30 December 2010; paper pending published: 07 February 2011; accepted: 28 April 2011; published online: 17 May 2011.
Citation: Homae F, Watanabe H, Nakano T and Taga G (2011) Large-scale brain networks underlying language acquisition in early infancy. *Front. Psychology* 2:93. doi: 10.3389/fpsyg.2011.00093
This article was submitted to *Frontiers in Language Sciences*, a specialty of *Frontiers in Psychology*.
Copyright © 2011 Homae, Watanabe, Nakano and Taga. This is an open-access article subject to a non-exclusive license between the authors and Frontiers Media SA, which permits use, distribution and reproduction in other forums, provided the original authors and source are credited and other Frontiers conditions are complied with.



Neural processing of repetition and non-repetition grammars in 7- and 9-month-old infants

Jennifer B. Wagner^{1*}, Sharon E. Fox^{1,2}, Helen Tager-Flusberg³ and Charles A. Nelson¹

¹ Laboratories of Cognitive Neuroscience, Division of Developmental Medicine, Children's Hospital Boston/Harvard Medical School, Boston, MA, USA

² MIT Health Sciences and Technology, Cambridge, MA, USA

³ Department of Psychology, Boston University, Boston, MA, USA

Edited by:

Judit Gervain, CNRS–Université Paris Descartes, France

Reviewed by:

Judit Gervain, CNRS–Université Paris Descartes, France

Hugh Rabagliati, Brown University, USA

*Correspondence:

Jennifer B. Wagner, Laboratories of Cognitive Neuroscience, Division of Developmental Medicine, Children's Hospital Boston, 1 Autumn Street, AU641, Boston, MA 02215, USA.
e-mail: jen.wagner@alum.mit.edu

An essential aspect of infant language development involves the extraction of meaningful information from a continuous stream of auditory input. Studies have identified early abilities to differentiate auditory input along various dimensions, including the presence or absence of structural regularities. In newborn infants, frontal and temporal regions were found to respond differentially to these regularities (Gervain et al., 2008), and in order to examine the development of this abstract rule learning we presented 7- and 9-month-old infants with syllables containing an ABB pattern (e.g., “balolo”) or an ABC pattern (e.g., “baloti”) and measured activity in left and right lateral brain regions using near-infrared spectroscopy (NIRS). While prior newborn work found increases in oxyhemoglobin (oxyHb) activity in response to ABB blocks as compared to ABC blocks in anterior regions, 7- and 9-month-olds showed no differentiation between grammars in oxyHb. However, changes in deoxyhemoglobin (deoxyHb) pointed to a developmental shift, whereby 7-month-olds showed deoxyHb responding significantly different from zero for ABB blocks, but not ABC blocks, and 9-month-olds showed the opposite pattern, with deoxyHb responding significantly different from zero for the ABC blocks but not the ABB blocks. DeoxyHb responses were more pronounced over anterior regions. A grammar by time interaction also illustrated that during the early blocks, deoxyHb was significantly greater to ABC than in later blocks, but there was no change in ABB activation over time. The shift from stronger activation to ABB in newborns (Gervain et al., 2008) and 7-month-olds in the present study to stronger activation to ABC by 9-month-olds here is discussed in terms of changes in stimulus salience and novelty preference over the first year of life. The present discussion also highlights the importance of future work exploring the coupling between oxyHb and deoxyHb activation in infant NIRS studies.

Keywords: infancy, auditory processing, NIRS, optical imaging, language

INTRODUCTION

The first year of life represents an important period for language development that culminates with infants producing their first words. Even before they begin speaking, however, research has shown that infants can pick up on important and complex linguistic cues from their auditory input, including rhythmic cues (Mehler et al., 1988), phonemic contrasts (e.g., Werker and Tees, 1984; Kuhl et al., 2006), structural regularities (e.g., Christophe et al., 1994), and transitional probabilities (e.g., Saffran et al., 1996; Teinonen et al., 2009; for a review of early infant language development, see Gervain and Mehler, 2010).

A rich set of behavioral work with infants during the first year of life has highlighted abilities for extracting rules from stimuli varying in repetition structure. Seminal work by Marcus et al. (1999) found that 7-month-old infants tracked an embedded syllable structure and later used this to generalize to novel stimuli. Subsequent work has focused on the generalizability of this rule-learning mechanism, including the ability to extract rule information when syllables are substituted with other auditory sounds, such as tones (e.g., Marcus et al., 2007; Dawson and Gerken, 2009), the use of redundant cues to support rule learning (e.g., Frank et al., 2009),

changing stimulus salience over development (e.g., Dawson and Gerken, 2009), and rule abstraction across visual domains (Saffran et al., 2007; Johnson et al., 2009). The behavioral literature points to shifting rule abstraction abilities between 5- and 12-months, but little work has examined the neural correlates of pattern detection abilities in this age group.

A growing set of infant language studies have utilized near-infrared spectroscopy (NIRS), an optical imaging method that assesses hemodynamic response in awake participants, to examine the brain regions activated during processing and discrimination of linguistically relevant auditory input within the first few months of life (e.g., Pena et al., 2003; Homae et al., 2006; Saito et al., 2007a,b; Gervain et al., 2008; Telkemeyer et al., 2009; see Obrig et al., 2010; Gervain et al., 2011, for reviews). For example, Pena et al. (2003) presented newborn infants with blocks of forward and backward speech and used NIRS to measure hemodynamic response in the left and right hemisphere during these auditory streams. During forward speech, Pena et al. (2003) found greater activation in left temporal areas. More specifically relating to the rule-learning literature, Gervain et al. (2008) also examined auditory processing in newborn infants using NIRS, presenting blocks of trisyllabic words

with either an ABB pattern or an ABC pattern. Newborns showed greater activation to the ABB grammar than the ABC grammar, providing evidence that infants are capable of extracting regularities from their speech stream from birth (Gervain et al., 2008).

Although NIRS has provided researchers with the opportunity to examine the neural bases of auditory processing, and in the case of Gervain et al. (2008), repetition detection in very young infants, few studies thus far have used this novel method to study auditory processing in infants beyond 6-months-of-age (e.g., Homae et al., 2007; Minagawa-Kawai et al., 2007; Sato et al., 2009; see Obrig et al., 2010, for a review). By looking at a single auditory processing task over a wider range of ages, researchers can develop a more comprehensive understanding of how auditory and linguistic processing change over the course of infancy. Work by Minagawa-Kawai et al. (2007) adopted this developmental approach and used NIRS to examine the discrimination of phonemic contrasts in infants ranging from 3- to 28-months-of-age. Minagawa-Kawai et al. (2007) identified a developmental shift in oxyhemoglobin responses to across-category and within-category phonemic contrasts, as well as changes in hemispheric lateralization over the first 2 years of life.

The present study aimed to examine the neural basis of linguistic processing of repetition- and non-repetition-based grammars during the first year of life, extending from the paradigm developed by Gervain et al. (2008). With a host of behavioral work showing subtle shifts in development with respect to abstracting rules from input, this work will contribute to a neural framework of learning and pattern extraction. Seven- and 9-month-old infants were presented with grammars that did or did not contain syllables with a repetition (ABB vs. ABC), and NIRS was used to capture hemodynamic changes throughout the task. This work aimed to provide a richer picture of the neural correlates of auditory processing and pattern detection, across the first year of life.

MATERIALS AND METHODS

PARTICIPANTS

The final sample consisted of thirteen 7-month-old infants (mean age = 208 days, SD = 17; six female infants) and fifteen 9-month-old infants (mean age = 286 days, SD = 11; nine female infants). An additional five 7-month-old infants and two 9-month-old infants participated in the experiment, but were excluded due to excessive noise due to hair and/or motion artifact that resulted in missing data for over 30% of channels. All infants included in the experiment were: (1) born after 36 weeks gestational age, (2) born weighing more than 2500 g, and (3) born without a known neurological abnormality. Project approval was obtained from the Institutional Review Board of Children's Hospital Boston. Written informed consent was obtained from the parents of all infant participants.

STIMULI

Stimuli consisted of trisyllabic sequences identical to those used in a prior study by Gervain et al. (2008) with newborn infants that presented a repetition based ABB artificial grammar (e.g., "balolo") and an unstructured ABC control grammar (e.g., "baloti"; see Gervain et al. for further details of the ABB and ABC grammar construction). ABB and ABC grammars were matched in syllabic repertoire, frequency of syllables, and transitional probabilities between syllables.

Trisyllabic sequences were grouped into blocks of 10, with intervals of 0.5–1.5 s of silence occurring between each sequence. Each block was, on average, 16 s in length. Blocks of the ABB and ABC grammars were presented in one of two semi-randomized sequences, and were separated by a silent pause of varying duration (15 s minimum).

APPARATUS

A Hitachi ETG-4000 NIRS system with 24 simultaneously recording channels was used to collect hemodynamic response during stimulus presentation. Two wavelengths of light (695 and 830 nm) were used to measure cortical levels of oxyhemoglobin and deoxyhemoglobin (oxyHb and deoxyHb). The near-infrared light was guided by optical fiber bundles that were 1 mm in diameter. On the ETG-4000 device, each pair of adjacent incident and detection fibers defines a single measurement channel, allowing the measurement of the hemodynamic changes in the brain corresponding to a specific stimulus. The NIRS probes consist of two 3 × 3 chevron arrays, each with five emitting and four detecting fibers held in place by a silicone support with 3 cm spacing, and attached to a soft cap designed for infants (see Figure 1A). The flexibility of the silicone supported by an adjustable neoprene band allowed for a bilateral placement of the probes spanning from anterior to posterior lateral regions on each side of the head (see Figure 1B).

PROCEDURE

Infants were seated on a parent's lap throughout the experiment. Infants passively listened to the auditory stimuli, which were presented through speakers placed behind a curtain in front of the infants. During presentation of the ABB and ABC blocks, infants were presented with a video containing shapes moving on a screen, and when uninterested in the videos, an experimenter used silent toys and bubbles to keep infants calm and still. Infants who became fussy were permitted to nurse, feed from a bottle, or eat finger foods to expose them to as many blocks as possible. While this had the potential to introduce motion artifacts, past work recording EEG during auditory stimuli (a method much more susceptible to motion artifacts than NIRS) with infants who were nursing or bottle-feeding obtained sufficient artifact-free data for analyses under similar circumstances (Thomas and Lykins, 1995; Little et al., 1999). Generally, trials were presented until the infant heard all 28 blocks, or until the infant became too fussy to continue.

DATA ANALYSIS

Based on the light intensity detected through each channel, relative concentrations of oxyHb and deoxyHb were calculated from absorbance at each wavelength using the modified Beer–Lambert law. This conversion, as well as further data analyses, was implemented through customized Matlab scripts (version 7.6, Mathworks Inc., Natick, MA, USA).

Timeseries corresponding to oxyHb and deoxyHb values were first processed using a fifth order Butterworth filter between 0.01 and 1.0 Hz, and additional artifacts were identified and extracted if the raw signal exceeded a threshold value (4.95), indicating saturation of the detector optodes, or if total hemoglobin change exceeded 0.3 mM*mm within a 0.7-s time window. For each subject with at least 10 trials of each condition (ABB and ABC) and no significant

artifacts after initial filtering, the data from each block were parsed into a 16-s time window with 0.1 s time resolution beginning at stimulus onset. In addition to inter-block intervals of at least 15 s to allow the hemodynamic response to return to baseline, each 16 s time window was corrected to a zero baseline value at stimulus onset in order to allow for standardized comparisons across conditions. Specifically, for each trial, the oxyHb and deoxyHb value in the first 0.1 s time bin was subtracted from itself to create a zero baseline in each channel at the start of every trial. This value was then subtracted from each subsequent time point for a given channel to adjust the entire trial by this baseline correction within the initial time bin.

Each block from a single participant was grouped by grammar type and averaged to obtain mean values of oxyHb and deoxyHb across time for each individual channel during the ABB and ABC blocks. On a subject-by-subject level, any channels with low oxyHb signal-to-noise (Mean/SD < 1.0) were excluded from subsequent analyses. If this analysis of noise revealed a loss of more than 30% of channels, infants were excluded from further analyses. On average, infants included in subsequent analyses lost 12% of channels due to low signal-to-noise (SD = 8%).

Based on the work of Gervain et al. (2008), two sets of analyses were conducted on the present data. The primary analyses examined the average value of oxyHb and deoxyHb across the 16-s time window for each grammar type at each channel. A second analysis calculated the average oxyHb and deoxyHb response from the first four blocks of each condition and the last four blocks of each condition for every channel to examine differential activation across time. For both analyses, channels were grouped into four regions, Left Anterior (channels 1–5), Right Anterior (channels 20–24), Left Posterior (channels 6–12), and Right Posterior (channels 13–19). The probes were positioned such that posterior regions were posterior to the infant's ear (see **Figure 1**).

RESULTS

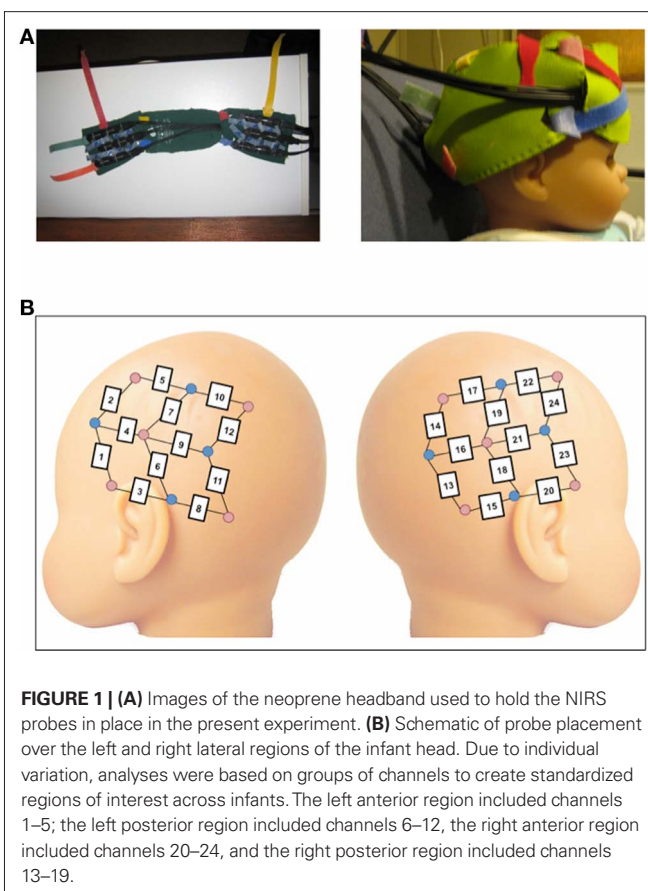
Figure 2 illustrates the grand average oxyHb and deoxyHb responses to the ABB and ABC grammars for each age group across all trials, collapsed across the 24 channels used in subsequent analyses. Preliminary repeated-measures ANOVAs were run to examine the between-subjects effect of gender for each of the four analyses outlined below (average oxyHb response, average deoxyHb response, oxyHb response over time, and deoxyHb response over time). No significant main effects were found and subsequent analyses collapsed across gender.

Initial analyses examined the changes in oxyHb and deoxyHb by the ABB and ABC grammars in each channel using paired *t*-tests for each age group. No single channel showed a significant difference between conditions in oxyHb or deoxyHb at either age (channel-by-channel *t*-maps for oxyHb and deoxyHb in 7- and 9-month-old infants are illustrated in the Appendix).

AVERAGE OXYHB AND DEOXYHB RESPONSE ACROSS ALL TRIALS

Average oxyHb response

Based on the average oxyHb concentration calculated across the 16-s time window for all trials, a 2 (Grammar: ABB, ABC) × 2 (Hemisphere: Left, Right) × 2 (Region: Anterior, Posterior) × 2 (Age: 7-month-old, 9-month-old) repeated-measures ANOVA



with grammar, hemisphere, and region as within-subjects factors and age as the between-subjects factor revealed a marginal three-way interaction between grammar, hemisphere, and region, $F(1,26) = 3.29$, $p = 0.08$, $\eta_p^2 = 0.11$, but no other significant main effects or interactions ($ps > 0.12$). This trend showed that oxyHb responses in the left hemisphere appeared similar in the anterior and posterior regions for the ABB grammar (ABB Left Anterior: $M = 0.018$ mmol*mm, $SD = 0.009$; ABB Left Posterior: $M = 0.014$ mmol*mm, $SD = 0.013$) and for the ABC grammar (ABC Left Anterior: $M = 0.020$ mmol*mm, $SD = 0.010$; ABC Left Posterior: $M = 0.018$ mmol*mm, $SD = 0.014$). In the right hemisphere, however, the ABB grammar trended toward larger responding in the anterior region ($M = 0.026$ mmol*mm, $SD = 0.019$) as compared to the posterior region ($M = 0.005$ mmol*mm, $SD = 0.015$), while the ABC grammar showed the opposite pattern (ABC Right Anterior: $M = 0.004$ mmol*mm, $SD = 0.011$; ABC Right Posterior: $M = 0.025$ mmol*mm, $SD = 0.017$).

Average deoxyHb response

Using the average deoxyHb concentration calculated across the 16-s time window for all trials, a 2 (Grammar: ABB, ABC) × 2 (Hemisphere: Left, Right) × 2 (Region: Anterior, Posterior) × 2 (Age: 7-month-old, 9-month-old) repeated-measures ANOVA with grammar, hemisphere, and region as within-subjects factors and age as the between-subjects factor revealed a main effect of region, $F(1,26) = 13.19$, $p = 0.001$, $\eta_p^2 = 0.34$, whereby a stronger negative response was found over

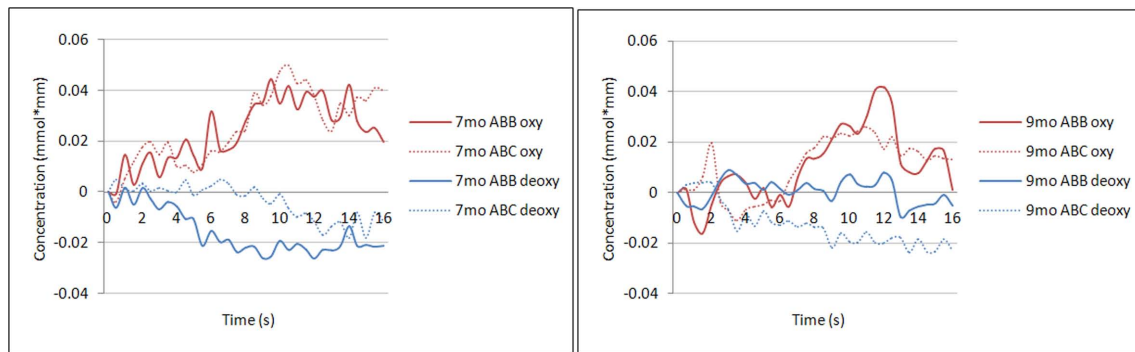


FIGURE 2 | Time course of oxyHb and deoxyHb. Illustration of the average oxyHb and deoxyHb concentration at each time point across the 16-s time window for the ABB and ABC grammar in the group of 7-month-old infants (left) and the group of 9-month-old infants (right). Concentrations are averaged across all 24 channels.

anterior regions ($M = -0.016$ mmol*mm, $SD = 0.016$) as compared to posterior regions ($M = -0.003$ mmol*mm, $SD = 0.018$). A significant interaction between hemisphere and region was also identified, $F(1,26) = 5.71$, $p = 0.024$, $\eta_p^2 = 0.18$. *Post hoc* pairwise comparisons revealed that for anterior regions, the right hemisphere showed a greater negative deoxyHb response ($M = -0.021$ mmol*mm, $SD = 0.023$) as compared to the left hemisphere ($M = -0.012$ mmol*mm, $SD = 0.016$), $t(27) = 2.11$, $p = 0.044$, $d = 0.46$, but in posterior regions, there was no significant difference between hemispheres, $t(27) = -0.92$, $p = 0.37$, $d = 0.19$ (Left Posterior: $M = -0.005$ mmol*mm, $SD = 0.021$; Right Posterior: $M = -0.001$ mmol*mm, $SD = 0.022$).

Further, a marginal interaction was observed between grammar and age, $F(1,26) = 3.18$, $p = 0.086$, $\eta_p^2 = 0.11$, whereby 7-month-old infants showed a greater negative deoxyHb response to the ABB grammar ($M = -0.018$ mmol*mm, $SD = 0.022$) as compared to the ABC grammar ($M = -0.005$ mmol*mm, $SD = 0.024$), and 9-month-old infants showed the opposite pattern, with larger negative deoxyHb responding to the ABC grammar ($M = -0.014$ mmol*mm, $SD = 0.025$) as compared to the ABB grammar ($M = -0.002$ mmol*mm, $SD = 0.023$). Follow-up one-sample *t*-tests revealed that in 7-month-olds, the ABB deoxyHb response was significantly different from zero, $t(12) = -2.879$, $p = 0.014$, $d = 0.80$, and in 9-month-olds, the ABC deoxyHb response was significantly different from zero, $t(14) = -2.220$, $p = 0.043$, $d = 0.57$. The 7-month-old ABC deoxyHb response and the 9-month-old ABB deoxyHb responses were not significantly different from zero ($ps > 0.45$; **Figure 3**).

AVERAGE OXYHB AND DEOXYHB RESPONSE: EARLY VS. LATE BLOCKS

Based on the work of Gervain et al. (2008), hemodynamic activity was compared between the first four blocks of each condition and the last four blocks of each condition to examine differential activation over time. These analyses included thirteen 7-month-old-infants, and eleven 9-month-old infants. The remaining four 9-month-olds who were included in the first analysis were excluded from the response-over-time analysis because they made it through less than 21 trials before becoming fussy.

OxyHb response across blocks

A 2 (Time: First 4 blocks, Last 4 blocks) \times 2 (Grammar: ABB, ABC) \times 2 (Hemisphere: Left, Right) \times 2 (Region: Anterior, Posterior) \times 2 (Age: 7-month-old, 9-month-old) repeated-measures

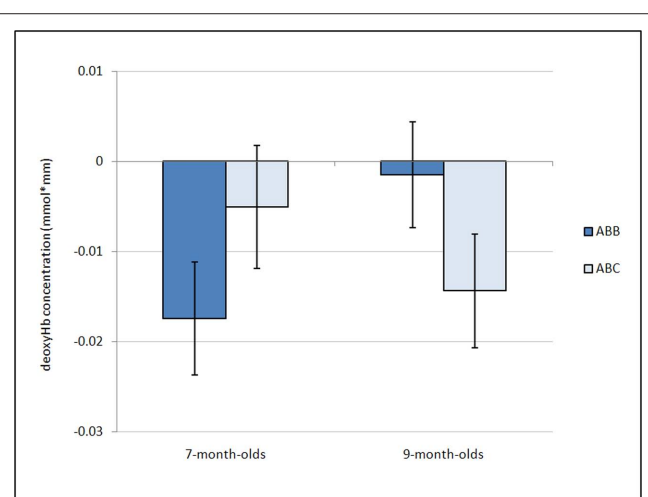


FIGURE 3 | Average deoxyHb concentration. Analysis of mean deoxyHb concentration revealed a marginal age by condition interaction ($p = 0.08$). *Post hoc* analyses revealed that 7-month-olds show a negative response significantly different from zero for the ABB grammar, while the 9-month-olds show a negative response significantly different from zero for the ABC grammar. Error bars represent $\pm SE$.

ANOVA with time, grammar, hemisphere, and region as within-subjects factors and age as the between-subjects factor revealed no main effects or interactions.

DeoxyHb response across blocks

Parallel to the analysis described above, a repeated-measures ANOVA with time, grammar, hemisphere, and region as within-subjects factors and age as the between-subjects factor examined deoxyHb responding within the first four and last four blocks of the experiment. This analysis revealed several findings. First, as in the average deoxyHb analysis, a main effect of region was observed, $F(1,22) = 19.09$, $p < 0.001$, $\eta_p^2 = 0.47$, with greater activation in anterior regions ($M = -0.015$ mmol*mm, $SD = 0.026$) than posterior regions ($M = -0.003$ mmol*mm, $SD = 0.020$). Additionally, there was a significant interaction between grammar and time, $F(1,22) = 5.80$, $p = 0.025$, $\eta_p^2 = 0.21$, such that the ABC grammar showed a significantly greater negative deoxyHb response during

the first four blocks than the last four blocks, $t(23) = 2.56$, $p = 0.017$, $d = 0.61$, but the ABB grammar showed no difference across time, $t(23) = 0.77$, $p = 0.45$, $d = 0.21$. One-sample t -tests revealed that the ABC response during the first four blocks was significantly different from zero, $t(23) = -2.94$, $p = 0.007$, $d = 0.60$, and that the ABB response during the last four blocks was marginally different from zero, $t(23) = -1.83$, $p = 0.08$, $d = 0.37$.

Finally, a marginal interaction between grammar and age was also found, $F(1,22) = 4.06$, $p = 0.056$, $\eta_p^2 = 0.16$, with 7-month-olds showing greater activation to the ABB grammar than the ABC grammar, and 9-month-olds showing greater activation to ABC than ABB, though neither difference reaches significance ($ps > 0.1$). One-sample t -tests revealed that overall activation to the ABB grammar in 7-month-olds was significantly different from zero, $t(12) = -2.78$, $p = 0.017$, $d = 0.77$, and activation to the ABC grammar in 9-month-olds was significantly different from zero, $t(10) = -2.14$, $p = 0.058$, $d = 0.64$ (see **Figure 4**).

DISCUSSION

In the present study we examined oxyHb and deoxyHb responses to two artificial grammars, one containing syllables in an ABB pattern and the other containing syllables in an ABC pattern. Analyses examining the average oxyHb response to each grammar across all trials of the experiment as well as across early and late blocks found no differential responses relating to hemisphere or region and no differentiation between the ABB and ABC grammars. Across analyses examining deoxyHb responding, overall greater negative response was found in anterior regions as compared to posterior regions, and 7- and 9-month-olds showed trends toward differential responding to the ABB and ABC grammars. Specifically, 7-month-olds showed deoxyHb responding significantly different from zero for the ABB grammar and not the ABC grammar,

while 9-month-olds showed the opposite effect, with activation significantly different from zero for the ABC grammar and not the ABB grammar.

Average deoxyHb across all trials also revealed an interaction between region and hemisphere, with posterior regions showing no differences in responding across hemispheres, but significantly greater activation in the right anterior region than the left anterior region. When examining deoxyHb activation across early and late blocks, a grammar by time interaction was observed as well. Between the early and late blocks, ABC activation significantly decreased, while no difference in ABB responding was found between the two time points. The deoxyHb response to ABC during the early trials was significantly different than zero.

With regards to localization of the neural response, parallel to the oxyHb findings from Gervain et al. (2008), the present study revealed stronger deoxyHb activation in anterior regions as compared to posterior regions. However, no left hemisphere bias was found, contrary to prior studies using auditory stimuli with infants from 0- to 4-months-of-age (e.g., Dehaene-Lambertz, 2000; Dehaene-Lambertz et al., 2002; Pena et al., 2003). In fact, the only significant finding relating to hemisphere in the present work identified greater deoxyHb responding in right anterior regions as compared to left anterior regions. Although the work of Gervain et al. (2008) showed no absolute left hemisphere bias to the syllables in newborns, the oxyHb response to the ABB grammar did show a significantly greater response in the left hemisphere, raising the question of why the 7- and 9-month-old infants in the present study show little influence of the left hemisphere on hemodynamic activity. In considering important changes in speech perception that occur between 8- and 12-months-of-age (e.g., Werker, 1989), it is possible that the present stimuli, spoken by a native French speaker, could mark these stimuli as less linguistically meaningful to children as they become older. Further, recent work by Homae et al. (2007) found stronger right hemisphere activation for flattened speech streams, and because the present stimuli contained monotonous prosody, this could perhaps result in right hemisphere activation that tempers any left hemisphere dominance that might otherwise be found relating to rule learning and pattern extraction processes.

By using NIRS to examine the neural response to ABB and ABC grammars in 7- and 9-month-old infants, the present work extends the newborn ABB and ABC study of Gervain et al. (2008). Gervain et al. (2008) found greater oxyHb activation in response to the ABB grammar than the ABC grammar, and this was more pronounced in the anterior regions of the newborn brain. The current study found a developmental shift in this response, with 7-month-old infants showing a significant response to ABB blocks but not ABC blocks, and 9-month-olds showing a significant response to ABC blocks but not ABB blocks. The heightened neural attunement to differing patterns at different ages is consistent with behavioral work illustrating shifts in sensitivity to repetition patterns during the first year of life (e.g., Johnson et al., 2009).

With larger negative deoxyHb responses expected to correspond with greater activation to that stimulus (see Obrig et al., 2010, for a review), the present findings suggest changes in the salience of the ABB and ABC grammars across the first year of life, as newborns (Gervain et al., 2008) and 7-month-olds show more negative deoxyHb responding to the ABB grammar and

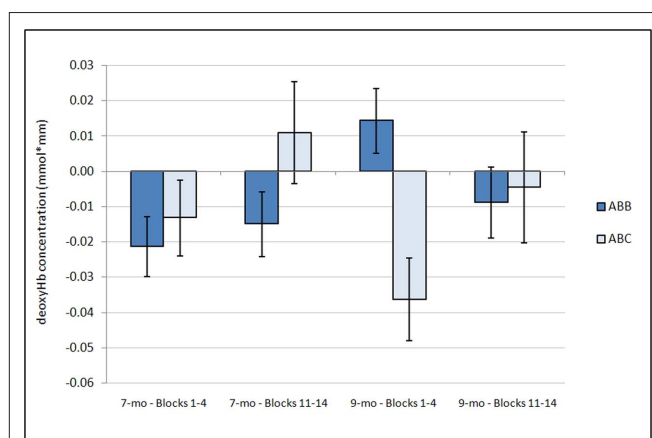


FIGURE 4 | Average deoxyHb concentration over time. Analysis of deoxyHb concentration during the first four blocks and the last four blocks revealed a significant interaction between condition and time ($p = 0.025$), with a significantly larger deoxyHb response for the ABC grammar during the early blocks than the later blocks ($p < 0.02$) and no change in ABB grammar responding across time ($p = 0.45$). Additionally, a marginal interaction was found between condition and age ($p = 0.056$). One-sample t -tests revealed that the 7-month-old ABB response during Blocks 1–4 and the 9-month-old ABC response during Blocks 1–4 are significantly different from zero. Error bars represent \pm SE.

9-month-olds show more negative responding to the ABC grammar. This shift in salience from ABB to ABC over the first year of life might parallel infant looking-time findings that show that, after a familiarization period, preference for the familiar stimulus is found in younger infants, while preference for the novel stimulus can be found in older infants (e.g., Ferry et al., 2010). Aslin (2007) describes this difference in terms of a balance, whereby "...stimuli can trigger a recognition response for familiarity that is stronger than the attentional response to novelty." The newborns tested by Gervain et al. (2008) could be neurally attuned to the familiarity and regularity of the repetition rule within the ABB blocks, while the ABC grammar could represent an overly complex stimulus to 1-day-old participants. On the other hand, in 7- to 9-month-olds who have been exposed to a wider range of auditory input and can proficiently extract regularities from the surrounding speech stream (e.g., Saffran et al., 1996), it is possible that the easily detectable repetition in ABB blocks is redundant and requires less neural processing than the consistently novel ABC input.

Gervain et al. (2008) found that over the course of their experiment, newborns showed an interaction between grammar and time, such that the ABB grammar showed increasing activation from the early blocks to the late blocks, while the ABC grammar showed no change. The present study also found an interaction between these two variables: 7- to 9-month-olds show a different response pattern, with significantly more activation to ABC during the first four blocks as compared to the last four blocks, and no difference between ABB activation across time. In the newborn study, infants become increasingly responsive to their preferred stimulus throughout the experiment, while 7- to 9-month-olds showed responding significantly greater than zero for only the ABC stimulus early on, but no significant difference from zero for either stimulus during late blocks. One probable explanation for this early differentiation followed by no differentiation in the 7- to 9-month-olds is that the experiment extended beyond their interest and led to diminished attention and responding to the differences between the ABB and ABC grammars. A recent review by Turk-Browne et al. (2008) discusses infant studies that resulted in either the habituation of responses across time or the enhancement of responses across time and the related phenomenon in adult neuroimaging work, but the factors that lead to habituation vs. enhancement remain unclear. In order to further delineate developmental shifts between the repetition enhancement found by Gervain et al. (2008) in newborns and the repetition suppression found in the present study with older infants, work is under way using the present auditory NIRS task with a wider age range of infants.

Finally, it is important to mention that while Gervain et al. (2008) found significant effects in their analyses of oxyHb activation, in the present study we only find effects in analyses of deoxyHb activation. Cortical activation is traditionally thought to result in both an increase in oxyHb and a decrease in deoxyHb; however, studies using NIRS often find stronger effects in one over the other. Many complex factors likely contribute to how robustly oxy- and deoxyHb activation are coupled in infant NIRS experiments, including the equipment used, the ages tested, and

the areas of the brain examined (see Gervain et al., 2011, for further discussion). While many infant NIRS studies have reported significant oxyHb effects alongside few or no significant deoxyHb effects (e.g., Saito et al., 2007a,b; Gervain et al., 2008; Lloyd-Fox et al., 2009; Sato et al., 2009), some studies have reported significant activation in both oxyHb and deoxyHb (e.g., Wilcox et al., 2005; Homae et al., 2006). The first year of life includes developmental changes in brain maturation such as changes in vascular tone and blood flow patterns, and NIRS work should continue to explore how this development differentially affects oxyHb and deoxyHb responding. The importance of understanding changes in deoxyHb has been emphasized by adult neuroimaging work that identifies deoxyHb as the species of hemoglobin correlated with the blood oxygenation level-dependent response found in fMRI studies (e.g., Kleinschmidt et al., 1996; see Steinbrink et al., 2006, for a review). This fMRI/NIRS literature supports data interpretation that gives at least as much weight to deoxyHb findings as oxyHb findings, an important consideration for future infant NIRS studies.

An important limitation of the present study is the varying attentional state of the infant participants throughout the session due to the range of silent activities permitted to keep them free of motion (e.g., watching videos, examining novel toys, drinking from a bottle, sleeping). Because of the more limited time span for keeping 7- and 9-month-old infants still during this 15-min task as compared to the newborn infants tested in Gervain et al. (2008), follow-up studies in older infants should consider a paradigm with fewer blocks of shorter duration in order to equate testing conditions across infants as carefully as possible.

CONCLUSION

In summary, the present work has identified bilateral regions in 7- and 9-month-old infants that show neural attunement to blocks of syllables containing a repetition (ABB) and blocks without a repetition (ABC). While newborn infants (Gervain et al., 2008) and 7-month-old infants showed heightened activity to ABB stimuli, 9-month-old infants showed greater responding to the ABC grammar, perhaps reflecting increased attention to the complexity of the always-changing stimulus. Current work with 3- and 12-month-old infants is underway to more clearly delineate changes in activation across the first year of life to stimuli with and without structural regularities. Specifically, this work will explore the shifting contributions of both oxyHb and deoxyHb in differentiating linguistically relevant information and contribute to our understanding of the factors influencing hemispheric lateralization in response to auditory stimuli. Moreover, this work will allow researchers to gain further insight into the neural bases of developmental shifts in abstract rule learning.

ACKNOWLEDGMENTS

This research was made possible, in part, by grants from the Simons Foundation and the NIDCD (R21 DC 08637 & R01 DC 10290), from the NIMH-funded Clinical Research Training Program at Harvard Medical School (to Jennifer B. Wagner), and from the Hugh Hampton Young Memorial Foundation & the NIH-funded Neuroimaging Training Program (to Sharon E. Fox).

REFERENCES

- Aslin, R. N. (2007). What's in a look? *Dev. Sci.* 10, 48–53.
- Christophe, A., Dupoux, E., Bertoncini, J., and Mehler, J. (1994). Do infants perceive word boundaries? An empirical study of the bootstrapping of lexical acquisition. *J. Acoust. Soc. Am.* 95, 1570–1580.
- Dawson, C., and Gerken, L. (2009). From domain-general to domain-specific: 4-month-olds learn an abstract repetition rule in music that 7-month-olds do not. *Cognition* 111, 378–382.
- Dehaene-Lambertz, G. (2000). Cerebral specialization for speech and non-speech stimuli in infants. *J. Cogn. Neurosci.* 12, 449–460.
- Dehaene-Lambertz, G., Dehaene, S., and Hertz-Pannier, L. (2002). Functional neuroimaging of speech perception in infants. *Science* 298, 2013–2015.
- Ferry, A. L., Hespos, S. J., and Waxman, S. R. (2010). Categorization in 3- and 4-month-old infants: an advantage of words over tones. *Child Dev.* 81, 472–479.
- Frank, M. C., Slemmer, J. A., Marcus, G. F., and Johnson, S. P. (2009). Information from multiple modalities helps 5-month-olds learn abstract rules. *Dev. Sci.* 12, 504–509.
- Gervain, J., Macagno, F., Cogoi, S., Pena, M., and Mehler, J. (2008). The neonate brain detects speech structure. *Proc. Natl. Acad. Sci. U.S.A.* 105, 14222–14227.
- Gervain, J., and Mehler, J. (2010). Speech perception and language acquisition in the first year of life. *Annu. Rev. Psychol.* 61, 191–218.
- Gervain, J., Mehler, J., Werker, J. F., Nelson, C. A., Csibra, G., Lloyd-Fox, S., Shukla, M., and Aslin, R. N. (2011). Near-infrared spectroscopy: a report from the McDonnell infant methodology consortium. *Dev. Cogn. Neurosci.* 1, 22–46.
- Homae, F., Watanabe, H., Nakano, T., Asakawa, K., and Taga, G. (2006). The right hemisphere of sleeping infant perceives sentential prosody. *Neurosci. Res.* 54, 276–280.
- Homae, F., Watanabe, H., Nakano, T., and Taga, G. (2007). Prosodic processing in the developing brain. *Neurosci. Res.* 59, 29–39.
- Johnson, S. P., Fernandes, K. J., Frank, M. C., Kirkham, N., Marcus, G., Rabagliati, H., and Slemmer, J. A. (2009). Abstract rule learning for visual sequences in 8- and 11-month-olds. *Infancy* 14, 2–18.
- Kleinschmidt, A., Obrig, H., Requardt, M., Merboldt, K., Dirnagl, U., Villringer, A., and Frahm, J. (1996). Simultaneous recording of cerebral blood oxygenation changes during human brain activation by magnetic resonance imaging and near-infrared spectroscopy. *J. Cereb. Blood Flow Metab.* 16, 817–826.
- Kuhl, P. K., Stevens, E., Hayashi, A., Deguchi, T., Kiritani, S., and Iverson, P. (2006). Infants show a facilitation effect for native language phonetic perception between 6 and 12 months. *Dev. Sci.* 9, F13–F21.
- Little, V. M., Thomas, D. G., and Letterman, M. R. (1999). Single-trial analyses of developmental trends in infant auditory event-related potentials. *Dev. Neuropsychol.* 16, 455–478.
- Lloyd-Fox, S., Blasi, A., Volein, A., Everdell, N., Elwell, C. E., and Johnson, M. H. (2009). Social perception in infancy: a near infrared spectroscopy study. *Child Dev.* 80, 986–999.
- Marcus, G., Fernandes, K., and Johnson, S. (2007). Infant rule-learning facilitated by speech. *Psychol. Sci.* 18, 387–391.
- Marcus, G., Vijayan, S., Bandi Rao, S., and Vishton, P. M. (1999). Rule-learning in seven-month-old infants. *Science* 283, 77–80.
- Mehler, J., Jusczyk, P., Lambertz, G., Halsted, N., Bertoncini, J., and Amiel-Tison, C. (1988). A precursor of language acquisition in young infants. *Cognition* 29, 143–178.
- Minagawa-Kawai, Y., Mori, K., Naoi, N., and Kojima, S. (2007). Neural attunement processes in infants during the acquisition of a language-specific phonemic contrast. *J. Neurosci.* 27, 315–321.
- Obrig, H., Rossi, S., Telkemeyer, S., and Wartenburger, I. (2010). From acoustic segmentation to language processing: evidence from optical imaging. *Front. Neuroenergetics* 2:13. doi: 10.3389/fnene.2010.00013
- Pena, M., Maki, A., Kovacic, D., Dehaene-Lambertz, G., Koizumi, H., Bouquet, F., and Mehler, J. (2003). Sounds and silence: an optical topography study of language recognition at birth. *Proc. Natl. Acad. Sci. U.S.A.* 100, 11702–11705.
- Saffran, J. R., Aslin, R. N., and Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science* 274, 1926–1928.
- Saffran, J. R., Pollack, S. D., Seibel, R. L., and Shkolnik, A. (2007). Dog is a dog is a dog: infant rule learning is not specific to language. *Cognition* 105, 669–680.
- Saito, Y., Aoyama, S., Kondo, T., Fukumoto, R., Konishi, N., Nakamura, K., Kobayashi, M., and Toshima, T. (2007a). Frontal cerebral blood flow change associated with infant-directed speech. *Arch. Dis. Child. Fetal Neonatal Ed.* 92, F113–F116.
- Saito, Y., Kondo, T., Aoyama, S., Fukumoto, R., Konishi, N., Nakamura, K., Kobayashi, M., and Toshima, T. (2007b). The function of the frontal lobe in neonates for response to a prosodic voice. *Early Hum. Dev.* 83, 225–230.
- Sato, Y., Sogabe, Y., and Mazuka, R. (2009). Development of hemispheric specialization for lexical pitch-accent in Japanese infants. *J. Cogn. Neurosci.* 22, 2503–2513.
- Steinbrink, J., Villringer, A., Kempf, F., Haux, D., Boden, S., and Obrig, H. (2006). Illuminating the BOLD signal: combined fMRI-fNIRS studies. *Magn. Reson. Imaging* 24, 495–505.
- Teinonen, T., Fellman, V., Naatanen, R., Alku, P., and Huotilainen, M. (2009). Statistical language learning in neonates revealed by event-related brain potentials. *BMC Neurosci.* 10, 21. doi: 10.1186/1471-2202-10-21
- Telkemeyer, S., Rossi, S., Koch, S. P., Nierhaus, T., Steinbrink, J., Poeppel, D., Obrig, H., and Wartenburger, I. (2009). Sensitivity of newborn auditory cortex to the temporal structure of sounds. *J. Neurosci.* 29, 14726–14733.
- Thomas, D. G., and Lykins, M. S. (1995). Event-related potential measures of 24-hour retention in 5-month-old infants. *Dev. Psychol.* 31, 946–957.
- Turk-Browne, N. B., Scholl, B. J., and Chun, M. M. (2008). Babies and brains: habituation in infant cognition and functional neuroimaging. *Front. Hum. Neurosci.* 2:16. doi: 10.3389/fnhum.09.016.2008
- Werker, J. F. (1989). Becoming a native listener. *Am. Sci.* 77, 54–69.
- Werker, J. F., and Tees, R. C. (1984). Cross-language speech perception: evidence for perceptual reorganization during the first year of life. *Infant Behav. Dev.* 7, 49–63.
- Wilcox, T., Bortfeld, H., Woods, R., Wruck, E., and Boas, D. A. (2005). Using near-infrared spectroscopy to assess neural activation during object processing in infants. *J. Biomed. Opt.* 10, 11010–11026.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 15 January 2011; accepted: 05 July 2011; published online: 25 July 2011.
Citation: Wagner JB, Fox SE, Tager-Flusberg H and Nelson CA (2011) Neural processing of repetition and non-repetition grammars in 7- and 9-month-old infants. *Front. Psychology* 2:168. doi: 10.3389/fpsyg.2011.00168
This article was submitted to *Frontiers in Language Sciences*, a specialty of *Frontiers in Psychology*.
Copyright © 2011 Wagner, Fox, Tager-Flusberg and Nelson. This is an open-access article subject to a non-exclusive license between the authors and Frontiers Media SA, which permits use, distribution and reproduction in other forums, provided the original authors and source are credited and other Frontiers conditions are complied with.

APPENDIX

Statistical maps (t -maps) comparing oxyHb and deoxyHb responses to the repetition (ABB) and non-repetition (ABC) grammars for 7-month-olds (**Figure A1**) and 9-month-olds (**Figure A2**). For ABB vs. ABC in oxyHb t -tests (left side of **Figures A1 and A2**),

warmer colors (reds) signify greater activation to ABB grammar as compared to ABC grammar; for deoxyHb t -tests (right side of **Figures A1 and A2**), colder colors (blues) signify greater activation to ABB as compared to ABC. No channels showed significant differences between the two grammars (all $ps > 0.05$).

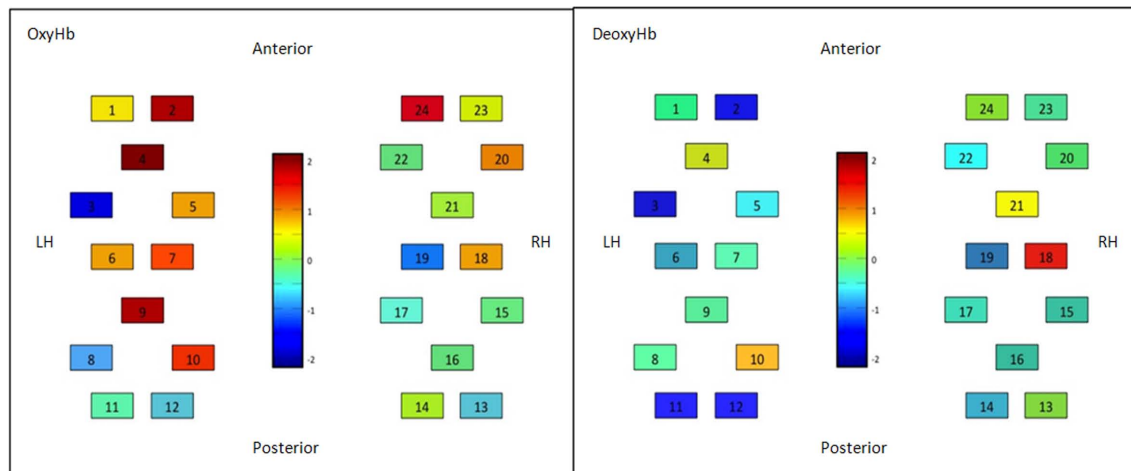


FIGURE A1 | Channel-by-channel t -tests for 7-month-old infants in the present study.

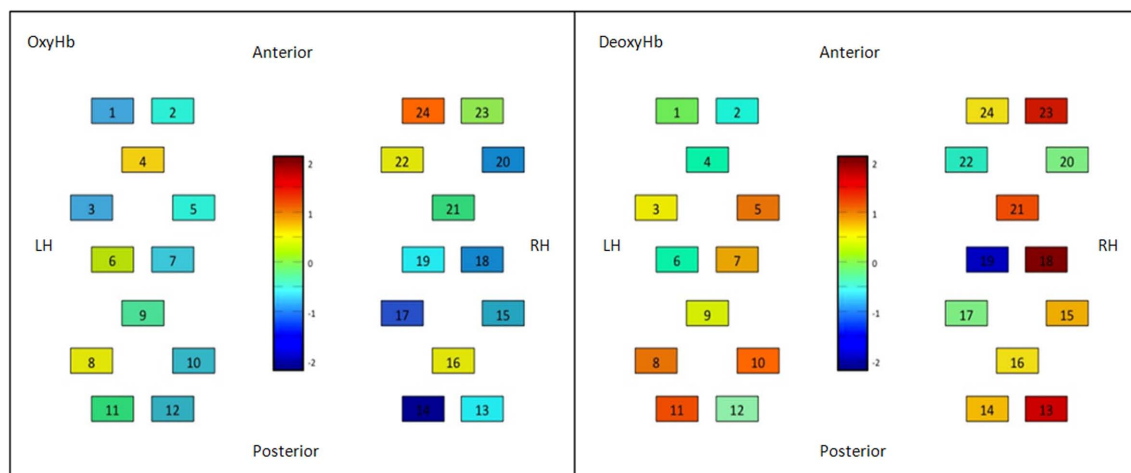


FIGURE A2 | Channel-by-channel t -tests for 9-month-old infants in the present study.



The role of orbitofrontal cortex in processing empathy stories in 4- to 8-year-old children

Tila Tabea Brink^{1,2*}, Karolina Urton¹, Dada Held¹, Evgeniya Kirilina^{1,3}, Markus J. Hofmann^{1,2}, Gisela Klann-Delius^{1,4}, Arthur M. Jacobs^{1,2,3} and Lars Kuchinke^{1,5*}

¹ The Cluster of Excellence "Languages of Emotion," Freie Universität Berlin, Berlin, Germany

² Neurocognitive Psychology, Freie Universität Berlin, Berlin, Germany

³ Dahlem Institute for Neuroimaging of Emotion, Freie Universität Berlin, Berlin, Germany

⁴ Department of Linguistics, Institut für Deutsche und Niederländische Philologie, Freie Universität Berlin, Berlin, Germany

⁵ Department of Psychology, Ruhr-Universität Bochum, Bochum, Germany

Edited by:

Judit Gervain, CNRS – Université Paris Descartes, France

Reviewed by:

Kalina J. Michalska, The University of Chicago, USA

Sarah Lloyd-Fox, Birkbeck, University of London, UK

*Correspondence:

Tila Tabea Brink, General and Neurocognitive Psychology, Freie Universität Berlin, Habelschwerdter Allee 45, 14195 Berlin, Germany.
e-mail: tila.brink@fu-berlin.de;

Lars Kuchinke, Experimental Psychology, Ruhr-Universität Bochum, Universitätsstraße 150, 44801 Bochum, Germany.
e-mail: lars.kuchinke@rub.de

This study investigates the neuronal correlates of empathic processing in children aged 4–8 years, an age range discussed to be crucial for the development of empathy. Empathy, defined as the ability to understand and share another person's inner life, consists of two components: affective (emotion-sharing) and cognitive empathy (Theory of Mind). We examined the hemodynamic responses of preschool and school children ($N = 48$), while they processed verbal (auditory) and non-verbal (cartoons) empathy stories in a passive following paradigm, using functional Near-Infrared Spectroscopy. To control for the two types of empathy, children were presented blocks of stories eliciting either affective or cognitive empathy, or neutral scenes which relied on the understanding of physical causalities. By contrasting the activations of the younger and older children, we expected to observe developmental changes in brain activations when children process stories eliciting empathy in either stimulus modality toward a greater involvement of anterior frontal brain regions. Our results indicate that children's processing of stories eliciting affective and cognitive empathy is associated with medial and bilateral orbitofrontal cortex (OFC) activation. In contrast to what is known from studies using adult participants, no additional recruitment of posterior brain regions was observed, often associated with the processing of stories eliciting empathy. Developmental changes were found only for stories eliciting affective empathy with increased activation, in older children, in medial OFC, left inferior frontal gyrus, and the left dorsolateral prefrontal cortex. Activations for the two modalities differ only little, with non-verbal presentation of the stimuli having a greater impact on empathy processing in children, showing more similarities to adult processing than the verbal one. This might be caused by the fact that non-verbal processing develops earlier in life and is more familiar.

Keywords: OFC, cognitive empathy, affective empathy, children, fNIRS, verbal, non-verbal

INTRODUCTION

As social interaction is central to the life of human beings, paying attention to and trying to understand the cognitive and affective processes of others is important for the prediction and interpretation of their behavior. These skills are studied under the label of social cognition. Studies concerning the neural basis of social cognition have mainly focused on adults (Amodio and Frith, 2006). Thus, relatively little is known about the neural processing of socio-emotional information in children, although it is well known that changes in many domains of cognition occur with development (Durstun and Casey, 2006). For example, the development of higher-level cognitive processes is discussed to covary with maturation of the prefrontal cortex throughout childhood and adolescence (Casey, 1999). In particular, the most anterior part of the prefrontal cortex, the orbitofrontal cortex (OFC), is supposed to support the processing of social cognition, while it is known to mature last in ontogeny (Barbey et al., 2009).

Empathy, the ability to understand and share another person's inner life, is an essential process in social cognition. It is a complex form of psychological inference, in which observation, memory, knowledge, and reasoning as well as affective sharing are combined (Ickes, 1997). Empathy describes both, sharing as well as understanding the emotional state of others in relation to oneself (Decety et al., 2008). Thus, previous research focused on two main approaches to empathy:

- (1) The ability to understand the intentions of another person, so-called *cognitive empathy* (Baron-Cohen and Wheelwright, 2004; Shamay-Tsoory et al., 2008), sometimes also termed perspective-taking (Lamm et al., 2007), "Theory of Mind" (ToM; Wimmer and Perner, 1983), or mentalizing (Frith and Frith, 2006a, 2006b).
- (2) In contrast, *affective empathy* focuses on the affective response to another person's presumed affective state (Eisenberg and Miller, 1987) and is sometimes also called affective sharing

(Decety and Jackson, 2006). Although empathy and affective empathy are sometimes used synonymously in literature, we will use the term “affective empathy” in the following to distinguish it from “cognitive empathy.”

Cognitive empathy is defined as the ability to imagine or “experience” a situation from another person’s point of view. At the age of two, typically developing children understand another person’s intentions, independently from their own intentions (Leslie, 1987; Flavell, 1999). Understanding another person’s beliefs, for example that a person is thinking wrongly about something, develops at the age of about four (see Wellman et al., 2001, for an overview). Children between two and a half and almost 4 years make the so-called “false belief”-mistake (Baron-Cohen et al., 1985). They are unaware of the fact that their knowledge is different from another person’s knowledge and do not yet understand that different people, depending on their perspective, can have different thoughts about the same situation (“first-order false belief”). With the age of four, most children solve the “false belief” task successfully and 5- to 6-year-olds are able to give correct answers in 90% of the cases (Baron-Cohen et al., 1985; Perner et al., 1987), and solve even higher-order abstraction tasks (Perner and Wimmer, 1985).

Affective empathy is defined as an affective response, derived from the apprehension and comprehension of another person’s affective state, which is identical or very similar to what the other person is expected to feel (Eisenberg, 2000). To empathize with another person does not only mean understanding *why* the other person is happy or sad, but also being able to *feel* with her or him, i.e., to mentally “simulate” the other person’s feelings. Affective empathy develops very early in life, and the mechanism underlying affective sharing is discussed to be present from birth on (Decety and Meyer, 2008). The earliest form of empathy is “reactive crying” (emotional contagion) in newborns, presumably lacking any cognitive component. Hamlin et al. (2007) suggested that already at about 6 months of age infants engage in rudimentary forms of social evaluation and preferentially interact with an agent who helped rather than hindered the actions of another character. Altruistic helping as a form of pro-social behavior also emerges early in childhood. Behavioral studies demonstrate that by 12 months of age infants begin to comfort victims of distress, and 14- to 18-month-old infants exhibit spontaneous, unrewarded instrumental helping behaviors (Warneken and Tomasello, 2009). These naturally emerging behaviors are thought to be motivated by sympathetic emotion or concern for others’ well-being. Affective empathy develops and is at its highest level when a person is able to empathize with another person’s experiences and feelings beyond the immediate situation (Hoffman, 2000).

At present, an important argument for distinguishing the two different approaches are the different roles affective and cognitive empathy play in psychiatric disorders like autism and psychopathy: While patients with autism spectrum disorder often show impairment in ToM (Baron-Cohen et al., 1985), affective empathy may be preserved (Blair, 2005; Dziobek et al., 2008). In contrast, psychopaths show a great lack of affective, but mostly no impairment in cognitive empathy (e.g., Baron-Cohen et al., 1985; Soderstrom, 2003). These findings support the assumption that

the two concepts, besides sharing similar features and common neural networks, may have specific neuronal correlates, distinct from each other.

Most imaging studies on cognitive empathy have examined mentalizing-tasks that did not include any affective component (Decety and Jackson, 2004). Only few studies have been carried out linking the two components by directly contrasting cognitive and affective empathy tasks in adults (e.g., Hynes et al., 2006; Völlm et al., 2006; Shamay-Tsoory et al., 2008), while studies in children are lacking. The relationship of affective and cognitive empathy has yet to be further determined, especially referring to their neuronal correlates in children.

A neural network often described in empathy research is the frontal mirror neuron system (MNS), including the pars opercularis of the inferior frontal gyrus (IFG; e.g., Singer, 2006; Pfeifer et al., 2008; Hooker et al., 2009). The main function associated with the MNS is that of a simulation mechanism: Perceiving the actions of another person elicits activity in neurons that are also active when we perform those actions ourselves (Gallese and Goldman, 1998), which makes it a suitable mechanism underlying both, affective and cognitive empathic processes (Völlm et al., 2006).

Based on studies on autism and psychopathy, some researchers still argue for a dissociation and define empathy as a general term for a collection of specific neurocognitive functions (Blair, 2005, 2008). The temporoparietal junction (Frith and Frith, 2006a; Decety and Lamm, 2007; Decety et al., 2008; Hooker et al., 2009), for example, is discussed to only support the processing of cognitive empathy and according to a recent meta-analysis on the neural correlates of cognitive empathy (Carrington and Bailey, 2009), OFC, temporoparietal junction and the superior temporal sulcus (Hein and Singer, 2008) are commonly found to be activated by cognitive empathy tasks. Somewhat surprisingly, the same study could not identify a single region that was consistently activated across all analyzed studies (Carrington and Bailey, 2009). This could partly be due to the diversity of the paradigms, but is also evidence for the complexity of social cognition. Moreover, orbitofrontal regions have been found active in both, affective (Hynes et al., 2006; Decety and Meyer, 2008; Decety et al., 2008) and cognitive empathy tasks (Carrington and Bailey, 2009). OFC functioning is known to be critical for social cognition processes, moral decisions, and emotion control.

In the present study we investigate the contribution of the OFC to empathic processing, both, affective and cognitive, in young children.

There is a great lack of developmental studies on empathy and a need for investigating the different levels of empathic responding in further detail, in particular with respect to their underlying neural networks. A neurodevelopmental approach on cognitive and affective empathy may therefore help to better dissociate the two mechanisms (Singer, 2006). A recent functional magnetic resonance imaging (fMRI) study examined the neural correlates of ToM judgments in 8- to 12-year-old children and in adults, using verbal and non-verbal cartoons. The results suggest that children and adults both activate the temporoparietal junction and the IFG for ToM judgments, but differ in their activation patterns depending on task modality, e.g., children have higher activation in left IFG when processing non-verbal cartoons, whereas adults show higher activation in left IFG when processing verbal stimuli. Kobayashi

et al. (2007) suggest that children adopt different strategies, and that the observed interactions with age may be linked to age-related refinement of the inferior frontal and posterior temporal regions. A further study by Pfeifer et al. (2008) focused on the role of simulation in affective empathy in 10-year-old children. Their findings show that activity in the posterior IFG (pars opercularis) correlated significantly and positively with empathy and social skills.

The present study examines the processing of affective and cognitive empathy in children between 4 and 8 years of age in a passive empathy recognition paradigm. Most studies on empathy have examined adults and those studies that investigated empathy processing in children examined samples aged seven (Decety et al., 2008; Decety and Michalska, 2010) and older (Kobayashi et al., 2007; Pfeifer et al., 2008). Morphometric studies have demonstrated that structural brain changes in gray and white matter are present at preschool age (Giedd et al., 1999). Similarly, the first and main developmental steps of cognitive empathy are completed about the age of four and affective empathy and forms of pro-social behavior are present from early childhood on.

Empathy development and its neural correlates in preschool children are poorly understood. Therefore, the age range examined in the current study includes preschoolers as well as young school children. One reason for the lack of developmental studies on neural correlates of affective and cognitive empathy is that fMRI studies with children are relatively rare. It is still a challenge to perform experiments with small children in the noisy and unfamiliar environment of an fMRI scanner. Another problem is the high level of motion artifacts, occurring during the measurement of young children (Karmiloff-Smith, 2010).

To measure the neural correlates of empathy processing in children, we used functional near-infrared spectroscopy (fNIRS). This is a non-invasive, functional optical imaging method, assessing changes in cortical oxygenation by applying near-infrared light to measure changes in tissue attenuation. It monitors the brain function by measuring changes in the concentrations of oxygenated [oxy-Hb] and deoxygenated hemoglobin [deoxy-Hb] and is based on the fact that hemoglobin changes its color when the oxygen content changes (Obrig and Villringer, 2003; Hofmann et al., 2008). In particular, brain activation is indicated by an increase in [oxy-Hb] and a decrease in [deoxy-Hb], and it is shown that the latter is highly correlated with an increase in fMRI blood oxygen level dependent (BOLD) response (Steinbrink et al., 2006). Near-infrared light is emitted into the cortex by light-sources and the re-emitted light is collected by another set of optic probes (so-called detectors) at a distance of 3 cm, making it possible to detect activated brain regions, which are approximately 1.5 cm under the cranium. NIRS is a high-potential tool in developmental research due to high light penetration depth in children's brains. fNIRS lowers the sensitivity to motion artifacts and is a less stressful procedure than fMRI (Lloyd-Fox et al., 2010).

There are no studies on empathy using fNIRS so far and generally imaging studies on empathy in children are rare. Therefore, a first aim of this study is to introduce fNIRS to the field of neurodevelopmental research on empathy processing. Functional brain responses in orbitofrontal and posterior temporal brain regions are investigated under affective and cognitive empathy conditions to reveal similarities and differences in the underlying cortical networks. Furthermore,

empathic responses in young children were examined in two stimulus modalities, a non-verbal set of cartoon stories and a set of verbal, auditorily presented stories. The aim was to control for modality-specific effects which allows us to identify modality-independent and -specific activations. Thus, we expected that brain regions that are activated in a modality-independent manner, like the posterior temporal regions reported by Kobayashi et al. (2007) for visual processing of ToM judgments, could also be identified across a visual and an auditory recording session related either to affective or cognitive empathy. Finding brain regions that are activated under both empathy conditions and across the different sensory channels would argue for a more general, supervisory role of these regions in empathy.

As most studies investigated adults, we had to derive our hypotheses mainly from those results, tentatively assuming analogous processing in children. The OFC supports affective as well as cognitive empathy, thus we expect specific activations in this region in both conditions. Hynes et al. (2006) already investigated the adult OFC's role in affective versus cognitive perspective-taking by means of fMRI. They observed neural activations in both conditions in bilateral OFC, but far more for the emotional than for the cognitive condition, especially concerning the medial OFC. Thus, we expect to observe neural activations for cognitive empathy in lateral and for affective empathy in medial OFC.

According to ToM developmental research, older children have developed higher empathy skills and should thus show different activation-patterns than younger subjects. As affective empathy starts to develop very early but seems to change in quality with cognitive development, one would expect age-related differences in both, affective and cognitive empathy responses. In particular, Decety and colleagues (Decety and Michalska, 2010; Decety, 2011) hypothesize that a higher involvement of executive functioning in empathy processing develops (relatively slowly) in parallel to brain maturation, pointing to a role of frontal regions in modulating empathetic responses (see also Hoffman, 2000). Focusing on a relatively young sample of participants provides the opportunity to investigate developmental differences in children's empathy after the age of four, when basic abilities are present but still develop further. Still, because of sparse experimental results, we are reluctant to hypothesize on age differences in more detail.

MATERIALS AND METHODS

SAMPLE

Forty-eight participants (22 male/26 female) aged between 4;0 and 8;8 (mean age 6;2) were recruited from daycare centers and primary schools in Berlin. To examine developmental changes, two age groups were formed relative to the age of 6;6. Because in the German educational system children start attending primary school at the age of 6, and one can assume that formal education has an important impact on social and cognitive development, the younger group mainly consisted of preschool children aged 4;0–6;6 ($N = 24$; mean age = 5;0) and the older group consisted of school children aged 6;6 and older ($N = 24$; mean age = 7;6).

All participants were native German speakers, never had injuries or operations on the brain, were not taking any medication and did not show behavioral or neurological conspicuities. Their parents received an information-brochure and obtained 20 Euros for participation. Both, children and parents were informed about

the background of the study and its procedure and gave their agreement. Parents attended the whole session and could observe it via video from a second laboratory room. The recruitment and experimental procedure was performed in accordance with the local ethical guidelines and consent and assent were obtained from the parent(s) or legal guardian(s) and the child.

A short form of the Kaufman assessment battery for children (K-ABC; Kaufman and Applegate, 1988) was used to make sure none of the participants had an estimated IQ lower than 85, which is 1 SD below the mean. The mean score was 109.38 (range: 85–134.10, SD = 11.90) for the mental processing composite (MPC) and 111.42 (range: 88.40–132.35, SD = 8.45) for the Achievement Scales (ACH). The two age groups did not differ significantly in K-ABC scores (young children: mean MPC score = 107.52, SD = 12.34; mean ACH score = 113.38, SD = 8.78; older children: mean MPC score = 111.23, SD = 11.40; mean ACH score = 111.42, SD = 8.45), as proven by a *t*-test [MPC: $t(46) = -1.038$, $p = 0.285$; ACH: $T(46) = 1.631$, $p = 0.110$].

MATERIAL

The material consisted of two sets of stimuli, a visual set of non-verbal cartoon stories and an auditory set of verbal listening stories. Every cartoon consisted of four pictures and every story consisted of four sentences, which were presented in a fixed order (see **Figure 2**). Four conditions were presented during the study:

1. Affective empathy (two characters)
2. Cognitive empathy (one character)
3. Neutral stories with one character
4. Neutral stories with two characters

The material is based on a study by Völlm et al. (2006). A professional illustrator was engaged to adapt these cartoons for the current study. Additionally, new stories were developed. The cartoon stories were drawn in black and white. Only the main character was wearing something colored to direct the attention to her or him (see **Figures 1A–D**, see also Appendix).

The verbal listening stories were derived from the cartoon stories by describing each picture by one sentence (**Figure 1E**). These stories were then read in by an actress. In analogy to Völlm et al. (2006), all stories were constructed in two versions, which differed only in the last picture/sentence. The affective empathy stories had either a positive ending or a negative one. The cognitive empathy stories ended logically or illogically (i.e., the behavior of the main character was reasonable or not). All neutral stories relied on the understanding of physical causalities and were shown in correct or scrambled order, the fourth picture being necessary to recognize the difference. Every story was only seen and heard in one version by each subject.

The final material was derived from a larger set of stories by means of a rating study. Sixteen to 24 adults rated all cartoon stories online, evaluating how strongly the course of the story elicits affective empathy on a scale from 1 (not at all) to 5 (very strongly), to control for the affective empathy elicited by these stories.

The final stimulus set consisted of 48 stories in each stimulus modality, 12 of each experimental condition, half of them positive, logical, or in correct order and half negative, not logical, or

scrambled. For the cartoon condition, an ANOVA revealed a significant main effect of affective empathy ratings [$F(3,44) = 25.206$, $p < 0.001$]. Pairwise *post hoc* comparisons using Scheffé's test revealed that the affective empathy cartoons ($M = 3.81$) differed from all other cartoon conditions (cognitive empathy: $M = 2.60$; neutral stories with one character: $M = 2.47$; neutral stories with two characters: $M = 1.97$). Similarly, an ANOVA for the verbal listening stories revealed a main effect of affective empathy [$F(3,44) = 34.917$, $p < 0.001$], again driven by higher ratings for affective empathy stories ($M = 3.78$) as compared to cognitive empathy ($M = 2.00$), neutral with one character ($M = 2.01$), and neutral with two characters ($M = 1.86$). The length of the spoken sentences and the average number of words of the verbal stories were balanced across conditions (mean number of words: 37.3, range 21–50; mean length: 11.64 s).

EXPERIMENTAL PROCEDURE

The stimuli were presented on a standard 17" PC screen, which was placed 60 cm in front of the participants. The experiment started with an instruction to listen to in the verbal session or watch the stories carefully in the non-verbal session, respectively, followed by an example story introducing each block of one of the four experimental conditions. Verbal listening stories were presented via earphones. Stimulus presentation and timing in both modalities were controlled using OGAMA (version 2.0; Voßkübler et al., 2008).

After clicking the mouse, a colored screen appeared in both modalities, and after a second mouse click, an example story started. Following the example, a question was asked ("Did you understand?"). This allowed time for more detailed instructions and answering questions. When all questions were answered to the satisfaction of the child, clicking the mouse once again started the actual setting.

For the non-verbal modality, each trial consisted of four pictures, each presented for 1.5 s, so that one cartoon story lasted 6 s. After the fourth picture, a fixation cross (+) appeared in the middle of the screen for 4 s, and the participants were asked to simply click the mouse. This mouse clicking after each story did not have an influence on the speed of presentation, but was used to keep the children concentrated.

The four blocks (affective, cognitive, neutral with one character, and neutral with two characters), each containing 12 stories, were presented in random order. Inside of each block the two types of stories were presented in equal number and pseudorandomized order: an affective empathy block contained six stories with positive and six stories with negative ending; a cognitive empathy block contained six stories with logical and six stories with non-logical ending, and in neutral blocks six stories were presented in correct and six in scrambled order. Between the blocks a colored screen appeared, so that the participants could take a break before continuing by clicking.

The verbal session worked exactly the same way, except that a fixation cross was displayed while the participants listened to the four sentences. The average length of one sentence was 2.9 s. The two runs (two modalities) were presented one after the other, with a short break in between. Half of the subjects started with the cartoons, the other half with the listening stories.

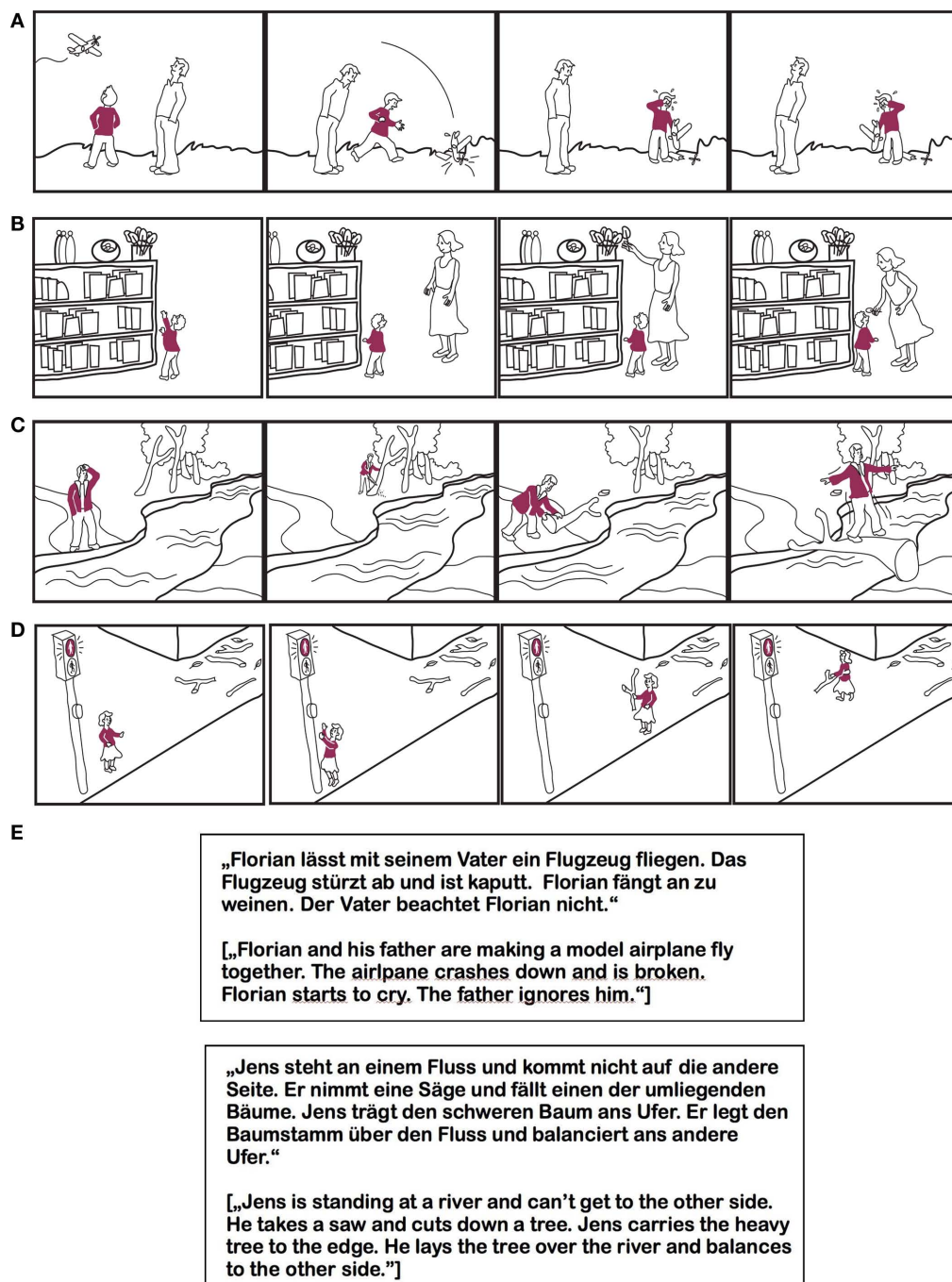


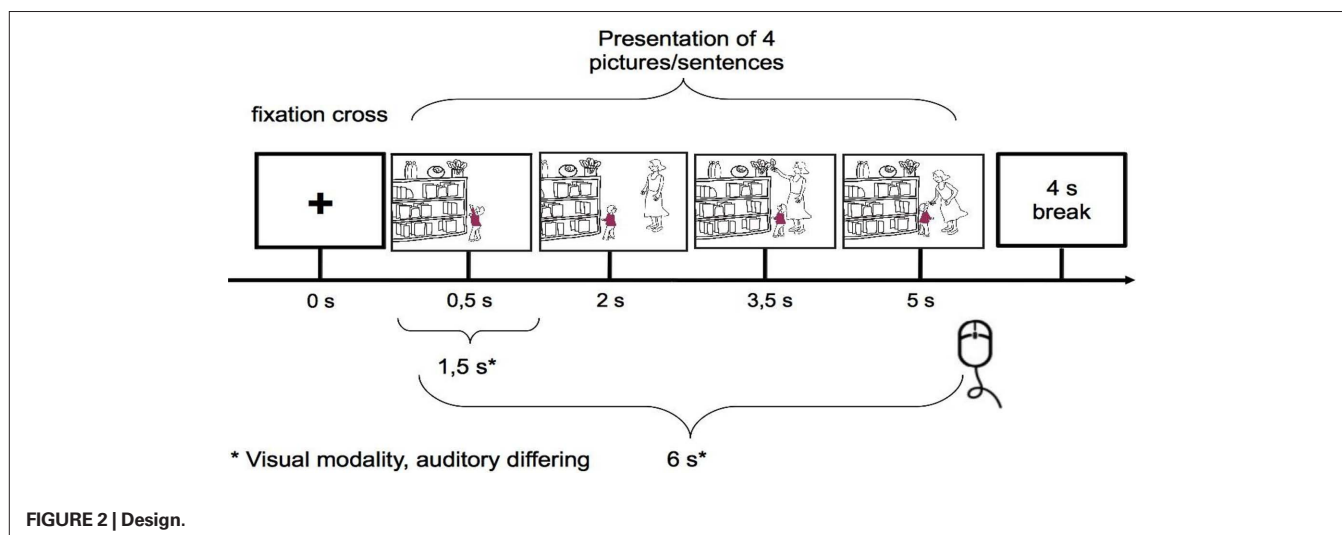
FIGURE 1 | (A) Example of a cartoon stimulus from the negative affective empathy condition. (B) Example of a cartoon stimulus from the positive affective empathy condition. (C) Example of a cartoon stimulus from the logical cognitive

empathy condition. (D) Example of a cartoon stimulus from the non-logical cognitive empathy condition. (E) Examples of auditory stimuli: negative affective and logical cognitive story, in German and English.

Every cartoon story took 6 s plus break, thus every block took 2 min (8 min for all four blocks) plus breaks and examples for the visual trials; the auditory blocks took 3.13 min on average each, 12.51 min altogether plus breaks and examples.

Throughout the whole measuring procedure, a supervisor stayed in the room, read instructions to the children and answered questions.

Parallel to the experimental sessions, all parents answered a German version of the “Griffith empathy measure (GEM)” scale (Dadds et al., 2008), a brief parent-report measure of child empathy. Overall, the children varied widely in their resulting empathy scores with a GEM total of 15.54 (SD = 18.24). For the affective empathy sub-scale, the mean score was 3.56 (SD = 7.24), for the cognitive empathy sub-scale it was 3.91 (SD = 9.14). A



comparison of the younger and older children did not reveal significant differences [GEM total: $T(46) = 0.564$, $p = 0.575$; GEM affective $T(46) = 0.500$, $p = 0.620$; GEM cognitive: $T(46) = 0.413$, $p = 0.681$].

NIRS DATA ACQUISITION

The fNIRS measurements were performed by a DYNOT System (32 sources/32 detectors, NIRx Medizintechnik, GmbH, Berlin, Germany) operating at two wavelengths (760 and 830 nm) at a sampling rate of 4.13 Hz. An optode-set of 11 sources and 21 detectors was used, making 39 source detector pairs (channels; see Figure 3B). The output ends of the sources and the input ends of the detector fibers were inserted into the electrode holes of a 52-cm Easycap (M16, equidistant system; Falk Minow Services), which was placed on the participants' heads. The optodes were placed on the orbitofrontal and temporal regions of the head with 3 cm averaged source detector-distance (equidistant system). To capture the orbitofrontal regions, the Easycap was turned by 180°, so that optodes could be placed on the forehead, positioned more orbital than usual. Synchronization with the experimental procedures was provided by marker signals (TTL), sent via the parallel port of the stimuli presenting computer, using OGAMA software.

In order to avoid time-consuming optode positioning procedures, the Easycap was prepared in advance by placing the optodes on it. Some sources and detectors had to be optimized by fixing hair with a small amount of gel (EASYPAC Supervisc, high-viscosity electrolyte-gel) to make sure no hair absorbed the light.

DATA ANALYSIS

Preprocessing

The NILAB toolbox (Koch et al., 2009) and Matlab (The MathWorks) were used for fNIRS-data analysis. The time courses of detected light intensities at both wavelengths were transformed into time courses of concentration changes of oxygenated [oxy-Hb] and deoxygenated [deoxy-Hb] hemoglobin by using the modified Beer Lambert law (Cope and Delpy, 1988). We used the following extinction coefficients: 830 nm oxy-Hb

2.3214 $\text{mM}^{-1}\text{cm}^{-1}$, deoxy-Hb 1.4866 $\text{mM}^{-1}\text{cm}^{-1}$, 760 nm oxy-Hb 1.7917 $\text{mM}^{-1}\text{cm}^{-1}$, deoxy-Hb 3.8437 $\text{mM}^{-1}\text{cm}^{-1}$ (using data from: W. B. Gratzer, Med. Res. Council Labs, Holly Hill, London), and a photon differential path-length factor of 5.98 cm for 830 nm and 7.15 for 760 nm. The obtained time series of [oxy-Hb] and [deoxy-Hb] concentration changes were low-pass filtered with a cut-off frequency of 0.4 Hz and visually inspected to correct for artifacts related to the subjects' motion. The trial was excluded from further analysis, if simultaneous step-like changes were observed for both, [oxy-Hb] and [deoxy-Hb], in at least two neighboring channels. Due to this procedure 21.03% (24.67/16.57% in younger/older group) of all data in the visual and 21.19% (21.57/20.78% in younger/older group) in the auditory condition were discarded.

The experimental time courses were block-averaged over six repetitions of every condition in the interval between -10 and $+20$ s around the stimulus onset of the last sentence (to cover the length of the hemodynamic responses to each presented story as a whole) and detrended to correct for linear baseline drift.

General linear modeling

Preprocessed data were subjected to a general linear model (GLM) analysis. Each of the four conditions – cognitive empathy, affective empathy, neutral story with one character, neutral story with two characters – was modeled with two predictors: The first predictor of each condition modeled cerebral activation at the onsets of the first three sentences/pictures of the story, the second predictor of each condition modeled the onsets of the last sentence/picture. Delta functions of the sentence/picture onsets were convolved with a hemodynamic response function (sum of two gamma functions with time constants of 5 and 16 s and with weights of 1 and $-1/6$). The predictors were subjected to the same block average procedure as the experimental time courses, accounting for blocks dismissed due to motion correction. An example of block-averaged data together with predictor time course is shown in Figure 4. Since each of the four conditions was presented in non-overlapping blocks, they were analyzed separately. Contrasts of interest were computed from the estimated beta values of the predictor of the last

sentence/picture and subjected to a second level analysis. Results for non-verbal cartoon stories and verbal stories and for both hemoglobin species are reported.

Optode locations as a basis for neuroanatomical labeling of the NIRS channels were obtained by means of an anatomical MRI template. Since the brain volume and geometry changes only slightly after the age of two (Faria et al., 2010), we acquired an anatomical T1 scan of an adult subject with a head circumference of 52 cm, the same circumference our sample had on average, wearing the Easycap (3T, SIEMENS-Trio, MPRAGE-Sequence, resolution 1 mm × 1 mm × 1 mm). The fiducial markers were placed in all optode positions and reference Cz. The anatomical volume was normalized to standard Talairach 3D space using Brain Voyager QX (v 1.7; Brain Innovation, Maastricht, The Netherlands). The channel-positions were defined, computing the exact midway between source and detector and the most probable anatomical label was obtained using the NFRI-toolbox (Okamoto et al., 2004; Singh et al., 2005).

Channels were clustered to define regions of interest, based on our initial hypotheses. One cluster comprised the average signal of three defining channels (see Figure 3; Table 1). Five clusters were defined: medial OFC, left OFC, right OFC, left PTR, and right PTR. In addition to this clustered region-of-interest analysis, significant activations of single channels which exceed a conservative threshold of $p < 0.01$ are reported, too. The positions of the channels and clusters can be seen in Figure 3.

For each stimulus modality one-sample t -tests (two-tailed) were computed at the second level for the following three main contrasts:

- (1) *Affective empathy* > Neutral (two characters): subtracting activations during the processing of neutral stimuli from those during processing affective empathy stimuli to reveal hemodynamic responses related to affective empathy processing only (Note that both conditions contained stories with two characters).
- (2) *Cognitive empathy* > Neutral (one character): subtracting activations during the processing of neutral stimuli from those during processing cognitive empathy stimuli to reveal hemodynamic responses related to cognitive empathy only.

(3) [*Affective empathy* – Neutral (two)] > [*Cognitive empathy* – Neutral (one)]: to directly contrast affective and cognitive empathy processing, contrast 1 and 2 are subtracted. As a result, activation increases are related to affective empathy processing, whereas decreases are related to cognitive empathy processing.

To detect effects of age, paired t -tests comprising “age group” as factor were computed. In addition, to identify whether or not neural activity in our *a priori* regions of interest correlated with the parent-reported empathy scores, a correlational analysis across all subjects was performed between the activity in the three contrasts and the affective and cognitive GEM sub-scores. For all analyses, significant results are reported that exceed a conservative threshold of $p < 0.01$ (uncorrected).

RESULTS

Due to unusually high residual variance in some subjects, data for seven cartoon and four listening sessions had to be excluded from further analysis. Cartoon ($N = 40$) and auditory listening story ($N = 44$) data were analyzed separately. Significant changes in oxygenated [oxy-Hb] and deoxygenated blood [deoxy-Hb] during the trials are given in Table 2 for visual and Table 3 for auditory presentation.

Table 1 | Cluster definitions.

Cluster	Channels
ORBITOFRONTAL CORTEX (OFC) COVERING BA10, BA11	
Medial OFC	Ch1, ch9, ch10
Left OFC	Ch2, ch7, ch8
Right OFC	Ch5, Ch11, Ch12
POSTERIOR TEMPORAL REGION (PTR) INCLUDING SUPERIOR TEMPORAL SULCUS AND TEMPOROPARIETAL JUNCTION	
Left PTR	Ch30, ch31, ch37
Right PTR	Ch33, ch35, ch38

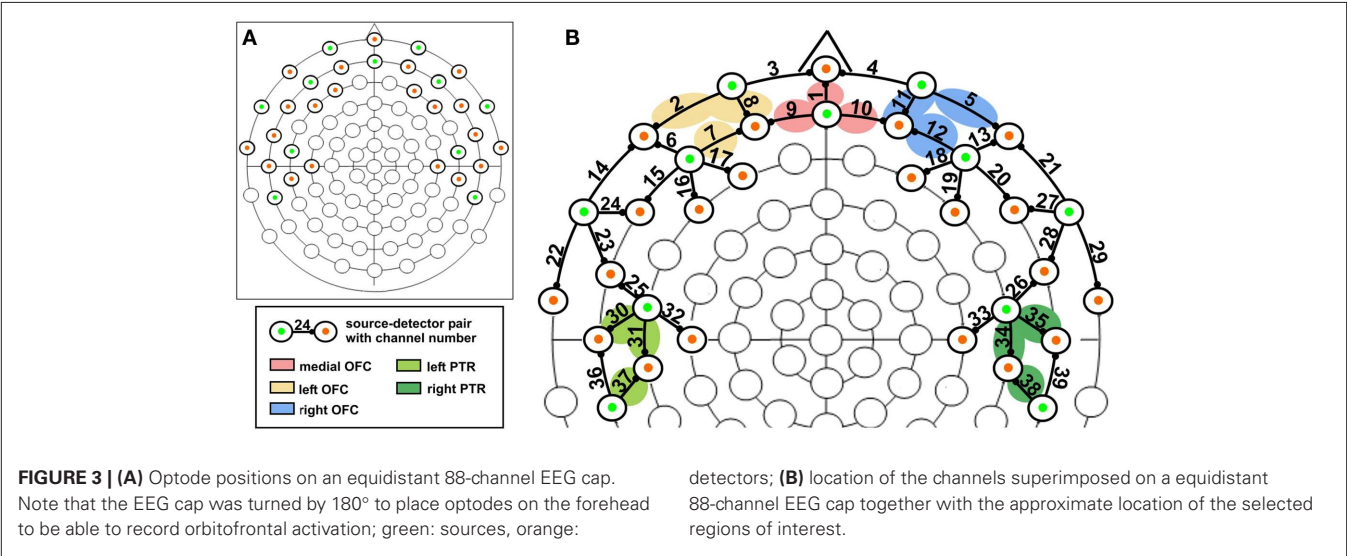


Table 2 | Significant findings for non-verbal cartoon stories.

	Main effects			
	[Oxy-Hb]		[Deoxy-Hb]	
	T-value	p-value	T-value	p-value
AFF > NEUTRAL				
Medial OFC	3.342	0.002**	0.146	0.884
Left OFC	5.048	0.000**	-3.504	0.001**
Right OFC	4.137	0.000**	-3.069	0.004**
Left dlPFC (channel 15)	0.747	0.472	-3.481	0.001**
Left dlPFC (channel 17)	2.823	0.007**	-1.771	0.084
Right dlPFC (channel 19)	3.007	0.005**	-0.546	0.588
COG > NEUTRAL				
Left OFC	2.800	0.008**	-3.937	0.000**
Right OFC	3.029	0.004**	-1.251	0.218
Left dlPFC (channel 16)	3.851	0.000**	-2.198	0.043

Two-tailed *t*-tests; ***p* < 0.01.

Table 3 | Significant findings for verbal listening series.

	Main effects			
	[Oxy-Hb]		[Deoxy-Hb]	
	T-value	p-value	T-value	p-value
AFF > NEUTRAL				
Left OFC	2.741	0.009**	-1.195	0.239
COG > NEUTRAL				
Left OFC	1.386	0.173	-4.035	0.000**
Right OFC	1.313	0.196	-2.952	0.005**
Right dlPFC (channel 20)	0.126	0.827	-2.951	0.005**

Two-tailed *t*-tests; ***p* < 0.01.

CONTRAST (1): AFFECTIVE > NEUTRAL WITH TWO CHARACTERS

Cartoons

Significant activations were observed in medial OFC and bilateral OFC. In addition, channels in dorsolateral prefrontal cortex (dlPFC) revealed significant activations: channel 15 [deoxy-Hb] and 17 [oxy-Hb] of the left and channel 19 [oxy-Hb] of the right dlPFC (see **Table 2**; **Figure 5**).

The *t*-test for age-group differences showed a significant age-effect in channel 17 [oxy-Hb], left dlPFC, demonstrating that older children activated this region to a greater degree than younger ones [$T(38) = -3.737$; $p = 0.001$].

Listening stories

The left OFC was significantly activated for the affective condition. Age differences were observed in [deoxy-Hb] channel 24 [$T(42) = 3.728$; $p = 0.001$] in the anterior left IFG. Again, older children showed higher activations (increased deactivations in [deoxy-Hb]) revealing age-related differences in left IFG involvement in affective empathy processing (see **Table 3**; **Figure 6**).

CONTRAST (2): COGNITIVE > NEUTRAL WITH ONE CHARACTER

Cartoons

Significant activations related to cognitive empathy processing were elicited in right and left OFC, as well as in channel 16 [oxy-Hb] of the left dlPFC. No significant age differences were observed in this contrast.

Listening stories

Significant activations were observed in left and right OFC. Additionally [deoxy-Hb] channel 20 of the right dlPFC showed a significant activation. No age-group differences were found.

CONTRAST (3): (AFFECTIVE – NEUTRAL WITH ONE CHARACTER) > (COGNITIVE – NEUTRAL WITH TWO CHARACTERS)

Cartoons

No region reached the level of significance in the direct contrast. No age differences were found.

Listening stories

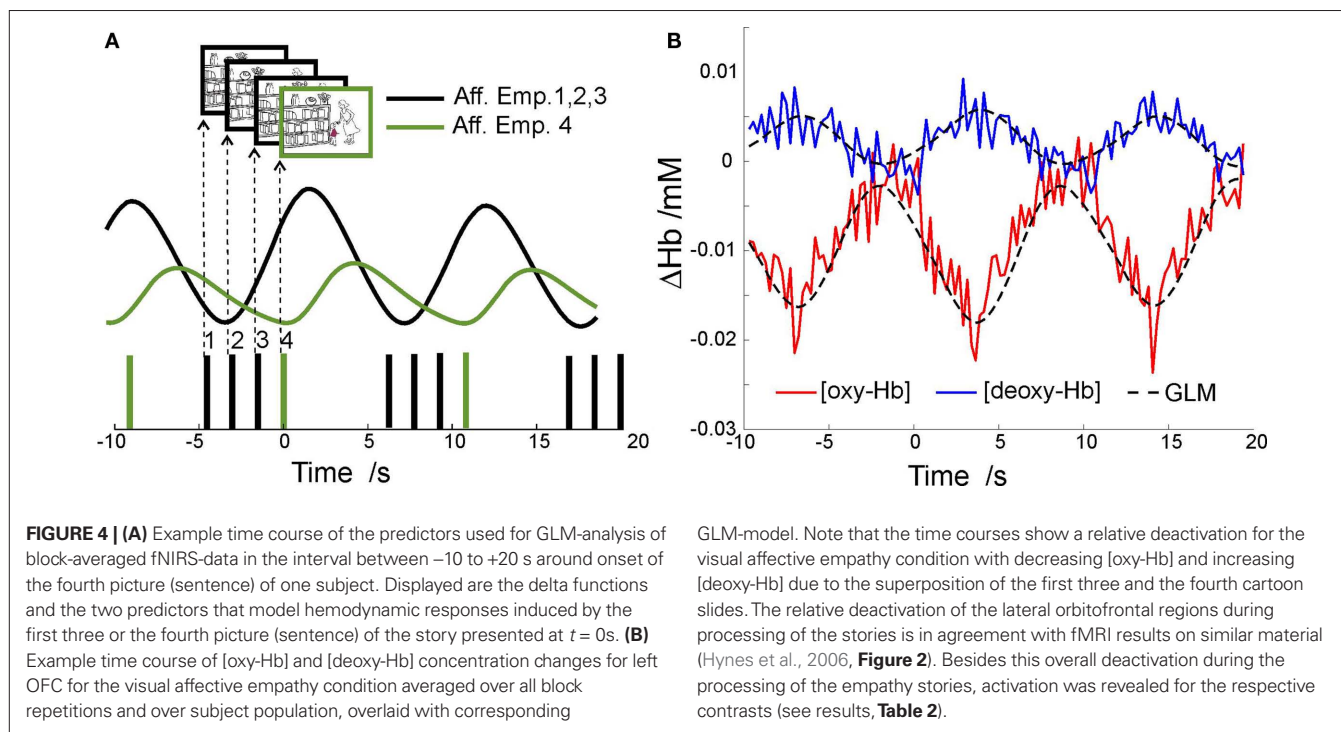
No region reached the level of significance. A significant effect of age group was observed in the medial OFC ([oxy-Hb]; $T(42) = -2.925$; $p = 0.006$). The effect was due to larger positive concentration changes in [oxy-Hb] in the older children's group in the affective empathy condition compared with higher activation of medial OFC in the cognitive empathy condition in younger children.

The correlational analysis with the affective and cognitive GEM scores revealed significant correlations with the parent-rated affective empathy measure only. Significant positive correlations were found for the non-verbal cartoon-modality in the medial OFC in contrast (1) ([Affective > Neutral]; [deoxy-Hb]; $r = 0.409$, $p = 0.009$) as well as in contrast (2) ([Cognitive > Neutral]; [deoxy-Hb]; $r = 0.460$; $p = 0.003$), revealing activation decreases in children with higher affective GEM scores in the medial OFC.

DISCUSSION

The present study investigated the neural correlates of affective and cognitive empathy in young children in orbitofrontal and posterior temporal regions. We measured hemodynamic responses of empathy processing in a non-verbal cartoon and a verbal story-listening task using fNIRS and observed activations related to empathy processing in medial and bilateral OFC for both types of stimuli.

Our results provide evidence for the hypothesis that empathy processing in young, healthy children requires a greater involvement of orbitofrontal brain regions, irrespective of whether the task elicits affective or cognitive empathy. As such, the present study extends recent findings from studies examining adult participants that also found orbitofrontal regions involved in empathy processing (Hynes et al., 2006; Decety and Meyer, 2008; Decety et al., 2008). Thus, although it is discussed that anterior frontal regions mature relatively late (until the age of 20, see Giedd et al., 1999), and maturation itself is suggested to be the reason for the late development of complex social cognition skills in ontogeny, our results suggest that also young children's empathic processing depends on orbitofrontal functioning comparable to what is known from adult studies.



MAIN FINDINGS

In line with our initial hypothesis, affective as well as cognitive empathy processing activates the OFC bilaterally. The occurrence of similar activation patterns for the two conditions is additionally supported by the finding that the direct (Affective > Cognitive) contrast showed no significant differences.

The role of the OFC has been examined by recent fMRI studies, investigating either affective (e.g., Decety et al., 2008) or cognitive (see Carrington and Bailey, 2009) empathic processing in adults, but without contrasting the two conditions directly. The neuropsychological lesion literature indicates, too, that the critical areas for empathic impairment of empathy processing are located in the OFC, as damage to those regions is associated with deficits in empathy (Eslinger, 1998; Shamay-Tsoory et al., 2004).

Only two fMRI studies, both on adults, were conducted that were comparable: Völlm et al. (2006) as well as Hynes et al. (2006) found OFC activations for both conditions bilaterally. In the latter study, affective empathic processing elicited higher activation and additionally activated the medial OFC (Hynes et al., 2006). Still, these studies differ from the present study in critical aspects and therefore their results are not directly comparable to ours: Both fMRI studies relied on adult participants, who were asked to make explicit empathy judgments, whereas the present study might have triggered empathic responses implicitly, more closely related to the concept of empathy as “Einfühlen,” defining empathy as a mode of inner imitation or “feeling into someone” (Eisenberg and Strayer, 1987; Barnes and Thagard, 1997).

A direct comparison of the auditory and visual modality has so far not been conducted with young participants and it seems likely that their empathic processing might differ from that in adults. Since no additional activations in temporal regions were

observed in the present study, differences between children’s and adults’ processing of empathic stories seem to be associated with differences in the engagement of temporal regions, and not the OFC. Still, it should be noted that the present definition of the PTR and its approximate location does not fully overlap with the often reported TPJ area in adults’ cognitive empathy processing. The examined PTR cluster is located in large parts along the superior temporal sulcus. The latter has reciprocal connections to the OFC (Barbas, 1988) and has been implicated in cognitive empathy processing and social cognition (Allison et al., 2000; Hein and Singer, 2008). **Figure 5** shows that activations in PTR occurred in both visual empathy conditions, but did not reach our *a priori* level of significance¹. Whether this trend occurred because of a loss in statistical power due to methodological issues or whether it indicates principal differences between an involvement of the posterior superior temporal sulcus and the neighboring TPJ in children’s empathy processing should be examined in future longitudinal neuroimaging studies using similar paradigms in adults and children that can help solving this issue.

The OFC is suggested to be a key region supporting social cognition in general: It was found to be critical for ToM-processing (Carrington and Bailey, 2009) and decision making (Kringelbach, 2005) on the one hand, and moral appraisals (Moll et al., 2002) and affective empathy (Decety and Jackson, 2004; Hynes et al., 2006) on the other hand. Thus, it is likely that social cognition, including empathy, is supported by OFC functioning, independent of how affective or cognitive the process might be.

¹left PTR [deoxy-Hb]: visual (Affective > Neutral) contrast: $T(39) = -1.950$; $p = 0.058$; left PTR [deoxy-Hb]: visual (Cognitive > Neutral) contrast: $T(39) = -2.176$; $p = 0.036$;

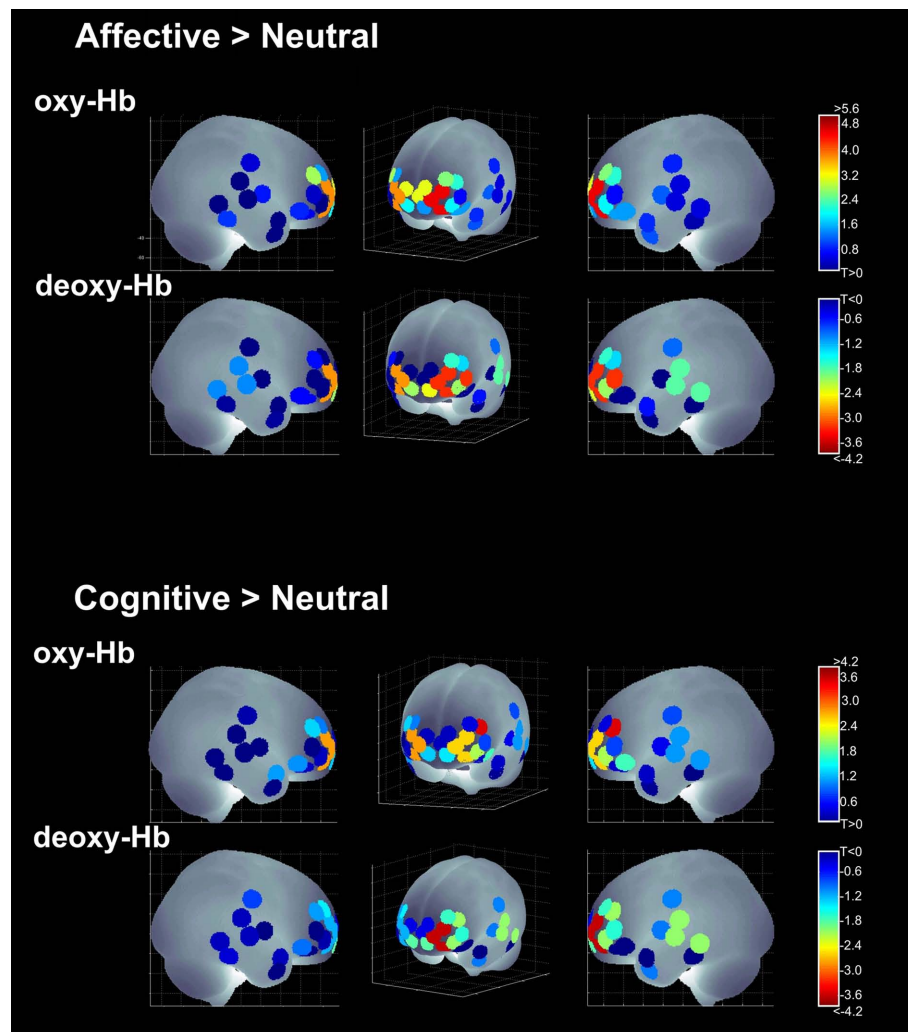


FIGURE 5 | Activations for the visual condition; [oxy-Hb] and [deoxy-Hb] results across all four main contrasts in the sagittal and a frontal view.

The observed involvement of the OFC in empathy processing is also in line with theories that propose a hierarchical structure of the prefrontal cortex, postulating more complex cognitive processes, the more anterior a region is located in the frontal lobe (Fuster, 2004; Botvinick, 2008; Barbey et al., 2009). As a first result, thus, complex socio-emotional processing as in the affective and cognitive empathy conditions of our study, demanding the mirroring of different plots, intentions and emotions, seems to require the most anterior part of the brain in young children.

It still remains an open question why other studies on the same issue did not find OFC activations for affective and cognitive empathy (see Hein and Singer, 2008). There is a lot of variance among activated areas within empathy paradigms (Carrington and Bailey, 2009). It is also obvious from the heterogeneity in the literature that there are several ways of defining and assessing empathy and the relationship of its components. For example, “affective ToM,” “affective perspective-taking,” and “affective empathy” are often used synonymously, but do not always really describe and measure the same process, as

these concepts can focus on different aspects. Thus, it can be discussed whether the notion of affective empathy in terms of a social cognition process differs from the notion of affective empathy in terms of pain in others (e.g., Decety et al., 2008; Hein and Singer, 2008).

Finally, the diversity of paradigms and materials may account for some of the variability in the results of different studies: Written scenarios or cartoons (Kobayashi et al., 2007), interpretation of facial expression from the eyes (Baron-Cohen et al., 1999), detection of faux-pas and irony (Shamay-Tsoory et al., 2005), judgments (Völlm et al., 2006), or pain in others (Decety et al., 2008) are all paradigms used in empathy research. The present study was designed to focus on the direct comparison of equivalent processing of affective versus cognitive empathy, while passively following the course of a stimulus story. We think a strength of the present study lies in the parallel measurement of a visual non-verbal and an auditory verbal condition, which shows a remarkably high degree of concordance in the results and thus is able to replicate the OFC findings in each presentation domain. However, across all conditions,

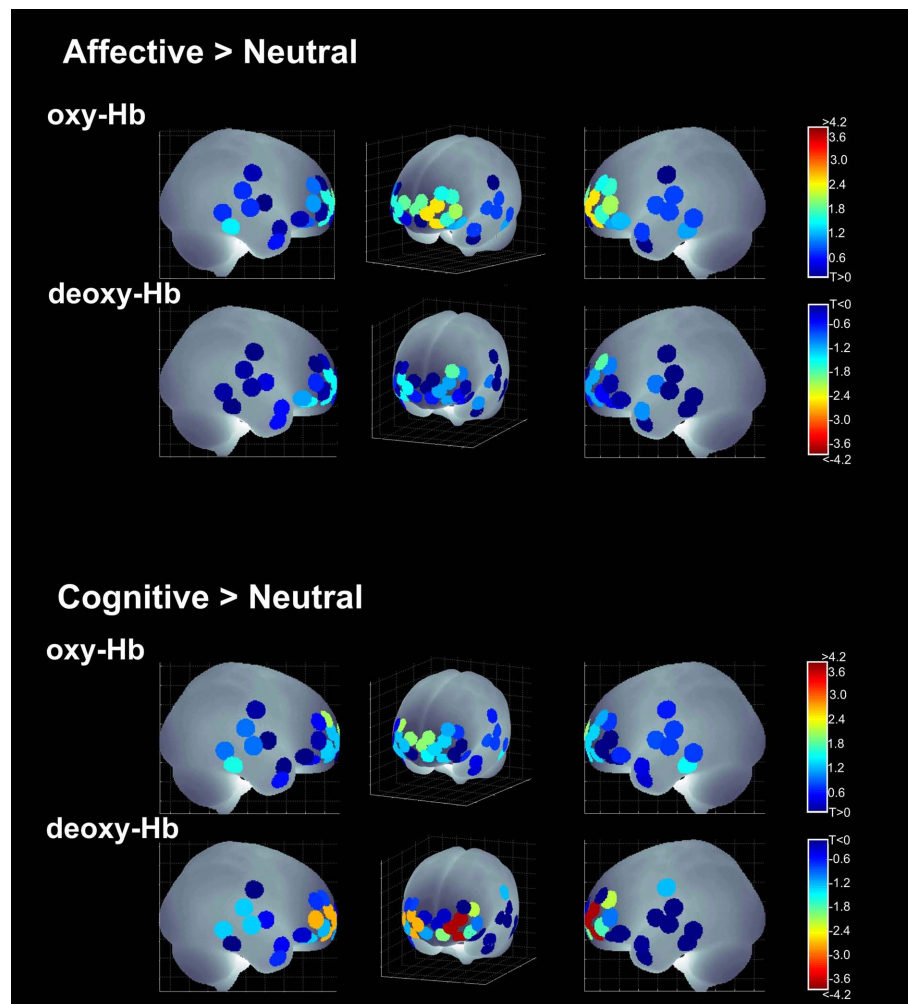


FIGURE 6 | Activations for the auditory condition; [oxy-Hb] and [deoxy-Hb] results across all four main contrasts in the sagittal and a frontal view.

cartoon stories elicited overall higher activations, recruited more anterior frontal regions in the OFC, and enhanced activations in dorsolateral regions. This points to the hypothesized visual dominance in younger children's processing (Guttentag, 1985), and a higher familiarity and external validity of cartoon-like stories. Thus, it seems likely that the visual modality facilitates the processing of its content and therefore the processing of empathic contents, too.

Only small differences are visible between affective and cognitive empathy processing in the present study and those did not survive the direct contrasting. fNIRS can only measure hemodynamic responses in the outer cortical regions. Thus, it is obvious that this method is not able to detect emotion-related differences in subcortical parts of the emotion processing circuits (Dolan, 2002), where further specific activations related to affective processing may be located. Still, our finding of additional medial OFC involvement, specific for affective empathy processing, is in line with the results of Hynes et al. (2006) for adults, even if we observed this activation for the visual modality only. Hynes et al. (2006) report medial OFC and medial PFC activations for affective processing, suggesting this

to be the difference between the two empathy-components. Again, it should be noted that medial PFC has not been the focus of the present study due to the low sensitivity of fNIRS to signals in these deeper brain regions along the medial wall.

We further observed an effect of children's age in medial OFC, related to affective empathy processing: Older children showed higher medial OFC activation for the direct (Affective empathy > Cognitive empathy) contrast than younger ones. We think that this main effect of age indicates a shift toward an involvement of additional medial parts of the anterior frontal lobe with older age in affective but not in cognitive empathy processing. This finding is supported by the neuroimaging literature on adults' affective empathy processing, who additionally require medial frontal regions (Decety and Jackson, 2004; Hynes et al., 2006; Hein and Singer, 2008). As such, our results show a pattern opposite to that of a recent neurodevelopmental study by Decety and Michalska (2010), who conducted an fMRI study with participants ranging from 7 to 40 years of age. Decety and Michalska (2010) report greater activation in the OFC in response to affective empathy-eliciting

scenarios (depicting intentional harm) that was shifted from its medial portion in younger participants, to the lateral portion in older participants. The authors discuss this pattern of developmental change in the OFC as reflecting a gradual shift from the monitoring of somatovisceral responses in young children to the executive control of emotion processing in older participants. Interestingly, Decety and Michalska (2010) report also age-related correlations in lateral frontal regions (enhanced involvement with older age) which seem to be more consistent with age-related findings in left dlPFC and IFG for the affective empathy condition.

Because Decety and Michalska (2010) investigated a much larger age range (with only a small amount of children at the age of our sample) and focused on somatovisceral empathic processing, whereas our study examined affective empathy processing in everyday social contexts, these two studies are not directly comparable. Still, we see the increased activation in medial OFC in older children as evidence in support of the hypothesis of an ongoing development of affective empathy processing in children aged between 4 and 8 years, with a further specialization of medial OFC in social-cognitive affective empathy processing in the older children. Whereas the younger age group shows no differential activation in medial OFC between the affective and cognitive condition, older children have a higher activation in medial OFC during affective empathy processing compared than during cognitive empathy processing².

Still, one might wonder why this age-effect was only found in the auditory condition. This finding suggests that visual and auditory processing of empathy stories in young children are qualitatively different due to their different everyday familiarity. We suggest this to result from a visual processing dominance in children. Children are probably more used to following visually guided cartoon stories than to listening mindfully to auditory empathy stories (e.g., Hayes and Birnbaum, 1980). Moreover, one could suppose that visual empathic processing along the OFC is developed earlier in ontogeny than the development of auditory empathic processing. For example, when children's auditory and visual perception of video-clips was tested, the results provide evidence that children, especially preschoolers, payed more attention to and recognized more details in visual than in auditory information (Grieve and Williamson, 1977; Hayes and Birnbaum, 1980). This is supported by the pattern of results for the visual condition which in general more closely mirrors the results obtained in studies with adults. The observed medial OFC activation in the visual affective empathy condition across the whole sample may result from facilitated and further-developed processing in the visual domain. Accordingly, it seems possible that it is easier to detect developmental changes in the auditory domain. Still, the adult fMRI studies relied on visual processing, and processing of empathy has not been investigated with auditory stimuli yet.

A second developmental effect was visible in affective empathy processing of auditory stimuli in the left IFG. As mentioned above, involvement of left IFG is often reported in adult studies on empathy processing and the MNS. These differences in the

activation of the left IFG probably point to a similar age-dependent shift in affective empathy processing in older compared with younger children. Older children's processing of auditory affective empathy stories is associated with higher involvement of the left IFG. Thus, children seem to increase the use of the frontal MNS more with ongoing development.

A different view comes from semantic memory research. Left IFG has been shown to be more active in tasks requiring retrieval of semantic knowledge and verbal recoding (Thompson-Schill et al., 1997; Poldrack et al., 1999). Thus, it is also reasonable to argue that younger and older children differ in the amount of verbal processing (e.g., rehearsal or recoding) when listening to affective empathy stories. As the IFG was only activated for age differences, the theory of a developing MNS seems to be a plausible account.

Finally, an age-effect was observed for channel 17 [oxy-Hb], showing higher activation in left dlPFC for older children for affective empathic processing in the visual domain. Apart from a general discussion on the role of dlPFC in working memory and executive functions (Miller and Cohen, 2001; Fuster, 2004; Petrides, 2005), an involvement of dlPFC in empathy processing has previously been reported in a lesion study (Shamay-Tsoory et al., 2003; also see Shamay-Tsoory and Aharon-Peretz, 2007). Patients with dlPFC lesions, as well as patients with ventro-medial PFC lesions, showed higher deficits in different ToM tasks, as compared to healthy control participants. Still, these authors also showed that empathy performance in the dlPFC group correlates with cognitive flexibility measures, again pointing more toward a role of dlPFC in executive functioning as a basis for empathy processing.

Taken together, the age-effects observed in the present study are all related to affective empathy processing, thus indicating developmental changes in the affective empathy component in the age-group investigated. Future research might use the same design in a study on adults to further our knowledge of developmental differences between children and adults.

ADDITIONAL FINDINGS

Apart from the finding of medial OFC involvement, no specific activations were found for the affective versus cognitive empathy comparisons. Also, no PTR cluster and no other channel in the temporal lobe became significantly activated. Temporal regions are discussed to be specifically activated for cognitive empathic processing in adults (Frith and Frith, 2006a). As Carrington and Bailey (2009) meta-analyzed, many different studies on ToM found different regions, but no single region to be activated concordantly. Only few other functional imaging studies on children's empathic processing (Kobayashi et al., 2007; Pfeifer et al., 2008; Decety and Michalska, 2010, age seven and above) reported a TPJ activation. In this study, school children processed ToM stories and were asked to judge false beliefs (Kobayashi et al., 2007). Thus, it is likely that passive following of empathic stories in young children, as in the present study, does not recruit additional neural circuits in anterior and posterior temporal regions. Empathic processing in this age range may be restricted to the most anterior parts of the brain.

For the visual modality, correlations of the affective GEM scores with medial OFC activation were found, revealing a negative correlation with affective and cognitive empathic processing, whereas cognitive GEM scores did not correlate significantly with

²To further evaluate the age-related effect in [oxy-Hb] in contrast 3 (Affective empathy > Cognitive empathy) *post hoc* simple *t*-test were computed for each age-group separately. In young children, no significant contrast effect was visible [$T(22) = -1.750$; $p = 0.094$], whereas older children showed a significant higher activation in affective empathy processing [$T(20) = 2.609$; $p = 0.017$].

any activated brain region in the present study. These results lead to the question to what extent children who score high in affective GEM activate the medial OFC less for affective and for cognitive empathic processing.

As parent-rated GEM scores and age did not correlate in the present study, this effect cannot simply be attributed to age. The finding suggests that children who are more affectively empathic process affective and cognitive empathy stories differently in the medial OFC (whose activation was only found to be related with affective empathic processing). Although the particular correlations are negative, this result mirrors well the above findings which point to the close relationship between medial OFC and children's affective empathy processing. Here, less activation in the medial OFC is found for children who are more empathic. This rather paradoxical finding might indicate a shift of empathic processing away from anterior frontal regions in empathic children, toward subcortical processing (as mirrored by the reported positive correlations in adult studies between empathic traits and anterior insula activation, e.g., Moriguchi et al., 2007; Silani et al., 2008, see Hein and Singer, 2008). However, this aspect remains speculative and should be investigated in future studies using fMRI.

With the present study, we introduce the method of fNIRS to the field of neurodevelopmental studies on empathy processing, and we believe it is a valuable tool for research in this field. fNIRS has several advantages over other imaging methods such as fMRI or PET, first of all, that no fixation of the head is needed. The child can sit on a comfortable chair in front of a screen, communicate with supervisor and parents during preparation and, if necessary, also during the trials. The optodes can easily be prepared beforehand, thus the actual preparation is comfortable and not so time-consuming. Because of its low constraints on the experimental environment, fNIRS is a good tool to investigate higher cognition, whereas an MRI environment may impair children's concentration and speech perception (Hofmann et al., 2008). It has already been established as a method in studying neuropsychological development in neonates and infants (Peña et al., 2003; Taga et al., 2004; Wartenburger et al., 2007). Additionally, fNIRS might be less prone than, e.g., fMRI to artifacts caused by the proximity to the air-filled sinuses of the OFC, when measuring OFC activations (Kringelbach, 2005). However, fNIRS cannot supply the spatial resolution of fMRI-based approaches (Hofmann et al., 2008).

There are some limitations to our study. Both limitations are based on the fact that we employed a design that was applicable to young children, having in mind not to overtax the children. First of all, the application of a passive following paradigm in children

does not allow to conclude that all children processed all stories deeply and with an intention to take the point of view of the main character which might have led to less activations in empathy processing regions. Secondly, one might ask whether the adult ratings to the empathy stories are also valid for children. Children's ratings are not yet available for the stimulus set and it is questionable to what extent young children are able to explicitly judge their empathic involvement – which is the reason why we had to base the study on adult rating data. Therefore, we cannot be certain, for example, whether the children processed all empathy stories in an affective manner or not. Still, given the post-experimental interviews, we believe, that the children processed the stories in the intended manner, and the high concordance of the results with that of previous adult studies seems to validate this assumption. Further studies are needed which directly address the differences between passively following and explicitly judging of empathy stories in young children.

CONCLUSION

Taken together, our findings provide evidence for higher medial and bilateral OFC activation in both, affective and cognitive empathy processing in a sample of young children 4–8 years of age. Thus, in a manner similar to what is known from adult OFC recruitment in complex social cognition tasks and empathy processing, orbitofrontal regions were involved in a task in which children passively followed empathic narratives – independently of whether these stories presented social situations where a character experiences affective outcomes of its own action or the plot required the mentalizing and prediction of further actions, and independently of whether these stories were presented visually or auditorily. Hence, our results support the idea that the OFC is a brain region associated with computing and evaluating predictions of other persons' actions and the comparison of these predictions with subjective states across both affective and non-affective situations.

Furthermore, in contrast to our initial hypotheses, developmental changes with increased brain activation in older children were observed in affective empathy processing as compared to neutral stories in left dlPFC in the visual condition and left anterior IFG in the auditory condition, but no age-related effects were observed in cognitive empathy processing. In contrast, medial OFC showed a higher activation when directly contrasting affective and cognitive processing conditions. Thus, the results support the idea of medial OFC being especially engaged in socio-affective processing and a development of medial OFC functioning toward a higher involvement with older ages during childhood.

REFERENCES

- Allison, T., Puce, A., and McCarthy, G. (2000). Social perception from visual cues: role of the STS region. *Trends Cogn. Sci. (Regul. Ed.)* 4, 267–278.
- Amodio, D. M., and Frith, C. D. (2006). Meeting of minds: the medial frontal cortex and social cognition. *Nat. Rev. Neurosci.* 7, 268–277.
- Barbas, H. (1988). Anatomic organization of basoventral and mediodorsal visual recipient prefrontal regions in the rhesus monkey. *J. Comp. Neurol.* 276, 313–342.
- Barbey, A. K., Krueger, E., and Grafman, J. (2009). An evolutionarily adaptive neural architecture for social reasoning. *Trends Neurosci.* 32, 603–610.
- Barnes, A., and Thagard, P. (1997). Empathy and analogy. *Dialogue Can. Philos. Rev.* 36, 705–720.
- Baron-Cohen, S., Leslie, A. M., and Frith, U. (1985). Does the autistic child have a “theory of mind.” *Cognition* 21, 37–46.
- Baron-Cohen, S., Ring, H. A., Wheelwright, S., Bullmore, E. T., Brammer, M. J., Simmons, A., and Williams, S. C. R. (1999). Social intelligence in the normal and autistic brain: an fMRI study. *Eur. J. Neurosci.* 11, 1891–1898.
- Baron-Cohen, S., and Wheelwright, S. (2004). The empathy quotient: an investigation of adults with Asperger syndrome or high functioning autism, and normal sex differences. *J. Autism Dev. Disord.* 34, 163–175.
- Blair, R. J. (2005). Responding to the emotions of others: dissociating forms of empathy through the study of typical and psychiatric populations. *Conscious. Cogn.* 14, 698–718.
- Blair, R. J. (2008). Fine cuts of empathy and the amygdala: dissociable deficits in psychopathy and autism. *Q. J. Exp. Psychol.* 61, 157.
- Botvinick, M. M. (2008). Hierarchical models of behavior and prefrontal function. *Trends Cogn. Sci. (Regul. Ed.)* 12, 201–208.
- Carrington, S. J., and Bailey, A. J. (2009). Are there theory of mind regions in the

- brain? A review of the neuroimaging literature. *Hum. Brain Mapp.* 30, 2313–2335.
- Casey, B. J. (1999). Images in neuroscience. Brain development. XII. Maturation in brain activation. *Am. J. Psychiatry* 156, 504–504.
- Cope, M., and Delpy, D. T. (1988). System for long term measurement of cerebral blood and tissue oxygenation on newborn infants by near infrared transillumination. *Med. Biol. Eng. Comput.* 26, 289–294.
- Dadds, M. R., Hunter, K., Hawes, D. J., Frost, A. D., Vassallo, S., Bunn, P., Merz, S., and El Masry, Y. (2008). A measure of cognitive and affective empathy in children using parent ratings. *Child Psychiatry Hum. Dev.* 39, 111–122.
- Decety, J. (2011). Dissecting the neural mechanisms mediating empathy. *Emot. Rev.* 3, 92–108.
- Decety, J., and Jackson, P. L. (2004). The functional architecture of human empathy. *Behav. Cogn. Neurosci. Rev.* 3, 71–100.
- Decety, J., and Jackson, P. L. (2006). A social-neuroscience perspective on empathy. *Curr. Dir. Psychol. Sci.* 15, 54–58.
- Decety, J., and Lamm, C. (2007). The role of the right temporoparietal junction in social interaction: how low-level computational processes contribute to meta-cognition. *Neuroscientist* 13, 580–593.
- Decety, J., and Meyer, M. (2008). From emotion resonance to empathic understanding: a social developmental neuroscience account. *Dev. Psychopathol.* 20, 1053–1080.
- Decety, J., and Michalska, K. J. (2010). Neurodevelopmental changes in the circuits underlying empathy and sympathy from childhood to adulthood. *Dev. Sci.* 13, 886–899.
- Decety, J., Michalska, K. J., and Akitsuki, Y. (2008). Who caused the pain? An fMRI investigation of empathy and intentionality in children. *Neuropsychologia* 46, 2607–2614.
- Dolan, R. J. (2002). Emotion, cognition and behavior. *Science* 298, 1191–1194.
- Durston, S., and Casey, B. J. (2006). What have we learned about cognitive development from neuroimaging? *Neuropsychologia* 44, 2149–2157.
- Dziobek, I., Rogers, K., Fleck, S., Bahnemann, M., Heekeren, H. R., Wolf, O., and Convit, A. (2008). Dissociation of cognitive and emotional empathy in adults with Asperger syndrome using the multifaceted empathy test (MET). *J. Autism Dev. Disord.* 38, 464–473.
- Eisenberg, N. (2000). “Empathy and sympathy,” in *Handbook of Emotions*, eds M. Lewis and J. M. Haviland-Jones (New York: Guilford Press), 677–691.
- Eisenberg, N., and Miller, P. A. (1987). The relation of empathy to prosocial and related behaviors. *Psychol. Bull.* 101, 91–119.
- Eisenberg, N., and Strayer, J. (1987). “Critical issues in the study of empathy,” in *Empathy and its Development*, eds N. Eisenberg and J. Strayer (New York: Cambridge University Press), 3–13.
- Eslinger, P. J. (1998). Neurological and neuropsychological bases of empathy. *Eur. Neurol.* 39, 193–199.
- Faria, A. V., Zhang, J., Oishi, K., Li, X., Jiang, H., Akhter, K., Hermoye, L., Lee, S., Hoon, A., Stashinko, E., Miller, M. I., van Zijl, P. C. M., and Mor, S. (2010). Atlas-based analysis of neurodevelopment from infancy to adulthood using diffusion tensor imaging and application for automated abnormality detection. *Neuroimage* 52, 415–428.
- Flavell, J. H. (1999). Cognitive development: children’s knowledge about the mind. *Annu. Rev. Psychol.* 50, 21–45.
- Frith, C. D., and Frith, U. (2006a). The neural basis of mentalizing. *Neuron* 50, 531–534.
- Frith, C. D., and Frith, U. (2006b). How we predict what other people are going to do. *Brain Res.* 1079, 36–46.
- Fuster, J. M. (2004). Upper processing stages of the perception–action cycle. *Trends Cogn. Sci. (Regul. Ed.)* 8, 143–145.
- Gallese, V., and Goldman, A. (1998). Mirror neurons and the simulation theory of mind-reading. *Trends Cogn. Sci. (Regul. Ed.)* 2, 493–501.
- Giedd, J. N., Blumenthal, J., Jeffries, N. O., Castellanos, F. X., Liu, H., Zijdenbos, A., Paus, T., Evans, A. C., and Rapoport, J. L. (1999). Brain development during childhood and adolescence: a longitudinal MRI study. *Nat. Neurosci.* 2, 861–862.
- Grieve, R., and Williamson, K. (1977). Aspects of auditory and visual attention to narrative material in normal and mentally handicapped children. *J. Child Psychol. Psychiatry* 18, 251–262.
- Guttentag, R. E. (1985). A developmental study of attention to auditory and visual signals. *J. Exp. Child. Psychol.* 39, 546–561.
- Hamlin, J. K., Wynn, K., and Bloom, P. (2007). Social evaluation by preverbal infants. *Nature* 450, 557–559.
- Hayes, D. S., and Birnbaum, D. W. (1980). Preschoolers’ retention of televised events: is a picture worth a thousand words. *Dev. Psychol.* 16, 410–416.
- Hein, G., and Singer, T. (2008). I feel how you feel but not always: the empathic brain and its modulation. *Curr. Opin. Neurobiol.* 18, 153–158.
- Hoffman, M. L. (2000). *Empathy and Moral Development*. Cambridge: Cambridge University Press.
- Hofmann, M. J., Herrmann, M. J., Dan, I., Obrig, H., Conrad, M., Kuchinke, L., Jacobs, A. M., and Fallgatter, A. J. (2008). Differential activation of frontal and parietal regions during visual word recognition: an optical topography study. *Neuroimage* 40, 1340–1349.
- Hooker, C. I., Verosky, S. C., Germine, L. T., Knight, R. T., and D’Esposito, M. (2009). Neural activity during social signal perception correlates with self-reported empathy. *Brain Res.* 1308, 100–113.
- Hynes, C. A., Baird, A. A., and Grafton, S. T. (2006). Differential role of the orbital frontal lobe in emotional versus cognitive perspective-taking. *Neuropsychologia* 44, 374–383.
- Ickes, W. J. (1997). *Empathic Accuracy*. New York: The Guilford Press.
- Karmiloff-Smith, A. (2010). Neuroimaging of the developing brain: taking “developing” seriously. *Hum. Brain Mapp.* 31, 934–941.
- Kaufman, A. S., and Applegate, B. (1988). Short forms of the K-ABC mental processing and achievement scales at ages 4 to 12 1/2 years for clinical and screening purposes. *J. Clin. Child Adolesc. Psychol.* 17, 359–369.
- Kobayashi, C., Glover, G. H., and Temple, E. (2007). Children’s and adults’ neural bases of verbal and nonverbal “theory of mind”. *Neuropsychologia* 45, 1522–1532.
- Koch, S. P., Steinbrink, J., and Obrig, H. (2009). “NILAB – a Matlab based processing tool for functional optical data,” in *Optical Imaging Workshop*, December 7–10, 2009, Bad Honnef.
- Kringelbach, M. L. (2005). The human orbitofrontal cortex: linking reward to hedonic experience. *Nat. Rev. Neurosci.* 6, 691–702.
- Lamm, C., Batson, C. D., and Decety, J. (2007). The neural substrate of human empathy: effects of perspective-taking and cognitive appraisal. *J. Cogn. Neurosci.* 19, 42–58.
- Leslie, A. M. (1987). Pretense and representation: the origins of “theory of mind”. *Psychol. Rev.* 94, 412–426.
- Lloyd-Fox, S., Blasi, A., and Elwell, C. E. (2010). Illuminating the developing brain: the past, present, and future of near infrared spectroscopy. *Neurosci. Biobehav. Rev.* 34, 269–284.
- Miller, E. K., and Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annu. Rev. Neurosci.* 24, 167–202.
- Moll, J., de Oliveira-Souza, R., Eslinger, P. J., Bramati, I. E., Mourão-Miranda, J., Andreiulo, P. A., and Pessoa, L. (2002). The neural correlates of moral sensitivity: a functional magnetic resonance imaging investigation of basic and moral emotions. *J. Neurosci.* 22, 2730–2736.
- Moriguchi, Y., Decety, J., Ohnishi, T., Maeda, M., Mori, T., Nemoto, K., Matsuda, H., and Komaki, G. (2007). Empathy and judging other’s pain: an fMRI study of alexithymia. *Cereb. Cortex* 17, 2223–2234.
- Obrig, H., and Villringer, A. (2003). Beyond the visible – imaging the human brain with light. *J. Cereb. Blood Flow Metab.* 23, 1–18.
- Okamoto, M., Dan, H., Sakamoto, K., Takeo, K., Shimizu, K., Kohno, S., Oda, I., Isobe, S., Suzuki, T., Kohyama, K., and Dan, I. (2004). Three-dimensional probabilistic anatomical cranio-cerebral correlation via the international 10–20 system oriented for transcranial functional brain mapping. *Neuroimage* 21, 99–111.
- Peña, M., Maki, A., Kovačić, D., Dehaene-Lambertz, G., Koizumi, H., Bouquet, F., and Mehler, J. (2003). Sounds and silence: an optical topography study of language recognition at birth. *Proc. Natl. Acad. Sci. U.S.A.* 100, 11702–11705.
- Perner, J., Leekam, S. R., and Wimmer, H. (1987). Three-year-olds’ difficulty with false belief: the case for a conceptual deficit. *Br. J. Dev. Psychol.* 5, 125–137.
- Perner, J., and Wimmer, H. (1985). “John thinks that Mary thinks that...”: attribution of second-order beliefs by 5- to 10-year-old children. *J. Exp. Child Psychol.* 39, 437–471.
- Petrides, M. (2005). Lateral prefrontal cortex: architectonic and functional organization. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 360, 781–795.
- Pfeifer, J. H., Iacoboni, M., Mazziotta, J. C., and Dapretto, M. (2008). Mirroring others’ emotions relates to empathy and interpersonal competence in children. *Neuroimage* 39, 2076–2085.
- Poldrack, R. A., Wagner, A. D., Prull, M. W., Desmond, J. E., Glover, G. H., and Gabrieli, J. D. (1999). Functional specialization for semantic and phonological processing in the left inferior prefrontal cortex. *Neuroimage* 10, 15–35.
- Shamay-Tsoory, S. G., and Aharon-Peretz, J. (2007). Dissociable prefrontal networks for cognitive and affective theory of mind: a lesion study. *Neuropsychologia* 45, 3054–3067.
- Shamay-Tsoory, S. G., Aharon-Peretz, J., and Perry, D. (2008). Two systems for empathy: a double dissociation between emotional and cognitive empathy in inferior frontal gyrus versus ventromedial prefrontal lesions. *Brain* 132, 617–627.
- Shamay-Tsoory, S. G., Tomer, R., Berger, B., and Aharon-Peretz, J. (2003). Characterization of empathy deficits following prefrontal brain damage: the role of the right ventromedial

- prefrontal cortex. *J. Cogn. Neurosci.* 15, 324–337.
- Shamay-Tsoory, S. G., Tomer, R., Berger, B. D., Goldsher, D., and Aharon-Peretz, J. (2005). Impaired “affective theory of mind” is associated with right ventromedial prefrontal damage. *Cogn. Behav. Neurol.* 18, 55–67.
- Shamay-Tsoory, S. G., Tomer, R., Goldsher, D., Berger, B. D., and Aharon-Peretz, J. (2004). Impairment in cognitive and affective empathy in patients with brain lesions: anatomical and cognitive correlates. *J. Clin. Exp. Neuropsychol.* 26, 1113–1127.
- Silani, G., Bird, G., Brindley, R., Singer, T., Frith, C., and Frith, U. (2008). Levels of emotional awareness and autism: an fMRI study. *Soc. Neurosci.* 3, 97–112.
- Singer, T. (2006). The neuronal basis and ontogeny of empathy and mind reading: review of literature and implications for future research. *Neurosci. Biobehav. Rev.* 30, 855–863.
- Singh, A. K., Okamoto, M., Dan, H., Jurcak, V., and Dan, I. (2005). Spatial registration of multichannel multi-subject fNIRS data to MNI space without MRI. *Neuroimage* 27, 842–851.
- Soderstrom, H. (2003). Psychopathy as a disorder of empathy. *Eur. Child. Adolesc. Psychiatry* 12, 249–252.
- Steinbrink, J., Villringer, A., Kempf, F., Haux, D., Boden, S., and Obrig, H. (2006). Illuminating the BOLD signal: combined fMRI–fNIRS studies. *Magn. Reson. Imaging* 24, 495–505.
- Taga, G., Asakawa, K., Hirasawa, K., and Konishi, Y. (2004). Hemodynamic responses to visual stimulation in occipital and frontal cortex of newborn infants: a near-infrared optical topography study. *Pathophysiology* 10, 277–281.
- Thompson-Schill, S. L., D’Esposito, M., Aguirre, G. K., and Farah, M. J. (1997). Role of left inferior prefrontal cortex in retrieval of semantic knowledge: a reevaluation. *Proc. Natl. Acad. Sci. U.S.A.* 94, 14792–14797.
- Voßkübler, A., Nordmeier, V., Kuchinke, L., and Jacobs, A. M. (2008). OGAMA (open gaze and mouse analyzer): open-source software designed to analyze eye and mouse movements in slideshow study designs. *Behav. Res. Methods* 40, 1150–1162.
- Völlm, B. A., Taylor, A. N., Richardson, P., Corcoran, R., Stirling, J., McKie, S., Deakin, J. F. W., and Elliott, R. (2006). Neuronal correlates of theory of mind and empathy: a functional magnetic resonance imaging study in a nonverbal task. *Neuroimage* 29, 90–98.
- Warneken, F., and Tomasello, M. (2009). The roots of human altruism. *Br. J. Psychol.* 100, 455–471.
- Wartenburger, I., Steinbrink, J., Telkemeyer, S., Friedrich, M., Friederici, A. D., and Obrig, H. (2007). The processing of prosody: evidence of interhemispheric specialization at the age of four. *Neuroimage* 34, 416–425.
- Wellman, H. M., Cross, D., and Watson, J. (2001). Meta-analysis of theory-of-mind development: the Truth about false belief. *Child Dev.* 72, 655–684.
- Wimmer, H., and Perner, J. (1983). Beliefs about beliefs: representation and constraining function of wrong beliefs in young children’s understanding of deception. *Cognition* 13, 103–128.
- Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 30 November 2010; paper pending published: 22 December 2010; accepted: 13 April 2011; published online: 28 April 2011.
Citation: Brink TT, Urton K, Held D, Kirilina E, Hofmann MJ, Klann-Delius G, Jacobs AM and Kuchinke L (2011) The role of orbitofrontal cortex in processing empathy stories in 4- to 8-year-old children. *Front. Psychology* 2:80. doi: 10.3389/fpsyg.2011.00080

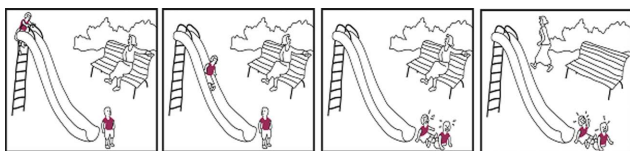
This article was submitted to *Frontiers in Language Sciences*, a specialty of *Frontiers in Psychology*.

Copyright © 2011 Brink, Urton, Held, Kirilina, Hofmann, Klann-Delius, Jacobs and Kuchinke. This is an open-access article subject to a non-exclusive license between the authors and Frontiers Media SA, which permits use, distribution and reproduction in other forums, provided the original authors and source are credited and other Frontiers conditions are complied with.

APPENDIX

VISUAL STIMULI

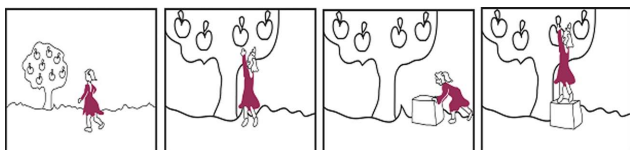
Affective empathy negative



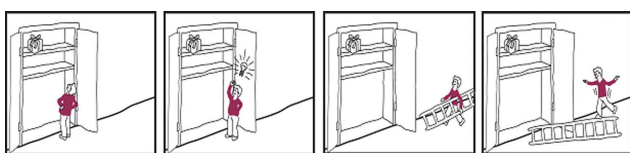
Affective empathy positive



Cognitive empathy logical



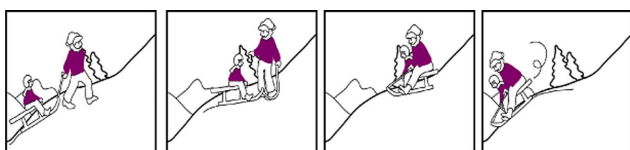
Cognitive empathy non-logical



Neutral story with one person



Neutral story with two persons



WRITTEN VERSIONS FOR THE AUDITORY STORIES, READ AND RECORDED BY AN ACTRESS, GERMAN ORIGINAL AND ENGLISH TRANSLATION

Affective empathy positive

Tim steht vor einer Bude und schleckt sein Eis.
Ein Mann kommt den Weg entlang gelaufen.
Er übersieht den Jungen und rempelt ihn so an, dass Tim das Eis aus der Hand fällt.
Der Mann schenkt Tim ein neues Eis.

Tim is standing in front of an ice-cream booth and licks his ice-cream
A man comes walking along.
He overlooks the boy and bumps into him, so that the ice-cream falls down from Tim's hand.
The man buys a new ice-cream for Tim.

Affective empathy negative

Lisa und Tom spielen zusammen auf dem Spielplatz und ihre Mutter schaut von einer Bank aus zu.
Tom steht vor der Rutsche, die Lisa gerade heruntergerutscht kommt.
Lisa rutscht Tom in den Rücken, beide fallen hin und weinen.
Die Mutter steht von der Bank auf und geht davon.

Lisa and Tom are playing together on the playground, their mother watches them from a bench.
Tom is standing in front of the slide, which Lisa is just sliding down.
Lisa slides into Tom's back, both of them fall down and start crying.
Their mother gets up from the bench and walks away.

Cognitive empathy logical

Anna läuft zu einem Apfelbaum an dem viele Äpfel hängen
Sie möchte einen Apfel pflücken aber kommt nicht dran
Sie holt sich eine Kiste und schiebt sie unter den Baum
Sie stellt sich auf die Kiste und pflückt einen Apfel

Anna walks towards an apple-tree full of apples.
She wants to pick an apple, but cannot reach any.
She fetches a box and puts it under the tree (gets places).
She steps up on the box and picks an apple

Cognitive empathy unlogical

Jannis entdeckt im obersten Fach des Schrankes seiner Eltern ein Geschenk.
Nach einigem Überlegen geht ihm ein Licht auf, wie er wie er daran kommen könnte.
Er holt sich eine lange Leiter.
Jannis legt die Leiter auf den Boden und beginnt auf ihr zu balancieren

Jannis discovers a present in the top shelf of his parents' closet.
After thinking a while, he gets an idea how to get it.
He fetches a long ladder.
Jannis puts the ladder on the ground and starts to walk on it.

Neutral with one person

Olli sieht eine Schaukel an einem Baum hängen
Er zieht die Schaukel zu sich
Er klettert auf den untersten Ast, damit er richtig Schwung holen kann
Mit Schwung fängt er an zu schaukeln

Olli sees a swing hanging on a tree.
He pulls the swing towards himself.
He climbs onto the lowest branch, to get a good start.
He begins to swing.

Neutral with two persons

Der Vater zieht Jan auf dem Schlitten den Berg hinauf.
Die beide kommen oben an.
Der Vater setzt sich zu Jan auf den Schlitten.
Sie rodeln gemeinsam den Berg hinunter.

The father pulls Jan on the sleigh up the hill.
They arrive at the top of the mountain.
The father sits down on the sleigh with Jan.
They sled down the mountain together.