Optical Brain Imaging Reveals General Auditory and Language-Specific Processing in Early Infant Development

Yasuyo Minagawa-Kawai1-4, Heather van der Lely4,5, Franck Ramus1, Yutaka Sato2, Reiko Mazuka2,6 and Emmanuel Dupoux1

1Laboratoire de Sciences Cognitives et Psycholinguistiques, ENS--DEE--EHESS--CNRS, Paris 75005, France, 2Laboratory for Language Development, Brain Science Institute, RIKEN, Wako, Saitama 351-0198, Japan, 3Graduate School of Humanities and Social Sciences, Keio University, Miwa-ku, Tokyo 108-8345, Japan, 4Centre for Developmental Language Disorders and Cognitive Neuroscience, University College London, London WC1N 1PF, UK, 5Department of Psychology, Harvard University, Cambridge, MA 02138, USA and 6Department of Psychology and Neuroscience, Duke University, Durham, NC 27708, USA

This study uses near-infrared spectroscopy in young infants in order to elucidate the nature of functional cerebral processing for speech. Previous imaging studies of infants' speech perception revealed lateralized responses to native language. However, it is unclear if these activations were due to language per se rather than to some low-level acoustic correlate of spoken language. Here we compare native (L1) and non-native (L2) languages with 3 different non-speech conditions including emotional voices, monkey calls, and phase scrambled sounds that provide more stringent controls. Hemodynamic responses to these stimuli were measured in the temporal areas of Japanese 4-month-old infants. The results show clear left-lateralized responses to speech, prominent to L1, as opposed to various activation patterns in the non-speech conditions. Furthermore, implementing a new analysis method designed for infants, we discovered a slower hemodynamic time course in awake infants. Our results are largely explained by signal-driven auditory processing. However, stronger activations to L1 than to L2 indicate a language-specific neural factor that modulates these responses. This study is the first to discover a significantly higher sensitivity to L1 in 4-month-olds and reveals a neural precursor of the functional specialization for the higher cognitive network.

Keywords: emotion, infant, laterality, NIRS, speech perception

Introduction

A i creasi g u ber cer of curoy agi g studies have docu
ted that the hu a la guage yst sy ov les a large et
work of perisylvia areas preo i a tly lateralized i the left
he isphere (Bra che a et al. 1964; Oje a et al. 1989; Hickok a d Poeppl 2000; Friederici 2002). Yet, the develop e tal
process u derly g such cerebral specialization for the guage re ai s co traversial. This is pri arly due to the tech ical
difficulties of apply g brai i agi g tech iqes to you g i fa ts. The i productio ut ultica ear i frad spec-
trosco py (NIRS) 10 years ago provided a pro isi g tech iq e i this e ergi g field of develop e tal euruscie ce (for a review, see Mi agawa-Kawai et al. 2008). Studies usi g NIRS have explored the develop e o the i fa t brai for several cog ative abilities i cludi g social cog ito (Gross a et al. 2008; Lloyd-Fox et al. 2009; Mi agawa-Kawai et al. 2009), object recog ito (Wata abe et al. 2008; Wilcox et al. 2008, 2009), a d speech processi g (Pe a et al. 2003; Hoe ae et al. 2006, 2007; Borfeld et al. 2007; 2009; Mi agawa-Kawai et al. 2007; Ger vae et al. 2008; Naka o et al. 2009; Telke eyer et al. 2009). NIRS is o i vasive, portable, a d sca er oise free a d has the adve tage of provi g reliab laocal-
izatio of cerebral respo ses to sou d sti uli, aki g it ideally
suitable for assessi g lateralizatio i fa ts. I the pres e t
study, we utilise NIRS to revisit the issue of brai lateralizatio
for speech a d o spee ch sou ds i 4-0 th-old i fa ts.

Neuroni agi g studi es o spee ch processi g i early i fa cy are gradually e ergi g. Usi g NIRS, Pe a et al. (2003) reported that ewbo i fa ts braal respo ses were stro ger i the left
t e poral areas for se te ces i the ater al la guage (L1) tha for the a sti uli played backwa rd (BW). Left
do i a ce for L1 speech processi g was also observed i older i fa ts aged 6-9 o ths by usi g both forward (FW)
speech a d sile t co tr l to get hether si al sti uli (Bortfeld et al. 2009). A fu ctio al ag etic reso a ce i agi g (MRI)
study of 3-0 th-old i fa ts (Dehae e-La bertetz et al. 2002) also reported do i a ce i the left te poral areas for both FW a d BW spee ch, with ore activatio for FW as co pared
with BW spee ch o ly i the left a gular gyrus. Sato, Hirabayashi et al. (2006) exte ded these fi ds by usi g NIRS to co pare braal
respo ses to L1 with those to a foreig la guage (L2) i ewbo i fa ts. They fou d that exclusively i the L1 co ditio , activatio s were sig ifica tly stro ger for FW tha
for BW spee ch i the left te poral area. However, for these
ewbo i fa ts, there were o sig ifica differ ent i activatio a plitudes betwee L1 a d L2. It is likely i the
course of la guage develop e t, L1 specific cerebral respo ses will e erge, but as yet, o study has i vestigated the differ e ce betwee L1 a d L2 activatio s i older i fa ts. Our study ai s to
fill this gap i the literatu re.

Further ore, although these early brai sig als i the left
he isphere for la guage a d BW spee ch are i trigu g, it is still
u clear how to relate these activatio s to la guage processi g
per se. O e of the plausible expla atio s for the left-do i a t
respo ses to spee ch is the acousti c properties of the sti uli: spee ch-co tai i g seg e ts such as co so a ts a d o vel o
ay e ha ce leftward braal activatio s due to their rapid
acousti c tra sitio s. Several studies (Zatorre a d Beli 2001; Boe io et al. 2005; Schü wies er et al. 2005; Jaiso et al. 2006) have revealed that the left he isphere is prefere tially
i volved i processi g rapid spectr o poral ch a ges, such as the
for a t tra sitio s i pho e es, whereas the right he isphere is ore e gaged i slow cha ges, such as those i the prosody. Based o this view, a left lateralizatio i you g i fa ts for la guage could be accou ted for by the acousti c
properties of spee ch sou ds that co vai rapidly cha gi g
li guistic seg e ts. A other plausible expla atio for stro ger
activatio s to BW tha a BW spee ch is the utural ess of the
sou ds; BW speech can be physically produced by a hu a vocal tract, a d so it is ot biologically atural. Specifically, BW speech differs fro FW speech i low-level acoustic characteristics: BW speech co tai s reversed e velope profiles that are i possible to obtai fro a vocal tract or a y atural physical syste . The proble is that the differenc e between causally possible a d i possible physical sou ds could be e coded early i the auditory pathway. As highlighted by Galbraith et al. (2004), bras ste-evoked respo ses differ betwee FW a d BW speech. Further ore, based o a i for atio theoretical odeli g study, S ith a d Lewicki (2006) clai ed that the spectrote poral properties of the cochlear code are opti al for atural sou ds a d FW speech, but ot for BW speech. If so, differenc e i activatio a plitude between FW a d BW speech, such as those fou d by Pe¨a et al. (2003) a d Sato, Hirabayashi et al. (2006), ight reflect low-level acoustic properties, rather th la guage processi g per se. Thus, usi g BW alo e a co a co trol co dito s for induc ing BW speech properties is pote tially proble atic. Cerebral respo ses to the FW sti uli also eed furth e i vestigatio , si ce it is possible that the hu a brai respo ds to voci zatio s rather th la to la guage per se. Although adult i agi g studies have co siste tly exhbitd greater cerebrual activatio s to hu a voci zatio s th a to thier a i al calls such as o keys a d birds (Fecteau et al. 2004; Hashi oto et al. 2006; vo Kriegstei et al., 2007), our study has exa i ed wheth e this species-specific trait exists at the begin i g of hu a develop e t.

Our pri cipal ai is to elucidate the ature of the guage lateralizatio a d the effects of L1 versus L2 i you g i fa ts. Buildi g o previous studies, we focus o i fa ts of 4 o ths old to test L1 specificit y i thier brai s. Nazzi et al. (2000) reported that i fa ts of 5 o ths old are able to behaviorally discri i ate a y la guage fro their L1 i co trast to ewbors who co fuse L1 a d L2 that have the sa e la guage rhythm s as each other. Because eural sig atures te d to appear earlier th behavioral a fjestatio s (Germain 2008), we predicted that the c erge ce of L1 specificit y iste i g to co ected speech ey at 4 o ths old. We therefore easured left a d right te poral activatio s to L1 (Japa ese) a d L2 (E glish) with NIRS. We exa i ed the possible factors that trigger stro g activatio s by e plo y g rigorous co trol co dito s. I stead of usi g BW speech, we e plo yed 3 differenc e sou d sti uli that vary i acoustic properties a d biological factors: 1) e otio al sou ds, which are produced by a hu a vocal tract but lack the seg e tal/ fast variatio structure of speech; 2) o key voci zatio s, which are produced by a o hu a vocal tract (Ghaza far a d Re dall 2008); a d 3) scra bled sou ds that are totally artificial o voci zatio s but are atech f or e gery a d lo g-ter spectrum with the other 4 co dito s. Usi g differenc e ki ds of co trols (see Table 1) elucidate the specificit y of the la guage fu ctio e t ul a relation to acoustic properties of the sti uli, species specificit y, a d voci zatio specificit y o left-dio a t brai activatio s.

Typically, a block desig is used i NIRS for sti ulus prese tatio , whereby relatively large brai respo se ca be obtai ed (Bortfeld et al. 2007, 2009; Wilcox et al. 2008; Mi agawa-Kawai et al. 2009). However, i order to test all these crucial differenc e t co dito s with the sa e i fa t, it is essencary for this study to use a fast eve t-relat ed desig (Zarath et al. 1997; Fristo et al. 1998). This e ables us to i crease the trial repetitio t i e a d allows us to take a ew approac h to the a lysis of these data fro the eve t-relat ed paradigm for i fa t NIRS. Specifically, we e plo yed a ge eral li ear odelm (GLM) approach (Fristo et al. 1994, 1995), which uses a odel of the expected he ody a ic respo se fu ctio s. However, it is questio able whether the sta dard adult HRF odelm is valid for i fa ts as the brai physiolog y of 4 o th-olds is quite differenc e t th that of adults, a d i deed i fa ts show a phase delay i thier evoked he ody a ic respo ses (Schroeter et al. 2004; Shi ada a d Hiraki 2006). We therefore e plo yed a tech ique based o fi ite i pulse respo se (FI) fu ctio s (Fristo et al. 1995) in order to reco strut the HRF of the i fa ts before e gagi g i GLM odeli g.

### Materials and Methods

#### Participants

We studied 12 full-ter i fa ts (4 girls a d 8 boys) with or al birth weights, belo g to o oli gual Japa ese fa ilies. At the ti e of testi g, which was co ducted at the BRKEN Brai Sci ce I stitute (BSI), their average age was 128 days (sta dard deviatio [SD] = 13.3). A additio al 18 i fa ts were tested, but they were excluded fro the fi al sa ple after co sideratio of artifacts i the data due to head ove e ts, fussi ess, a d hair obstructio (N = 10); refusal to wear the NIRS holder (N = 4); poor positiio i g of the probe (N = 1); a d sleepi g or drowsi ess duri g testi g (N = 5). We e plo yed strict criteria to deter i e our fi al data set because oise-free he oglo (HB)sig als are essencary for a reliable GLM alysis. I accorda ce with our criteria, first, we excluded all the blocks co ta i ated with otio artifacts (for details, see Data A alyses); we also discarded whole blocks obta ied fo sleepi g i fa ts. The , o ly participa ts for who at least 8 blocks survived (of 12 blocks) without otio artifacts for each of the 5 co dito s were i cluded i the fi al data set (see sti uli for details of block a d co dito s ). This e sured that the fi al data set co tai ed ore th a 40 blocks for each i fa t. Our experi e tal desig , with its 5 co dito s, resulted i a prese tatio ti e that was lo gert e ti h th a is typically fou d i i fa t studies. As a result, there were a y participa ts who did ot co plete all the sessio s or who were u able to eet our strict criteria for i chusio i the data set as do cu e ted above. Pare ts provided i for ed co se i t co plia ce with a protocol approved by the ethics co ittee of Rike BSI (Wako 3rd-16-12 (10)).

#### NIRS Recording

This study used NIRS (ETG-4000, Hitachi Medical Co., Tokyo, Japa ), which e ures the HB co cc tratio cha ges of the optical paths i the brai betwee the earrest pairs of i cide t a d detectio probes separated by 3 c o the scalp surface (Wata abe et al. 1996; Ya ashta et al. 1996). This separatio e ables us to euse he ody a ic cha ges i the brai 25-3 c deep fro the head surface, which correspod ds to the gray utter o the outer surface of the brai (Fukui et al. 2003). The i str u e t e its 2 wavele gths (ca. 695 a d 830 ) of co ti uous ears fi ared lasers, oculat ed at differenc e t frequenc cies depe di g o the cha els a d the wavele gths a d detected with the sharp freque cy filters of lock-i a pilifiers (Wata abe et al. 1996).

Five i cide t a d 4 detectio probes arra ged i a tra sfor ed 3 3 3 grid (12 cha els, Fig. 1) were fitte d o the te poral a d fro tal areas

---

**Table 1**

<table>
<thead>
<tr>
<th>Stimulus type</th>
<th>Native language</th>
<th>Linguistic segment</th>
<th>Same species</th>
<th>Vocally produced</th>
<th>Sound</th>
</tr>
</thead>
<tbody>
<tr>
<td>Language (L1)</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Language (L2)</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Emotional voice</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Monkey call</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Scrambled</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>
of each side of the head usi g the i ter atio al 10-20 syste . Specifically, we placed the id-botto detector to T3 a d T4 positiio for each side a d the asio -i o li e a d d the vertet- to-tragus li e were used either for the horizo tal or the vertical axis (Fig. 1). Brai regio s correspo di g to NIRS cha el s were esti usi g the virtual registratio (oka to et o 2004; t suzuki et o 2007) by taki g the s all head size of 4 o th-olds i to accou t.

Stimuli
The auditory sti uli co prised 5 co dito s: ative speech (Japa ese), o - ative speech (British E glis h), e o - atio voices, o key calls (Macaque), a d scra bled cro ts. All speec h sti uli were ade of a co cate atio of short se ce ces reco rded fro fil dialogues or a specch database (Corpus of S po to aus Japa e) (mackaw et o 2004) by either fe ale a d ale speakers. The e oto al voices were huu a vocalizatio s with o li guastic co te e, et e with a positive (e.g., ad iratio a d laugi g) or with a egatio e oto al vale ce (e.g., cryi g a d sigh). Mo key calls were si larly divided i to those with positive (e.g., coos, gir eys, a d har o ic arches) a d those with egatio e oto al vale ce (e.g., scra e a d shi ll barks). The o key calls were take fro a sou d library co piled by Marc D. Hauser (Harvard U iversity; sti uli reco rded fro the Isla d of Cayo Ti ago, Puerto Rico, USA). Spectra scra bli g was ap plied to the typer of sti uli above to ake a atched scra bled cro ts (Japa ese, E glis h, e oto al voices, a d o key calls). Spectra scra bli g sou d s that are totally u i telligible sou ds co sot of the ra do a d i depe de t excha ge of ha ds of spectra withi a give sti uli. More precisely, they were sy thesized by processi g all i individual sti uli through a ga a to e filter ba k with 64 cha els. 1 each cha el, the sig al was wi dowed with overlappi g ha i g wi dow of 25- s duratio . The wi dow were the shuffled ra do ly withi a cha el a d d displa ed withi a ti e a d ra ge of 2500 s arou d its origi al te poral positi o. The scra bled sou d sou ds were fi ally ob tai ed by putti g all fereqe cy cha els back together. Thus, the average a plitude a d lo g-te weather specy of scra bled cro ts is exactl y the sa e a s for their origi al sti uli. All sti uli ra ged fro 700 a 1200 s i le g h a d equal he rgy (root ea square). These i individual sti uli were used to co struct sti uli blocks that lasted 10 s o average for each co dito . O e block co ta i ed a co cate atio of 9-12 i individual sti uli of the sa e type (short se ce ces i Japa ese, Mo key vocalizatio s, etc.) with 200- s sile ce between each sti uli. Fi ally, the blocks were asse bled i to sessio s. O e sessio co ta i ed 4 blocks of each co dito (20 blocks per sessio ) that were prese ted to the i fa ti a ra do orde r with a sile ce period (8-14 s) be twee between the blocks. All i fa ts had at least 3 sessio s, each of which lasted for 6.5-6.8 i.

Figure 1. Probe arrangement to attach to the bilateral temporal areas. Black circles on the brain indicate channel positions. Red probes are emitters, and blue ones are detectors. Black circles are NIRS channels between one emitter and detector.
activatio i the right side for the o speech vocal co ditio .
Separate a alysis for each co ditio revealed various activa-
ti o i the two he ispheres (Table 2). While a stro g
respo se was o ly fou d i the right auditory/STG for the
e oti o al vocalizatio , sig ifica t cha els were restric ted to
the left side for the Japa ese, E glish, a d scra bled sou ds
c o ditio s. The o key calls broadly activated bilateral
te poral areas. I cludi g the a terio STG. I order to co pare
brai activatio s for ative a d o - ative la guages, a alysis of
varia ce with factors of laterality a d la guage was perfor ed
(Fig. 4). The results revealed sig ifica t ai effects of laterality
($F_{1,44} = 4.41, P = 0.047$) a d la guage ($F_{1,44} = 5.46, P = 0.029$).

Discussion

This study exa i ed the develop e tal specificity of la guage
processi g i i fa ts brai s by usi g a carefully co trolled set
of sti uli as well as a ew a alysis ethod for i fa t data. The
results revealed localized brai activatio s with larger left-
 lateralyzed respo ses to speech as opposed to bilateral or
rightward activatio for various co trol sti uli i the early
i fa t developi g brai. Our data provide ew evide ce i at
least 3 respects: 1) they reveal fu ctio al asy etries for
speech a d other o speech vocalizatio s such as e oti o al
voices i you g i fa ts; 2) they de o strate europhysiologi-
cal evidence ce for la guage-specific eural plasticity i 4 o th-
olds by co pari g the brai respo ses to L1 a d L2; a d 3)
they u cover a slower he ody a ic ti e course i awake
i fa ts brai s co pared with that of adults.

Before we focus o the left do i a ce i speech processi g,
we first review the results of the o speech co ditio s. For the
e oti o co ditio , a sig ifica t activatio was observed o ly i
the right te poral area. This rightward te poral activatio is
c o siste t with a y adult i agi g studies o vocal e oti o s
(Meyer et al. 2002, 2004; Wildgruber et al. 2002; Wiethoff et al.
2008). So far, o ly a few studies have exa i ed the eual basis
of the perceptio of e oti o al prosody i i fa ts. For i sta ce,
7 o th-olds showed differe t ephysiologi-cal respo ses to
diffe t e oti o al voices, with a greater egative co po-
et o to a gy voices (Gross a et al. 2005). Although o
previous study has revealed the specific brai regio associated
with e oti o processi g i you g i fa ts, our study de o-
strated a tre d for the right do i a ce i processi g e oti o al
prosody i 4 o th-olds. Our results for the o speech
c o ditio s for i fa ts ge eraly replicat e previous i agi g
studies with adults. However, we co sider it likely that these
activatio s chiefly reflect processi g the acoustic properties i
the differe t sti ulus co ditio s. More co crety, sou d strea s
with seg e tal features or fast spectral cha ges ay
ea ha ce activatio s i the left he isphere, whereas prosodic
pitch co tours ay e ha ce activatio s i the right he-
isphere a d that this is respo sible for the left/right do i a ce
for the differe t sti uli (Zatorre a d Beli 2001; Poeppel 2003;
Zatorre a d Ga 2008). We observed that the e oti o al
voices that have slow prosodici cha ges a d fewer li guistic
seg e ts evoked rightward activatio s, while the scra bled
sou ds with a lot of rapidly cha gi g seg e ts elicit ed
sig ifica t leftward activatio s.

Further to the acoustic factor, our co trol sti uli e a bled us
to disti guish 2 additi o factors: species specificiti y a d
vocalizatio. I respo se to the o key vocalizatio s, hu a
4 o th-olds showed stro g a d broad activatio s i the

Figure 2. (A) FIR results in 24 channels for all the conditions. Dashed line indicates the canonical HRF. (B) Time course of Hb changes elicited by the FIR analyses. Dashed line indicates the canonical HRF. Zero point in the time line is a stimulus onset.

Brain Responses to Speech and Nonspeech Stimuli

We perfo r ed a alysis with a progressively refi ned level of
detail. I the first a alysis, all 5 sti ulus types were grouped
together i a i gle sou d co ditio . I the seco d a alysis, we
declared 5 co ditio s: Japa ese, E glish, e oti o , o key, a d
scra bled. All these a alyses used GLM odels of the oxy-Hb
sig al for each co ditio , cha el, a d i fa t separately, usi g
regressors of a boxcar, t i e locked to each sti ulus o set i
a give co ditio , a d co volved with the i fa t HRF odel.
The regressio coeficie ts ($\beta$) for each co ditio , cha el, a d i
fa t were the tested agai st zero ($t$-test across i fa ts)
a d corrected for multiple co pariso s across cha els by
usi g the false discovery rate ethod. The resulti g $t$-tests,
were corrected a d corrected, are show i Table 2. The global
sou d co ditio sig ifica tly activated the bilateral auditory
areas. All of the 5 sti uli activated various regio s i cludi g the
superior te poral gyri (STG) a d i ferio fral gyri as i
dicated by a ap of across-i fa t sig ifica t $\beta$s for each
c o ditio (Fig. 3). Although activatio patter s arou d the
perisylvia areas were si ilar for speech a d o speech vocal
sou ds, activatio levels were higher i the left side for the
speech co ditio i co trast to the larger a d broader

Page 4 of 8 Optical Brain Imaging of Infant Development Minagawa-Kawai et al.
Bilateral temporal areas. Exclusively for this condition, there were significant activations around the bilateral anterior STG. These results are consistent with those of adult fMRI studies examining perception of animal calls (Fecteau et al. 2004; Hashimoto et al. 2006; von Kriegstein et al. 2007). However, human adults typically show weaker brain responses to animal calls than to human vocalizations. Our finding, showing large responses to monkey calls, indicates that 4 month-olds are still sensitive to vocalizations of other species, which is characterized by higher pitch and resonance (Ghazanfar and Rendall 2008). This is congruent with their early visual abilities to discriminate monkey faces in 6 month-olds (Pascalis et al. 2002). Furthermore, recent behavioral results (Vouloumanos et al. 2010) revealed newborns’ preferences for both human speech and monkey calls that sounded like human vocalization. These initial abilities may be gradually narrowed during development within the first year of life (Pascalis et al. 2002), as infants tune their sensitivity to conspecific stimuli or their ambient environment such as L1. In contrast to brain responses to vocalizations including monkey calls, artificial scrambled sounds showed only one significant channel (Table 2). Even though vocal versus scrambled conditions have identical long-term spectrum, artificial nonvocal sounds induced fewer activations than the vocal sounds did. This suggests that 4 month-olds’ brains are more responsive to biologically natural sounds than to unnatural sounds. This neuronal evidence may explain weaker brain activations for BW speech in the previous imaging studies (Peneta l. 2003; Sato, Hirabayashi et al. 2006). The results are also consistent with infants’ behavioral responses showing a preference for speech sounds than for their analogous artificial tones from birth to 6.5 months (Vouloumanos and Werker 2004, 2007). Finally, it should be noted that these data from the control stimuli should be taken with careful consideration because although each of these stimuli elicited a different pattern of significant activations, we lacked the power to demonstrate that these patterns of activations were statistically different from one another.

### Table 2

<table>
<thead>
<tr>
<th>Conditions</th>
<th>Brain region</th>
<th>Channel</th>
<th>Side</th>
<th>β</th>
<th>SE</th>
<th>t</th>
<th>P (unc)</th>
<th>P (FDR)</th>
</tr>
</thead>
<tbody>
<tr>
<td>All sounds</td>
<td>Auditory</td>
<td>9</td>
<td>R</td>
<td>0.07</td>
<td>0.08</td>
<td>4.87</td>
<td>0.0003</td>
<td>0.008</td>
</tr>
<tr>
<td>All sounds</td>
<td>Auditory</td>
<td>21</td>
<td>L</td>
<td>0.07</td>
<td>0.09</td>
<td>3.53</td>
<td>0.0004</td>
<td>0.046</td>
</tr>
<tr>
<td>Vocal</td>
<td>pSTG/SMG</td>
<td>6</td>
<td>R</td>
<td>0.05</td>
<td>0.07</td>
<td>3.0</td>
<td>0.01</td>
<td>0.065</td>
</tr>
<tr>
<td>Vocal</td>
<td>Auditory</td>
<td>9</td>
<td>R</td>
<td>0.1</td>
<td>0.11</td>
<td>5.11</td>
<td>0.0002</td>
<td>0.005</td>
</tr>
<tr>
<td>Vocal</td>
<td>Auditory</td>
<td>21</td>
<td>L</td>
<td>0.1</td>
<td>0.13</td>
<td>3.29</td>
<td>0.006</td>
<td>0.085</td>
</tr>
<tr>
<td>JP</td>
<td>Precentral</td>
<td>16</td>
<td>L</td>
<td>0.06</td>
<td>0.08</td>
<td>3.49</td>
<td>0.004</td>
<td>0.033</td>
</tr>
<tr>
<td>JP</td>
<td>pSTG/SMG</td>
<td>19</td>
<td>L</td>
<td>0.09</td>
<td>0.12</td>
<td>3.68</td>
<td>0.003</td>
<td>0.033</td>
</tr>
<tr>
<td>JP</td>
<td>Auditory</td>
<td>21</td>
<td>L</td>
<td>0.15</td>
<td>0.16</td>
<td>7.03</td>
<td>0.0001</td>
<td>0.002</td>
</tr>
<tr>
<td>ENG</td>
<td>pSTG/SMG</td>
<td>19</td>
<td>L</td>
<td>0.06</td>
<td>0.09</td>
<td>2.97</td>
<td>0.011</td>
<td>0.131</td>
</tr>
<tr>
<td>ENG</td>
<td>Auditory</td>
<td>21</td>
<td>L</td>
<td>0.08</td>
<td>0.11</td>
<td>3.64</td>
<td>0.003</td>
<td>0.075</td>
</tr>
<tr>
<td>EM</td>
<td>Auditory</td>
<td>9</td>
<td>R</td>
<td>0.11</td>
<td>0.14</td>
<td>3.61</td>
<td>0.003</td>
<td>0.078</td>
</tr>
<tr>
<td>MN</td>
<td>pSTG/SMG</td>
<td>6</td>
<td>R</td>
<td>0.09</td>
<td>0.12</td>
<td>3.29</td>
<td>0.006</td>
<td>0.029</td>
</tr>
<tr>
<td>MN</td>
<td>Auditory</td>
<td>9</td>
<td>R</td>
<td>0.09</td>
<td>0.11</td>
<td>4.99</td>
<td>0.0003</td>
<td>0.005</td>
</tr>
<tr>
<td>MN</td>
<td>pSTG/SMG</td>
<td>18</td>
<td>L</td>
<td>0.07</td>
<td>0.09</td>
<td>4.5</td>
<td>0.0001</td>
<td>0.005</td>
</tr>
<tr>
<td>MN</td>
<td>pSTG/SMG</td>
<td>19</td>
<td>L</td>
<td>0.12</td>
<td>0.15</td>
<td>4.07</td>
<td>0.001</td>
<td>0.008</td>
</tr>
<tr>
<td>MN</td>
<td>Auditory</td>
<td>21</td>
<td>L</td>
<td>0.11</td>
<td>0.14</td>
<td>4.5</td>
<td>0.0001</td>
<td>0.005</td>
</tr>
<tr>
<td>SC</td>
<td>Auditory</td>
<td>21</td>
<td>L</td>
<td>0.1</td>
<td>0.14</td>
<td>3.4</td>
<td>0.005</td>
<td>0.116</td>
</tr>
</tbody>
</table>

JP = Japanese, ENG = English, EM = emotional voices, MN = monkey calls, SC = scrambled sounds, SE = standard error, FDR = false discovery rate, IFG = inferior frontal gyrus, SMG = supramarginal gyrus, pSTG = posterior part of superior temporal gyrus, aSTG = anterior part of STG, R = right, L = left, unc = uncorrected.

---

**Figure 3.** Brain activations in different 5 conditions for 24 channels in the left and right side. P values are corrected for multiple comparisons.

**Figure 4.** Different brain activation levels for native (L1) and foreign (L2) languages in the left and right auditory areas.
specific language acquisition. In the course of language activation may be considered a neural precursor for L1-acquisition. In this regard, this enhanced left-dominant and grammatical structures in the later stages of language properties, which is necessary to integrate phonemic, lexical, structures, such as prosody specific to L1 (Mehler et al. 1988; month-old infants cannot process fine linguistic content such as in NIRS studies (Schroeter et al. 2004; Schilbach et al. 2008) a d esti ated the peak late cy by pooli g all the sou d co ditio s together. We fou d a respo se re arkyb close to that of the adult, but delayed by approx-i ately 3 s. If we had used a HRF without a delay, the peak late cy of the he ody a ic respo ses would have u fairly i flue ce the brai activatio levels. Such a delayed ti e course could be partly derived fro i ature vascular regulatio a d yeii atio i i fa ts (Chuga i et al. 1987). However, a other possible reaso for the delay could be atte tio al shifts due to the particular paradigm we used: i our study, the i fa ts atte tio was captured by visual sti uli (i.e., sile t toys), so that it could have take the a few secos ds to otice the sou ds played to the , i duci g a delay i auditory-related activatio . Si ce o clear details have bee provided for th e i fa ts vascular echa is i relatio to eural activities, ore physiological studies are required for a co clusiva atio atio . However, our ethos provides a pri cipled way to deter i e the odel for GLM i fa ts i steady of usi g the ca o ical HRF, or arbitrary wi downs of a alysis for i fa ts euroi agi g, give that variatio s i age, sti uli, tasks, a d wakel es ay i flue ce the respo se late cy (Meek et al. 1998; Taga et al. 2003; Shi ada a d Hiraki 2006; Mi agawa-Kawai et al. 2008).

By exa i g he ody a ic respo ses i 4 o th-olds to speech a d various o speech sou ds, the curre t NIRS study clearly showed a left-lateralized cerebral basis for speech processing. Although these activation patterns could be ostly explai ed by sig al-drive hypotheses of brai lateralizatio , we revealed a li guistic factor that ay co to the activatio of the la guage ewerk o the left side a d that provides a parsi o ious expla atio . Moreover, we have provided data releva t to the brai physiology i you i fa ts by showi g slow late cy of fu ctio al he ody a ic respo ses. Overall, this study illustrated that 4 o th-olds' eural develop e t is at a stage where their processi g of speech is based o a i teractio betwee ge eric auditory syste s a d lear i g echa is s that start to extract regularities regardi g their ative la guage.

**Funding**

Europea U io 's sixth-fra ework progra ( euro al origi s of la guage a d co u icatio : NEUROCOM (Project o. 012738); Gra t-i -Aid for Scie tific Research (A) (Project o. 21682002); Welco e Trust (063713 to H.L.).

**Notes**

We tha k K. Shirasawa, Y. Sogabe, a d other staf i the la guage develop e t laboratory; RIKEN BSI for assista ce with NIRS testi g; a d A. Shestakova, E. Kush ere ko, Y. J. Cha g, J. Meek, a d other e bers of the DLDCN Ce tre for their assista ce i co ducti g the pilot study for this work at the DLDCN Ce tre, Lo do . We also tha k M. D. Hauser for ki dly provi g the o key vocalizatio sti uli. **Conflict of Interest:** No e declared.
Wata abe H, Ho ae F, Taga G. 2010. Ge ral to specific develop e t of fu ctio al activatio i the cerebral cortexes of 2- to 3- o th-old i fa ts. Neuro iage. 50:1536-1544.
Wilcox T, Bortfeld H, Woods R, Wruck E, Ar stro g J, Boas D. 2009. He ody a ic respo se to featural cha ges i the i fa t cortex duri g the processi g of featural a d spatiote poral i for atio . Neuropsychologia. 47:657-662.