Cerebral lateralization and early speech acquisition: A developmental scenario

Yasuyo Minagawa-Kawai [a,b,c,*], Alejandrina Cristià [a], Emmanuel Dupoux [a]

[a] Laboratoire de Sciences Cognitives et Psycholinguistique, ENS-DEC-EHESS-CNRS, 29 rue d'Ulm, Paris 75005, France
[b] Keio University, Graduate School of Human Relations, 2-15-45 Mita, Minato-ku, Tokyo 108-8345, Japan
[c] Global COE program, Centre for Advanced Research on Logic and Sensibility, Keio University, 3-1-7 Mita, Minato-ku, Tokyo 108-0073, Japan

ABSTRACT

During the past ten years, research using Near-infrared Spectroscopy (NIRS) to study the developing brain has provided groundbreaking evidence of brain functions in infants. This paper presents a theoretically oriented review of this wealth of evidence, summarizing recent NIRS data on language processing, without neglecting other neuroimaging or behavioral studies in infancy and adulthood. We review three competing classes of hypotheses (i.e. signal-driven, domain-driven, and learning biases hypotheses) regarding the causes of hemispheric specialization for speech processing. We assess the fit between each of these hypotheses and neuroimaging evidence in speech perception and show that none of the three hypotheses can account for the entire set of observations on its own. However, we argue that they provide a good fit when combined within a developmental perspective. According to our proposed scenario, lateralization for language emerges out of the interaction between pre-existing left–right biases in generic auditory processing (signal-driven hypothesis), and a left-hemisphere predominance of particular learning mechanisms (learning-biases hypothesis). As a result of this completed developmental process, the native language is represented in the left hemisphere predominantly. The integrated scenario enables to link infant and adult data, and points to many empirical avenues that need to be explored more systematically.

© 2011 Elsevier Ltd. All rights reserved.

Contents

* Corresponding author at: University, Graduate School of Human Relations, 2-15-45 Mita, Minato-ku, Tokyo 108-8345, Japan. Tel.: +81 0 3 5427 1156; fax: +81 0 3 5427 1209.

E-mail address: myasuyo@bea.hi-ho.ne.jp (Y. Minagawa-Kawai).

1. Introduction

The hemispheric asymmetries evidenced in language have long been the object of debate (Lenneberg, 1966; Zatorre and Gandour, 2008). Until recently, studying the development of cerebral lateralization in response to speech perception was nearly impossible for two reasons. First, available methods were not ideally suited to investigate lateralization accurately. Thus, the various approaches used, including neuropsychological observation, dichotic listening (Glanville et al., 1977; Bertoncini et al., 1989; Best et al., 1982; Vargha-Khadem and Corballis, 1979), and electroencephalography (EEG; Novak et al., 1989; Molfese and Molfese, 1988; Duclaux et al., 1991; Simos et al., 1997), had rather poor spatial resolution. Although these methods can in principle capture hemispheric lateralization, results from dichotic listening and early EEG studies on infants were often inconsistent showing, for **Q1** instance, right-dominant (Novak et al., 1989; Molfese and Molfese, 1989), left-dominant (Dehaene-Lambertz and Baillet, 1998) and symmetrical responses (Simos and Molfese, 1997) to phonemic contrasts. Nearly half a century after Lenneberg (1966), we are now much closer to understanding the development of the functional lateralization for speech, as the use of imaging techniques such as functional magnetic resonance (fMRI), multi-channel event-related potentials (ERP) and near-infrared spectroscopy (NIRS) provide more reliable evidence regarding the cerebral bases of language development (Werker and Yeung, 2005; Dehaene-Lambertz et al., 2006; Minagawa-Kawai et al., 2008; Obrig et al., 2010; Gervain et al., 2011). In particular, during the past ten years, research using NIRS to study the developing brain has rapidly expanded to provide crucial evidence for the emergence of lateralization. A second roadblock to the study of the development of functional lateralization for speech was that the biases driving this lateralization in adults were not fully understood. Nowadays, adults' imaging data for the cerebral bases of language is rapidly accumulating and the picture for functional cerebral lateralization in adults is clearer than the one before (see recent reviews in Tervaniemi and Hugdahl, 2003; Zatorre and Gandour, 2008). As these two roadblocks are removed, we are now in a position to provide a principled account of that development. In this paper, we review imaging studies on infants and adults to compare activations in the developing brain with the cerebral organization of the mature language system (e.g. Friederici, 2002; Scott and Johnsrude, 2003; Minagawa-Kawai et al., 2005).

To this end, we present three competing classes of hypotheses (i.e. signal-driven, domain-driven, and learning biases hypotheses) regarding the causes of hemispheric specialization for speech processing, and the adult neuroimaging data supporting them, in Section 2. We assess the fit between each of these classes of hypotheses and neuroimaging evidence in infant speech and non-speech perception in Section 3, and show that none of the three hypotheses can account for the entire set of observations on its own. However, we argue that they provide a good fit when combined within a developmental perspective. Based on this discussion, in Section 4 we propose a model where cerebral lateralization for language emerges out of the interaction between biases in general auditory processing and a left-hemisphere bias associated to certain learning subsystems recruited in language acquisition. It should be noted that when we speak of left/right dominance or left/right lateralization, we mean that the degree of activation is larger in one or the other hemisphere, not that activation is found exclusively in one of them.

2. Three hypotheses for language lateralization

It has long been known that the left and right hemispheres differ structurally in ways that could map onto functional differences, including larger micro-anatomical cell size, greater thickness of myelin, wider micro-columns, and larger spacing of macro-columns in the left hemisphere (Hayes and Lewis, 1993; Penhune et al., 1996; Seldon, 1981; Galuske et al., 2000). Furthermore, the patterns of connectivity across brain regions also differ between the two hemispheres, with a larger volume of fiber tracts in the arcuate fasciculus in the left hemisphere (e.g., Duffau, 2008). These differences have been hypothesized to enable the left hemisphere to function efficiently by implementing a large number of subsystems, which would facilitate or enable language acquisition and processing (Stephan et al., 2007; Friederici, 2009). Additionally, these differences between the hemispheres have sometimes been deemed one of the evolutionary innovations by which humans have come to develop a language system (Devlin et al., 2003). Importantly, such anatomical asymmetries are not as marked in non-human animals (e.g., Buxhoeveden et al., 2001). Additionally, left-lateralization is often reported at a number of linguistic levels, including syntactic processing and semantic access (for a review of recent data, see Price, 2010). Here, however, we will focus only on speech processing at the phonetic/phonological level, which in a majority of right-handed adults elicits left-dominant responses (e.g., Furuya and Mori, 2003; Turkeltaub and Coslett, 2010).

There are two leading hypotheses postulated to account for how the two hemispheres, with their different structures, could lead to the functional specialization for speech in the adult human brain. The signal-driven hypothesis puts a strong emphasis on the low level spectral or temporal

properties characteristic of speech sounds (Boemio et al., 2005; Jamison et al., 2006; Schonwiesner et al., 2005; Zatorre and Belin, 2001). Specifically, the left hemisphere is said to be preferentially involved in processing rapid durational changes such as those that distinguish phonemes, whereas the right hemisphere is more engaged in fine spectral processing such as that required for discrimination of slow pitch changes or emotional vocalizations. In contrast, the *domain-driven hypothesis* puts a strong emphasis on the fact that speech sounds are part of a highly complex communicative/expressive system specific to the human species, which recruits dedicated brain networks (e.g., Dehaene-Lambertz et al., 2005, 2010; Fodor, 1985). Specifically, this hypothesis predicts that language relevance, rather than the acoustic properties of a stimulus, underlie patterns of neural recruitment when processing sounds. A third view, which has received less attention lately, emphasizes to a larger extent the fact that speech processing is first and foremost the outcome of a learning experience; we call it the *learning-biases hypothesis*. We put forward one instantiation of the learning-bias hypothesis, according to which language acquisition recruits several specialized (but not necessarily domain-specific) learning subsystems (Ashby and O'Brien, 2005; Friederici et al., 2006; Zeithamova et al., 2008), each of them implicating distinct brain networks (as in Ashby and Ell, 2001). Specifically, we claim that the establishment of feature-based, categorical phonetic units and the extraction of words and rules on the basis of hierarchical and adjacent regularities require specific learning algorithms that are especially efficient in the left hemisphere and, as a result, speech perception comes to be left-lateralized as a function of experience.

We review each of these hypotheses in the light of neuroimaging evidence of speech and non-speech perception in human adults and non-human animals. In fact, one should recognize that there is considerable variation across authors on the precise formulation of these hypotheses, and we should rather refer to them as *classes* of hypotheses. However, for the purposes of this exposition, we will take into account the most extreme version of each of the three classes of hypotheses, without the intention of caricaturizing them. This strategy serves to evaluate the hypotheses in their strongest stance, even though it is clear that, within each group of researchers (or even within the same researcher), some combination of the three biases is expected. As will be evident in the final section, we agree that the right answer likely involves a combination of those hypotheses. We would like to, again, point out that this review is focused on the brain networks involved in the perception of speech sounds. While it is clear that other components of language (morphology, syntax, semantics), and other processing modalities (speech production) are also left-lateralized, we consider these components to fall outside the narrow scope of the present review.

2.1. The signal-driven hypothesis

2.1.1. Identifying features

Several studies using a variety of non-speech stimuli with fMRI or PET (positron emission tomography) relate differences in lateralization to differences in the low-level physical characteristics of stimuli, particularly along a temporal dimension ("fast" versus "slow"), but also along a spectral dimension ("simple" vs "complex"). Most studies document a clear asymmetry in the temporal cortex as a function of spectro-temporal features of the stimuli, with greater leftward responses to quickly changing spectral signals and more rightward responses to slowly modulated or spectrally rich signals (Jamison et al., 2006; Schonwiesner et al., 2005; Zatorre and Belin, 2001), although others report rather bilateral engagement when processing fast modulated stimuli (e.g., Belin et al., 1998; Boemio et al., 2005; Poeppel et al., 2008). The dichotomy between fast vs. slow temporal features resembles the well-established local vs. global dichotomy documented in the literature in the visual cognitive field (Ivry and Robertson, 1998; Koivisto and Revonsuo, 2004). Moreover, some neuropsychological and electrophysiological studies find similar asymmetries in response to local vs. global auditory changes (Peretz, 1990; Horvath et al., 2001). Finally, there is some evidence for signal-driven neural processing also in non-human animals. For instance, lesions in the right auditory cortex affect the discrimination of rising and falling tones in Mongolian gerbil (Wetzel et al., 1998) and in rats (Rybalko et al., 2006), while rapidly changing auditory stimuli are processed in the left temporal area of rats (Fitch et al., 1993: but see Fitch et al., 1994).

2.1.2. Adult data

How does the signal-driven hypothesis account for a predominant left lateralization for language? In fact, authors disagree on whether the relevant parameter involves spectral complexity, temporal complexity, or a combination of the two. Even among those who emphasize the temporal dimension primarily, the notion of fast/slow varies across authors and may therefore map onto different linguistic structures. As seen on Table 1, the durations of stimuli or the period of oscillations for 'fast signals' typically varies between 20 and 40 ms, whereas for 'slow signals' it varies between 150 and 300 ms. However, measurements of running speech show that segment duration typically fall in between the fast and the slow range: in French, stops like/b,k/last 77–112 ms; fricatives like/v,s/80–128 ms; sonorants like/m,j/55–65 ms; vowels between 72 and 121 ms (Duez, 2007). Other researchers emphasize the notion of spectral and temporal 'complexity' (rather than duration per se), captured through change over successive sampling windows. Indeed, some (e.g., Rosen, 1992) have proposed that acoustic landmarks of 20–50 ms could be sufficient for phoneme identification. However, a wealth of research shows that listeners integrate information over substantially longer windows. For instance, formant transitions and duration both influence vowel perception (Strange and Bohn, 1998), preceding and following vowels influence sibilant place of articulation (Nowak, 2006), the duration of a following vowel influences perception of articulation manner of a consonant (Miller and Liberman, 1979). In other words, the information relevant for the identification of a given phoneme is recovered from sam-

Table 1
Selected studies illustrating the different conceptions of signal-driven biases. All measures have been converted to durations in milliseconds.

Stimuli	Study	Left H Bias		Right H Bias	
Non-Speech		Fast tone/formant changes		Slow tone/formant changes	
	Belin et al. (1998)	Fixed duration 40 ms		Fixed duration 200 ms	
		Temporal complexity		Spectral complexity	
	Schönwiesner et al. (2005)	variable duration	5–20 ms	Fixed duration 33 ms	
	Zatorre and Belin (2001)	variable duration	21–667 ms	Fixed duration 667 ms	
		Small integration window		Large integration window	
	Poeppel (2003)	Window duration	20–40 ms	Window duration	150–250 ms
(none)		Gamma band spontaneous oscillation		Theta band spontaneous oscilation	
	Giraud et al. (2007)	Oscillation period	25–36 ms	Oscillation period	167–333 ms
Speech	Shankweiler and Studdert-Kennedy (1967), Haggard and Parkinson (1971), Ley and Bryden (1982), Zatorre et al. (1992) Furuya and Mori (2003)	Temporal coding of phonemes or words		Tonal pitch and prosody	
		Phoneme duration[a]	80 ms	Tone event[c]	80 ms
		Word duration[b]	200–300 ms	Sentential/emotional prosody[d]	1000–1800 ms

[a] In French (Duez, 2007), stops like/b,k/last 77–112 ms; fricatives like/v,s/80–128 ms; sonorants like/m,j/55–65 ms; vowels between 72 and 121 ms.
[b] Range computed over average word length in English, Japanese, Italian, French (Pellegrino, Coupé and Marcico 2007).
[c] Based on average vowel duration (see note 1).
[d] Based on average sentence duration in Childes in French and Japanese.

pling acoustic cues distributed over adjacent phonemes. Therefore, if taken literally, one should map phonetic events on the left hemisphere and phonological processing on the right hemisphere, which is obviously not the case.

Despite this diversity in instantiations, this set of hypotheses is prevalent in the field, and it is empirically interesting to investigate whether linguistic structures are lateralized along this temporal continuum. Rapidly changing speech components (including consonant–vowel (CV) stimuli) activated predominantly the left auditory area in many studies (e.g., Jancke et al., 2002; Dehaene-Lambertz and Gliga, 2004; Zatorre et al., 1992, 1996), whereas stimuli with richer pitch (tone, intonational prosody) modulates a right-dominant activation (Furuya and Mori, 2003; Meyer et al., 2002; Zatorre, 1988; Zatorre et al., 1992). However, not every published paper has found such hemispheric specialization in accordance with the temporal continuum. For example, CV stimuli activated the brain symmetrically in Binder et al. (2000), Joanisse and Gati (2003) and Benson et al. (2006). Nonetheless, research involving both speech and non-speech lends support to signal-driven explanations. Jancke et al. (2002), for instance, found greater involvement of the left planum temporale in processing CV rather than a tone or a vowel in isolation. That this was due to the enhanced temporal nature of the voiceless consonant in CV was confirmed in a second study, where a non-speech stimulus with similar rapid temporal changes (such as a "gap") tended to activate the left auditory region (Zaehle et al., 2004). In summary, these studies as well as other imaging literature (Zaehle et al., 2008) suggest that, at an early stage of auditory perception, speech and non-speech processing share a similar neuronal pathway that is driven by signal properties, and that, at this stage, lateralization responds to the differential hemispheric receptivity to rapid vs. slow variation in the acoustic signal. A complete picture would necessarily involve additional processing stages in order to account for left lateralization in response to 'slow' phonological features, such as lexical tones (Gandour et al., 2002, 2004; Xu et al., 2006).

2.2. The Domain-driven hypothesis

2.2.1. Identifying features

The basic feature of this set of hypotheses is that there is a single (left-lateralized) brain network which responds to the linguistic characteristics of the input. Fodor (1985) proposed that human language is a 'module' that implicates a set of innately specified, automatic, dedicated processes. Chomsky and Lasnik (1993) claimed that the species-specificity of language resides in a set of abstract properties that takes the form of a Universal Grammar (UG), i.e. a set of abstract parameters and principles. Translated into brain networks, this yields the idea that there is a human-specific, domain-specific, left-lateralized processing architecture that is initially present, independent from experience. The domain of this processing architecture is not defined by low-level stimulus characteristics, but rather, the abstract principles of (human) UG. This includes spoken and sign language, but excludes music or computer languages. However, the basic intuition that left lateralization arises from the *linguistic* characteristics of the stimuli is not *always* associated to networks that are human-specific, domain-specific, and learning-independent, as shown on Table 2.

Although the hypothesis of a dedicated brain network for language has been formulated for humans, similar hypotheses have been proposed for other species suggesting some phylogenetic continuity. For instance, a right ear (left hemisphere) advantage has been observed in response to conspecific calls in rhesus monkeys (Hauser and Andersson, 1994), sea-lions (Boye et al., 2005) and rats (Ehret, 1987). Furthermore, recent imaging data in rhesus monkeys shows that, in contrast to temporal lobe activities that were basically right-lateralized for various types of stimuli, only conspecific calls significantly activated the left temporal pole (Poremba et al., 2004). These results, however, should be interpreted cautiously, as much counterevidence has been reported (e.g., Gil-da-Costa et al., 2004). Regarding specificity to language, sign language involves very similar networks to spoken language, despite the fact that it rests on a different modality

Table 2

Selected quotes to represent the variety of theoretical stances within the domain-driven set of hypotheses, depending on whether the neural bases are specific to humans, whether they are used only for language, and whether learning is unnecessary.

Reference	Quote	Human-specific	Domain-specific	Present from birth
Dehaene-Lambertz and Gliga (2004)	Therefore, we hypothesize that in the case of phoneme processing, there is continuity between neonates and adults, and that from birth on infants are able to spon- taneously compute phonemic representations [.]This phonemic network, effective from the first days of life, is adequately configured to process the relevant properties of the speech environment and to detect any inherent regularities present in input.[.] It is not exposure to speech that creates the capabilities described in infants.	Yes	Yes	Yes
Peña et al. (2003)	[These results imply] that humans are born with a brain organization geared to detect speech signals and pay attention to utterances produced in their surroundings.		Yes	Yes
Dehaene-Lambertz et al. (2006)	We do not know yet whether another structured stimulus, such as music, would activate the same network. [. T]he similarity between functionally immature infants and competent mature adults implies a strong genetic bias for speech processing in those areas. This 'bias' might partially result from recycling of auditory processes observed in other mammals (e.g. rhythmic sensitivity or perceptive discontinuities along some acoustic dimension) but is not limited to them.	Partially	Partially	Yes
Dehaene-Lambertz et al. (2010)	Acknowledging the existence of strong genetic constraints on the organization of the perisylvian regions [for speech perception] does not preclude environmental influences.			No

(Poizner et al., 1987; MacSweeney et al., 2002; Campbell et al., 2008; but see Neville et al., 1998; Newman et al., 2002; where it is found that RH activation in native ASL signers seeing ASL is greater than that found in English speakers hearing English). A priori, this fits well with the idea that it is abstract properties, not low level signal properties, which are responsible for the pattern of specialization for language. Similarly, left-dominant activations have been recorded in response to whistled Spanish in a group of people who frequently used it, even though whistled speech has signal properties similar to music (Carreiras et al., 2005). Of course, this previous research with signed and whistled languages typically used words or sentences, stimuli that had morphology, syntax, and semantics. Hence, it is possible that the left dominance documented there did not reflect phonological processing. Clearer evidence to this effect would come from studies using meaningless, but phonologically and phonetically well-formed signs and whistles, which would be comparable to the spoken non-words/pseudowords typically used when neuroimaging spoken language phonological processing. Pseudosigns have been used in behavioral research in order to better isolate phonetic/phonological processing from lexical treatment, and they can reveal differences in perception between native signers, late learners, and non-signers, suggesting they tap a linguistic level of representation (e.g., Baker et al., 2005; Best et al., 2010).

2.2.2. Adult data

In its strongest form, the domain-driven hypothesis predicts left-dominant responses to any linguistic stimulation, regardless of input modality and previous experience. Contrary to this view, Mazoyer et al. (1993) and Perani et al. (1996) reported symmetrical activation of superior temporal regions for the presentation of a completely unknown language. This has also been found by MacSweeney et al. (2004) for sign language. Note, however, that such conclu-

sions may depend on the control conditions used, because when compared to backward speech, an unknown spoken language elicited a significantly larger leftward activation in the inferior frontal gyrus, inferior parietal lobule, and mid-temporal gyrus (Perani et al., 1996). Nonetheless, present evidence in favor of a lateralized network for an unknown language is not very strong.

Another line of evidence relevant to the domain-driven hypothesis comes from studies where the same stimuli elicit differential brain activations depending on whether they are perceived as speech or not, or whether the participant is focusing on the linguistic aspect of the signal (Dehaene-Lambertz et al., 2005; Mottonen et al., 2006; Vouloumanos et al., 2001). Mottonen et al. (2006), for instance, demonstrated an enhanced left-lateralized STS activation only for participants who were able to perceive sine-wave stimuli as speech. In addition, the same acoustic stimulus can yield a different pattern of lateralization, depending on whether the task is to differentiate the acoustic/voice or linguistic characteristics (Bristow et al., 2009; Meyer et al., 2002). This shows that hemispheric lateralization is not only determined by the acoustic characteristics of the stimuli; instead, the brain can be set into a language or a non-language processing mode, and that the former specifically involves left lateralized structures (Dehaene-Lambertz et al., 2005; Meyer et al., 2002; Mottonen et al., 2006). Naturally, such a processing mode could itself result from learning. This is what we explore next.

2.3. The Learning biases hypothesis

2.3.1. Identifying features

Contemporary studies of cognitive development favor the view that biological systems rely neither on a single, general-purpose learning mechanism, nor on domain-specific hard-wired solutions, but rather on a series of specific learning mechanisms that are "distinguished by

their properties – for example, whether or not they depend on temporal pairing – [and] not by the particular kind of problem their special structure enables them to solve" (Gallistel, 2000, pp. 1179). If different learning mechanisms require the computational resources of distinct brain areas and networks (Ashby et al., 1998; Davis et al., 2009), functional specialization for speech perception could be a side effect of learning. In other words, lateralization patterns could be the result of having recruited lateralized networks during the learning process.

Within the general framework of learning-based accounts, we propose a specific instantiation of a learning biases hypothesis whereby the units and relationships learned during phonological acquisition require a set of highly specialized learning mechanisms, some of which are more efficient on the left. Such mechanisms are not necessarily specific to language, and can also be recruited in other domains, but language is probably the only domain which recruits each and every one of them. According to linguistic theory and behavioral research, (a) spoken phonetic units are *abstract categories* composed of *features* (Chomsky and Halle, 1968; Hall, 2001; Holt and Lotto, 2010; Kenstowicz and Kisseberth, 1979; Maye et al., 2008; White and Morgan, 2008; Cristia et al., 2011), and (b) acceptable wordforms are made up of legal *sequences* (Kenstowicz, 1994; Mattys et al., 1999; Graf Estes et al., 2011) of sounds determined within *prosodic (hierarchical) structures* (Coleman and Pierrehumbert, 1997; Nespor and Vogel, 1986). Abstract categories composed of features, and sequencing and hierarchical structures, are found in domains other than language, and can be examined with non-linguistic material and non-human animals. Thus, this hypothesis is not strictly speaking domain-driven. Similarily, given that these mechanisms are involved in learning with non-auditory input they are not strictly speaking signal driven either. What defines these mechanisms, rather than their function or input, is the internal *representations* and *computations* they use in order to extract regularities.

In the next subsection, we summarize studies that explore some of the learning mechanisms that could sustain the emergence of such units. Furthermore, we also review further evidence that left-lateralization for speech is the result of learning, since it is stronger for (better) known languages.

2.3.2. Adult data

There is much evidence that left-dominance is associated with abstract, categorical processing, even when the categories are nonlinguistic, as illustrated in Table 3. For example, results using both visual and auditory stimuli document a right eye/right ear/left hemisphere advantage for categorical, abstract processing and a left eye/left ear/right hemisphere for exemplar-based processing both in adult humans (Curby et al., 2004; Marsolek and Burgund, 2008) and non-human animals (Yamazaki et al., 2007). For example, in Marsolek and Burgund (2008) human adults were presented with 2 novel 3-D shapes sequentially, and had to perform one of two tasks: In the same-category task, they should decide whether the 2 shapes shared enough features or parts to belong to the same category; in the same-exemplar task, whether they were the

exact same shape. When the sequences were presented to the left eye/RH, responses were faster for the same-exemplar task than the same-category task, whereas the reverse was true for right-eye/LH presentations. Since this RH-exemplar advantage is even evidenced by long-term repetition priming of environmental sounds (Gonzalez and McLennan, 2009), it is apparent that the RH is at a disadvantage for abstract category processing. In addition, individual variation in proficiency in category learning predicted the degree of left-hemisphere involvement in recent training studies with non-speech (Leech et al., 2009) and visual categories (Filoteo et al., 2005), furnishing some evidence that further left-hemisphere involvement results in more efficient learning. On the other hand, the precise role of *features* in such a pattern of LH dominance is still not well understood.

As for learning of sequencing and hierarchical regularities, the LH appears to be more efficient than the RH in learning both types. Notice that some of the evidence comes from artificial grammar learning studies that were originally geared towards syntax. In that work, it is often said that adjacent regularities of the type captured by finite state grammars are not properly linguistic, whereas more interesting aspects of language structure can only be represented through the more complex phrase structure grammars. This description may be more appropriate to syntax, whereas much of phonetics and phonology could be described through regular grammars (or perhaps even subregular ones; Heinz, in press-a, in press-b; Rogers and Hauser, 2010). Regardless of the computational algorithm would best capture phonology, current descriptions state that phonological regularities respond to both adjacent constraints and hierarchical properties.

We now turn to the predictive power of the learning bias hypothesis for adult speech processing. A major prediction of the learning bias hypothesis is that left lateralization should only be found with stimuli that can be parsed using the categories and regularities of a known language. As mentioned above, the presentation of sentences in a completely unknown language activates a restricted region, close to the auditory areas in a largely symmetrical fashion in adults (Mazoyer et al., 1993; but see Perani et al., 1996). In contrast, a second language mastered late but with high proficiency activates a left lateralized network that is almost superimposed to that of the native language (Perani et al., 1996, 1998), whereas a language acquired late with low to medium proficiency activates a network of extension similar to that of the native language, but less lateralized and presenting greater individual variability (Dehaene et al., 1997). In other words, as a language is learned in adulthood, the brain recruitment varies with proficiency from an (almost) symmetrical representation to that of the full left lateralized network typical of first language.

The same results are found with the discrimination of isolated sounds: a pair of sounds elicits asymmetrical activation in the temporal area only when the sounds form a contrast in listeners' native language. This has been documented for consonants (Rivera-Gaxiola et al., 2000), vowels (Dehaene-Lambertz, 1997; Naatanen et al., 1997; Minagawa-Kawai et al., 2005), tones (Gandour et al., 2002,

Table 3

Selection of training and perceptual studies association left-dominance with some of the characteristics attributed to phonological knowledge.

Level	Characteristics	Evidence				
		Stimuli type	Task/Stimuli	Areas involved	Population	Reference
Sound units	Feature-based	non-speech	Categorization	Individual variation correlated with L pSTS activation	Adults	Leech et al. (2009)
		Visual	Categorization	Individual variation correlated with L frontal and parietal	Adults	Filoteo et al. (2005)
	abstract (resilient to physical changes; excludes exemplar information)	Visual (feature-based, 2-D)	Categorization of trained vs. novel exemplars	R eye: feature-based; L eye: exemplar-based, configural processing	Pigeons	Yamazaki et al. (2007)
		Visual (not feature based; novel objects)	Viewpoint processing	Reduced viewpoint-specific effects when presented to the R eye (but only when objects associated to labels)	Adults	Curby et al. (2004)
		Visual (feature-based, 3-D)	Category identification vs. exemplar identification	R eye advantage for category; L eye advantage for exemplar	Adults	Marsolek and Burgund (2008)
		Environmental sounds (not feature based)	Long-term repetition priming	Exemplar priming only when presented to the L ear	Adults	Gonzalez and McLennan (2009)
Sound patterns, wordforms	Regularities describable in terms of adjacency	Written letters or syllables	Rule-based (versus item-based) trials over the course of learning	L prefrontal cortex	Adults	Fletcher et al. (1999)
			Illegal > legal strings	L operculum, R STS	Adults	Friederici et al. (2006)
			illegal > legal strings	L IFG	Adults	Forkstam et al. (2006)
		Tone sequences	Tones that had co-occurred vs. random tones	L IFG	Adults	Abla and Okanoya (2008)
		Spoken syllables	Variation frequency of co-occurrence	L STG, IFG	Adults	McNealy et al. (2006)
		Synthetic syllables	Immediate repetition within trisyllables > no repetition	L parieto frontal	Newborns	Gervain et al. (2008)
	Regularities describable in terms of hierarchical structure	Written letters, syllables, or words	Illegal > legal strings	L operculum, L IFG, L MTG, R STS	Adults	Friederici et al. (2006)
			Illegal > legal strings	L IFG	Adults	Opitz and Friederici (2003)
			Rule change > word change	L ventral premotor	Adults	Opitz and Friederici (2004)

2004; Xu et al., 2006), and syllable structure (Jacquemot et al., 2003). Recent evidence further shows that initially symmetrical electrophysiological responses to non-native contrasts shifted to left-dominant after intensive training (Zhang et al., 2009). Training to categorize non-phonemic auditory signals also enhanced fMRI activation in left posterior superior temporal sulcus (Liebenthal et al., 2010). Inversely, training to associate a brief temporal distinction (along the same dimension that distinguishes/d/from/t/) to talker identity, rather than speech categories, can cause the opposite shift from left- to right-lateralization (Francis and Driscoll, 2006). Brain morphometric studies also support a critical role of the left temporal area for efficient language learning. By studying individual differences when learning a new phonetic contrast, Golestani et al. (2002, 2007) showed that faster learners have more white matter volume in the left Heschl's gyrus and parietal lobe than slow learners.

In summary, there is some evidence for the claim that lateralization increases with familiarity to the language or contrast being processed. The learning mechanisms responsible for this lateralization may be related to the particular properties of phonological categories in speech (compositionality, abstractness, sequential and hierarchical structure), but more research is needed to pinpoint the brain circuits sensitive to these separate properties.

3. Fit between the hypotheses and developmental data

The three hypotheses reviewed so far are difficult to distinguish based on adult data only, because part of the lateralization observed in adults could be the consequence of developmental processes rather than due to an intrinsic difference in the processing function of the 2 hemispheres. This is why we now turn to developmental data, the central topic of the present paper. As mentioned in the introduction, cognitive neuroscience in infancy has greatly benefited from technical advances in neuroimaging methods, including NIRS (see Minagawa-Kawai et al., 2008, for a review). In the following sections, extant developmental neuroimaging data including both NIRS and fMRI are evaluated in terms of the signal-driven, the domain-driven, and the learning-biases hypotheses. This examination reveals how these hypotheses account for the neural substrates involved in speech processing in infancy, both when exposed to running speech (Section 3.1) and when tested with specific sound contrasts (Section 3.2).

3.1. Processing running speech

The studies reviewed below measure the brain activation in infants between 0 and 10 months of age using NIRS or fMRI in response to the presentation of sentences, either natural or modified, or artificial sounds. Table 4 shows a classification of 11 studies on the basis of the predictions drawn from the three sets of hypotheses. The fit of the hypotheses is represented by the match between the colors in the bottom of each column and the colors found within the cells.

According to the signal-driven hypothesis, stimuli consisting of rapid temporal changes (i.e., pure segmental information, coded in blue) should elicit left-dominant auditory activations; slow spectral changes associated with pitch in intonation (in red) should activate the right temporal cortex and normal speech containing both fast and slow signals (in lylac) should activate both hemispheres to the same extent. The predictions are globally sustained for slow signals: With only three exceptions, slow, spectrally rich signals activate more the right hemisphere (slowly changing tones, Telkemeyer et al., 2009; emotional vocalizations, Minagawa-Kawai et al., 2011; Grossmann et al., 2010; normal versus flattened prosody, Homae et al., 2006). Two of the exceptions involve music; the remainder concerns flattened prosody in 10-month-olds (Homae et al., 2007), which activates the RH to a larger extent than normal prosody, contrary to what happens in 3-month-olds (Homae et al., 2006). The latter exception could be captured by proposing that 10-month-olds have learned that flat prosody is abnormal, and thus requires extra processing. The prediction of a greater involvement for signals involving fast changes is less clearly sustained by the data. Speech seems to be more left-lateralized than expected based on the fact that it contains a mix of fast and slow signals. In addition, the only experiment using pure fast nonspeech signals (Telkemeyer et al., 2009) reports a response that is symmetrical. In short, the signal driven hypothesis correctly predicts RH dominance for slow signals, prosody and emotion, but the dominance of LH for fast signals is less well established empirically. If, as claimed by Boemio et al. (2005), fast signals turn out to elicit mostly *symmetrical* activation, LH dominance of language can no longer be accounted for by a signal-driven hypothesis.

As for the domain-driven classification, speech stimuli (blue) should involve the left hemisphere to a larger extent than comparable non-speech stimuli (red), and non-native speech stimuli may be more symmetrical. The results do not fall neatly within these predictions. Although it is true that most of the studies report left or symmetrical results for normal speech, it appears to be the case that non-speech analogues are also processed in a left-lateralized manner.

Knowledge of language-specific prosody, phonetic units, sound patterns, and wordforms is not evident in behavior before 5 months, and becomes increasingly language-specific over the first year (e.g., prosody: Nazzi et al., 2000 for language discrimination at 5 months; phonetic units: Kuhl et al., 1992 for vowel knowledge at 6 months; word-level stress and phonotactics by 9 months, Jusczyk et al., 1993a, 1993b; consonants by 10–12 months, Werker and Tees, 1984). In view of this behavioral evidence, the learning bias hypothesis predicts symmetrical processing and little difference between native and foreign languages before 4 months, and increasing left lateralization only for the native language after. Since no research has compared L1 and FL in the second half of the first year, this prediction cannot yet be falsified. However, extant data suggest that some neural tuning to the native language commences before 6 months, although this may not translate into differences in lateralization at this stage. Specifically, while there is no difference in activation for L1 as compared to FL at birth (Sato et al., 2006), there is greater

Table 4

Neuroimaging data on infants exposed to blocks of running speech or speech analogues, classified on the basis of the signal driven hypotheses in the first set of columns, and on the basis of the domain driven hypothesis in the second set.

Age (months)	Signal-Driven			Domain-Driven			Learning Driven			References
	Left	Bilateral	Right	Left	Bilateral	Right	Left	Bilateral	Right	
types of stimuli	segment rapid	segment + prosody	prosody slow	speech	non-speech		native speech (L1)	non-native (FL)		
0	L1			L1, L1 vs BW			L1			Peña et al. (2003)
		L1, FL, BW		L1 vs BW	L1, FL, BW, FL vs BW			L1, FL, L1 vs FL		Sato et al. (2006)
		25 ms tones	300 ms tones							Telkemeyer et al. (2009)
		Music, L1			L1			L1		Kotihlati et al. (2010)
2	L1, L1 vs Music	Music	Music vs. L1	L1, L1 vs Music			L1			Dehaene-Lambertz et al. (2010)
3	L1, BW			L1, BW			L1			Dehaene-Lambertz et al. (2002)
		Flattened, L1	L1 vs Flattened		L1			L1		Homae et al. (2006)
4	L1, FL, Scrambled		Emotional voc.	L1, FL, Scrambled			L1, FL	L1 vs FL		Minagawa-Kawai et al. (2011)
6 ~ 9	L1			L1			L1			Bortfeld et al. (2009)
7			Emotional voc. vs. other sound							Grossmann et al. (2010)
10	L1		Flattened		L1			L1		Homae et al. (2007)

In the signal-driven classification, blue codes for stimuli containing rapid/segmental content only; red, stimuli containing prosodic oppositions; lylac, stimuli containing both prosody and segments. In the domain-driven classification, blue codes for (native) speech stimuli; red for non-speech; lylac for non-native speech. In both sets of columns, blue indicates left bias predicted, red right bias predicted, lylac bilateral predicted. *Abbreviations*: L1: first language, FL: foreign language, BW: backward speech, Flattened: flattened speech, Emotional voc.: emotional vocalization, Scramble: scrambled sound.

activation to L1 than FL by 4 months of age (Minagawa-Kawai et al., 2011). Further data from our lab suggests that dialect discrimination elicits greater activation in the left hemisphere at 5 months (Cristia et al., submitted for publication), by which age this task recruits language-specific knowledge according to behavioral research (Nazzi et al., 2000). These data underline the importance of studying language processing throughout the first year.

Due to the large variability in extant results, none of the 3 hypotheses received overwhelming support. Nonetheless, this is not at all unexpected, since the data is very sparse: while we review the mere 11 studies reported to date, a recent review paper on left lateralization in adult language processing had the advantage of looking at 100 data points published within a single year (Price, 2010).

Additionally, the stimuli used in these 11 studies typically combined a number of features, and we were thus not ideally positioned to adjudicate between the three competing hypotheses. To this end, we now focus on studies using more controlled stimuli in the next section.

3.2. The development of phonological contrast

In this section, we examine in detail the neurodevelopment of the processing of individual contrasts. Unlike papers discussed in the previous section, speech processing here is gauged through the comparison of two types of blocks, one where two stimuli alternate, versus one where a single stimulus is repeated. This enables the study of the brain networks involved in speech sound discrimina-
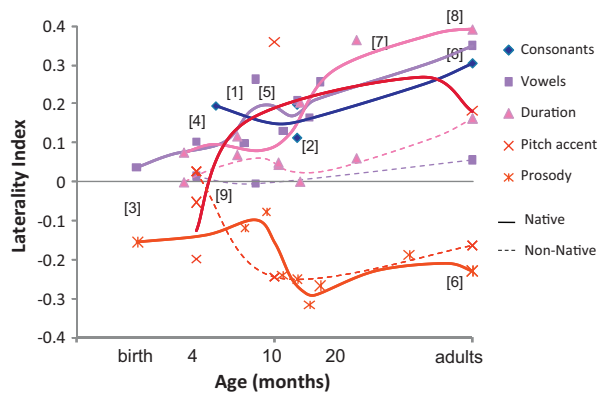
Fig. 1. Developmental changes of laterality index in various phonological contrasts. Original data are from [1] Furuya et al. (2001) and [2] Sato et al. (2007) for consonants, [3] Arimitsu et al. (in preparation), [4] Minagawa-Kawai et al. (2009a), [5] Sato et al. (2003) and [6] Furuya and Mori (2003) for vowels, [7] Minagawa-Kawai et al. (2007) and [8] Minagawa-Kawai et al. (2005) for durational contrast, [1,2] and [9] Sato et al. (2010) for pitch accent and [3,5,6] for prosody. All studies use the same change detection paradigm to examine the cerebral responses around auditory area. A laterality index was calculated using the formula $(L(R))/(L+R)$, where L and R are the maximal total Hb changes in the left and right auditory channels, respectively. Laterality index is above zero for left dominance and below zero for right dominance.

tion. The 9 studies are presented in Fig. 1 and listed in the supplementary materials/appendix. Before reviewing that data, let us draw out the predictions for each of the three hypotheses, as follows. If infants' neural responses to sound contrasts depended only on the contrasts' physical properties, we would expect a right–left lateralization gradient, with lexical pitch involving primarily the right hemisphere, vowel quality involving symmetrical processing, and consonants involving more leftward networks. In contrast, according to the domain-driven hypothesis, it is to be expected that all linguistic contrasts would elicit larger left-hemisphere activations from birth (with, perhaps, left-dominance decreasing for non-native contrasts with additional experience). Finally, the learning-biases hypothesis predicts that left lateralization should emerge as a consequence of acquisition, and therefore would only concern contrasts that can be captured using categories and rules developed from exposure to the ambient language(s).

As shown in Fig. 1, before 6 months, infants exhibit significantly rightward activations for prosodic and pitch accent in contrast to the leftward activation for consonants and consistently symmetrical activation for vowels. These results generally fit the predictions from the signal-driven hypothesis, as slow, spectrally rich signals (prosodic, pitch accent) elicit right-dominant, consonants left-dominant, and vowels symmetrical activations. However it should be noted that not all the data before 6 months is in accordance with the signal-driven hypothesis, and that there are very few data points for consonants.

Developmental results provide support to the learning bias hypothesis, as contrasts become increasingly left-lateralized only if they are part of the native phonology, while non-native contrasts and non-speech analogues continue to be represented symmetrically/right-dominant. In consonance with previous behavioral research, the timing

of acquisition appears to vary according to the contrast type, such that vowel quality (for monosyllabic stimuli, behavioral: 6 months, Kuhl et al., 1997; NIRS: 7 months, Minagawa-Kawai et al., 2009a; MEG: 6 months, Imada et al., 2006) may be acquired earlier than lexical prosody (behavioral: 9 months, Mattock et al., 2008; NIRS: 11–12 months, Sato et al., 2003; although notice that the stimuli used by Sato et al. were bisyllabic, whereas Mattock and Burnham used monosyllables) and vowel duration (behavioral: 18 months, Mugitani et al., 2009; NIRS: 14 months, Minagawa-Kawai et al., 2007). It is uncertain why some contrasts are learned earlier than others, but it may be the case that acoustically salient ones require less exposure (Cristia et al., in press). Although there is little work on consonants,[1] these showed activations that were somewhat left-dominant at the early age of 5 months.

In all, it appears that a combination of the signal-driven hypothesis and the learning bias hypothesis, with their relative contributions varying with the infant age/experience, may provide a good fit of the data on sound contrasts, as these results document an increase in lateralization as a function of development and experience from an initial state where lateralization responds to signal factors. To take a specific example, let us focus on the case of pitch accent. Initially rightward/symmetrical activations gradually change to left-lateralized only if the contrast is phonological in the infants' ambient language, whereas non-native/non-linguistic analogues continue to elicit right lateralized responses (with one exception: a symmetrical response has been reported for pitch contrasts in Japanese 4 month-olds; Sato et al., 2010). Furthermore, these results complement those in the previous section, as they underline the importance of learning for lateralization in response to isolated words, stimuli that allow a much greater control over the factors influencing lateralization.

4. A developmental scenario

As summarized in Table 5, the three hypotheses (signal-driven, domain-driven and learning biases) capture some of the infant and adult lateralization results we reviewed, but none fully account for all of them. Thus the signal-driven hypothesis provides a principled account for right-dominant activations in response to prosodic content, early lateralization patterns for some sound contrasts in early infancy, and the to-be-confirmed left bias for language in newborn. However, it cannot account for the following 4 sets of results. First, although lexical prosody (tone, stress, pitch accent) relies on the same 'slow' acoustic dimensions involved in sentential prosody, single words differing along those dimensions elicit left-lateralized activations in adults who speak a language where tone, stress, and pitch accent are contrastive. Second, the signal-driven hypothesis by itself does not predict the existence of task effects; for example, when the same

---

[1] Dehaene-Lambertz and Gliga (2004) reported in an ERP study that left-lateralized responses to consonantal contrasts were evident in newborns and 3-month-olds, but similar left-dominant activations were elicited by non-speech analogues.

Table 5
Main findings of the adult and infant literature review carried out in previous sections. As evident, no single hypothesis covers all of the evidence.

| Finding		Signal-driven	Domain-driven	Learning bias
1	Adults: Slow signals activate more LH if linguistically contrastive	(	+	+
2	Adults: Language mode activates more LH (task effects)		+	+
3	Adults: Sign language activates more LH		+	+
4	Adults: LH involvement proportional to proficiency	(	(	+
5	Adults: FL contrast elicits RH if slow, LH if fast	+	(	
6	Newborns: L1 vs non-speech only in LH in the absence of extensive experience		+	(
7	Infants: Slow signals activate more RH	+		
8	Infants: L-dominance increases with development and experience	(	(	+
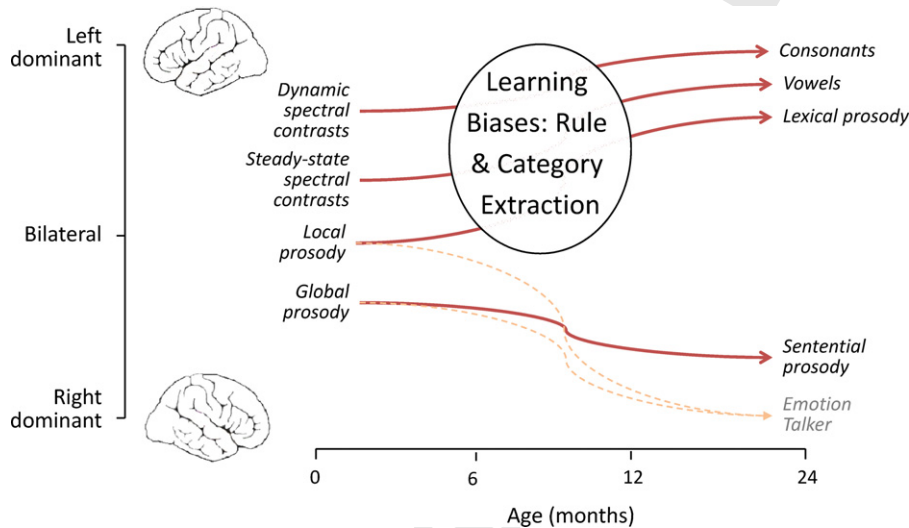
Fig. 2. A schematic model of developmental hemispheric lateralization.

physical stimulus gives rise to left- or right-dominant activation depending on whether the task relates to language comprehension or talker identification.[2] Third, lateralization of sign language processing remains unaccountable within the signal-driven hypothesis. Finally, since learning cannot affect the physical characteristics of the signal, this hypothesis have little to say about developmental *changes*, including the fact that speech involves more left-dominant responses with increased language exposure.

In contrast, both the domain-driven and the learning bias hypotheses provide a parsimonious account for the first three sets of findings listed on Table 5. The last one specifically supports the learning bias hypothesis, together with the differences in brain representation for L1 versus L2 or FL in adults. Finally, if an initial asymmetry for language in newborn were confirmed, this would not be incompatible with a learning bias provided that the effect could be traced back to *in utero* experience.

Even though we presented the three hypotheses as exclusive alternatives, they are not incompatible with one another. As has been proposed in the case of face perception (Morton and Johnson, 1991), signal-based orienting mechanisms can channel particular stimuli to a domain-general learning mechanism, which eventually results in a mature domain-specific system. Additionally, a signal-based approach can be reconciled with a domain-driven hypothesis if low-level biases are supplemented with higher level perceptual biases (Endress et al., 2009; Mehler et al., 2008). Therefore, we propose a developmental scenario in three steps for the unfolding of lateralization which combines the signal-driven and learning bias hypotheses, to result in processing that appears to be domain-driven in the adult state (Fig. 2). First, at the initial stage, neural recruitment for speech processing is chiefly influenced by temporal and spectral properties of speech; thus, rapidly changing sounds would yield left-dominant or bilateral activations and slowly changing, spectrally rich sounds right-dominant ones. This correctly predicts right-dominant activations for sentential prosody in 3-month-old infants (Homae et al., 2006), and a possible left-right gradient for segments and suprasegments. Second, as infants are exposed to language, the left hemisphere learning systems of phonological category capture the newly learned sounds into the lexical circuits around the left temporal areas. Finally, in the stable state, L1 speech processing has become basically left dominant, giving rise to domain-specific language networks (although ones that only apply to known languages, or to novel streams that can be captured with structures from the known language). A key prediction of this scenario arises in the example

---

[2] One could explain some of the task effects through attentional amplification of particular signal characteristics: for instance, attending to a phoneme change versus talker change. The representation of fast versus slow would enhance aspect of speech.

of tone systems (e.g., in Thai and Yoruba) or pitch accent (used, for instance, in Japanese and Swedish). Our scenario predicts that such lexical prosody contrasts should initially elicit right-dominant activations, in consonance with the signal-driven hypothesis. However, in languages where such contrasts are phonological, the infant must come to learn that pitch patterns function like a phonological feature in the composition of wordforms, and that their distribution can be hierarchically and sequentially bound (e.g., in tone sandhi). As a result, the involvement of left-dominant mechanisms recruited for this learning will eventually result in left dominant activations. The end product is that in adults, non-native listeners process these contrasts with right-dominant or symmetrical activations, while native listeners evidence left-dominant ones (Gandour et al., 2002; Xu et al., 2006). At present, this prediction for the leftward shifts through development has been confirmed in a longitudinal study on the processing of Japanese pitch accent. However, as mentioned above, evidence is still sparse, particularly at early ages, and there are only two studies showing asymmetrical or right-dominant activations in response to pitch contrasts at an early age (Fig. 1). This calls for further research, particularly with neonates.

In a nutshell, according to our proposed scenario young infants' laterality can be accurately described taking only the acoustic properties of the stimuli into account, but learning-based changes are necessary to account for the eventual domain-driven organization for the first language in the left hemisphere. This does not implicate that signal-driven processing ceases to function for L1 in adults, as it may still be at work at lower-level auditory processing (as assumed in the dual pathway model, Friederici and Alter, 2004). For instance, patients with lesions in the corpus callosum are able to correctly process acoustic cues of grammatical prosody on the RH as a lower-level of auditory processing, but such cues are simply not available for linguistic interpretation due to a failure of transfer to the LH (Friederici et al., 2007). It would be interesting to study a similar effect for lexical prosody (tones, pitch accent, stress).

5. Open questions and conclusion

As noted in the introduction, the discussion has been restricted to the hemispheric specialization of speech perception; however, perception and production are closely linked each other, according to adult neuroimaging studies (Iacoboni, 2009; Morillon et al., 2010; Kell et al., 2011). For instance, when a phoneme is perceived in adults, activations of inferior frontal gyrus (IFG) are frequently reported in addition to those in the auditory area and posterior STG, suggesting that a dorsal network associated with a sensory-motor loop of phoneme processing (Buccino et al., 2001; Dehaene-Lambertz et al., 2005). Although the precise contribution of motor representations in speech perception (and vice versa) is a matter of debate (see e.g., Alexander Bell et al., 2011; Hickok et al., 2009; Yuen et al., 2009, and references therein, for diverse perspectives on the matter), it is incontestable that infants' language experience is multimodal: Infants will often see the movement of at least

some articulators in the talking caregiver; and their experience of speech will necessarily involve the auditory and somatosensory channels as soon as they babble, which can be as early as 4 months (Vihman, 1996). Indeed, a recent connectivity study with NIRS on 3 month-old infants documented that activation measured in channels over frontal regions correlated significantly with that registered in temporal regions during and after exposure to speech stimuli (Homae et al., in press). Nonetheless, a MEG study focusing specifically on Broca's area failed to find consistent evidence for speech-specific activations before 12 months (Imada et al., 2006). Here again, the description of language networks would greatly benefit from more work being carried out over the first year of life, as this sparse evidence leaves important questions unanswered, such as the type of experience necessary to establish action–perception loops.

Similarly, the present review mostly dealt with left/right asymmetries within the auditory areas (including the planum temporale and STG). It is likely that such areas for initial auditory processing are connected to lexical networks and their connectivity is strengthened by phonological acquisition. We speculate that such network involves the ventral route from STG, middle temporal gyrus to IFG (Hickok and Poeppel, 2004, 2007). At the same time, phonological/phonetic representations encoded around the auditory area will be further connected to the dorsal pathway which may involve phonological short-term memory and sensory or articulatory processing of speech (Hickok and Poeppel, 2007). We hope that future empirical and theoretical research is able to enrich our developmental scenario of hemispheric specialization with considerations of infants' language acquisition beyond the auditory areas.

To this point, we have largely overlooked learning *in utero*. However, it may be the case that some learning occurs before birth, where the infant has some access to phonetic information (DeCasper et al., 1994; Querleu et al., 1988; see also Granier-Deferre et al., 2011, for a recent summary; and for a report that forward and backwards L2 evoked similar patterns of heart-rate decelerations in 38-week-old fetuses). As mentioned above, there is one fact that is only accountable through the domain-driven hypothesis, namely that a comparison of L1 and non-speech is only significant in the LH in newborns (Peña et al., 2003; replicated in Sato et al., 2006). However, such asymmetry is not evident in the L2 versus non-speech comparison reported in Sato et al. (2006), which would fit with a learning-biases account. Moreover, a recent fMRI study shows greater leftward posterior STG activation to the mother's speech than to a stranger's speech in 2-month-olds (Dehaene-Lambertz et al., 2010), lending further support to experience-based asymmetries. Therefore, future theoretical investigations should incorporate a consideration of the effects of *in utero* experience.

One final consideration is in order: Speech perception is left-lateralized in most individual cases, but not universally. An oft-cited case involves plasticity, whereby young children who have lost the left hemisphere come to develop language quite typically (Liegeois et al., 2004). In contrast, atypical lateralization has been observed in disordered development, for example in children with

autism spectrum disorder (e.g., see Minagawa-Kawai et al., 2009a,b). But even within the normal population, there is a statistically significant variation in the degree of lateralization (e.g., Szaflarski et al., 2002; Whitehouse and Bishop, 2009). Future work should also consider how genetic factors may shape signal-driven biases present by birth, and how genetic factors, experience, and their interaction may shape the learning-driven bases that impact lateralization over the course of development, in order to better understand variation in left-dominance during speech processing.

In conclusion, we have reviewed adult and infant neuroimaging data on asymmetrical activation in response to the processing of speech characteristics in the absence of lexical, semantic, and syntactic characteristics. Three hypotheses were found insufficient to capture these data: a signal-driven explanation and a domain-driven hypothesis, explored to some extent in previous work, and a novel proposal based on learning biases. Therefore, we put forward a developmental model, that combines the signal-driven and learning-biases explanations to account for most of the extant results, and which further allows to make some important predictions for future work.

Q4 Uncited references

Bortfeld et al. (2009), Dehaene-Lambertz et al. (2002), Goldstein and Schwade (2008), Grossmann et al. (in press), Kotilahti et al. (2010), Mehler et al. (1978), Rogers and Hauser (in press) and Scott et al. (2000).

Acknowledgements

References

Abla, D., Okanoya, K., 2008. Statistical segmentation of tone sequences activates the left inferior frontal cortex: a near-infrared spectroscopy study. Neuropsychologia 46, 2787–2795.

Alexander Bell, C., Morillon, B., Kouneiher, F., Giraud, A.-L., 2011. Lateralization of speech production starts in sensory cortices—A possible sensory origin of cerebral left-dominance for speech. Cereb. Cortex 21, 932–937.

Arimitsu, T., Uchida-Ota, M., Yagihashi, T., Kojima, S., Watanabe, S., Hokuto, I., Ikeda, K., Takahashi, T., Minagawa-Kawai, Y. Functional hemispheric specialization in processing phonemic and prosodic auditory changes in neonates, in preparation. Q5

Ashby, F.G., Alfonso-Reese, L.A., Turken, A.U., Waldron, E.M., 1998. A neuropsychological theory of multiple systems in category learning. Psychol. Rev. 105, 442–481.

Ashby, F.G., Ell, S.W., 2001. The neurobiology of human category learning. Trends Cogn. Sci. 5, 204–210.

Ashby, F.G., O'Brien, J.B., 2005. Category learning and multiple memory systems. Trends Cogn. Sci. 9, 83–89.

Baker, S.A., Idsardi, W.J., Golinkoff, R.M., Petitto, L.A., 2005. The perception of handshapes in American sign language. Mem. Cognit. 33, 887–904.

Belin, P., Zilbovicius, M., Crozier, S., Thivard, L., Fontaine, A., Masure, M.C., Samson, Y., 1998. Lateralization of speech and auditory temporal processing. J. Cogn. Neurosci. 10, 536–540.

Benson, R.R., Richardson, M., Whalen, D.H., Lai, S., 2006. Phonetic processing areas revealed by sinewave speech and acoustically similar non-speech. Neuroimage 31, 342–353.

Bertoncini, J., Morais, J., Bijeljac-Babic, R., McAdams, S., Peretz, I., Mehler, J., 1989. Dichotic perception and laterality in neonates. Brain Lang. 37, 591–605.

Best, C.T., Hoffman, H., Glanville, B.B., 1982. Development of infant ear asymmetries for speech and music. Percept. Psychophys. 31, 75–85.

Best, C.T., Mathur, G., Miranda, K.A., Lillo-Martin, D., 2010. Effects of sign language experience on categorical perception of dynamic ASL pseudosigns. Atten. Percept. Psychophys. 72, 747–762.

Binder, J.R., Frost, J.A., Hammeke, T.A., Bellgowan, P.S., Springer, J.A., Kaufman, J.N., Possing, E.T., 2000. Human temporal lobe activation by speech and nonspeech sounds. Cereb. Cortex 10, 512–528.

Boemio, A., Fromm, S., Braun, A., Poeppel, D., 2005. Hierarchical and asymmetric temporal sensitivity in human auditory cortices. Nat. Neurosci. 8, 389–395.

Bortfeld, H., Fava, F., Boas, D.A., 2009. Identifying cortical lateralization of speech processing in infants using near-infrared spectroscopy. Dev. Neuropsychol. 34, 52–65.

Boye, M., Gunturkun, O., Vauclair, J., 2005. Right ear advantage for conspecific calls in adults and subadults, but not infants, California sea lions (Zalophus californianus): hemispheric specialization for communication? Eur. J. Neurosci. 21, 1727–1732.

Bristow, D, Dehaene-Lambertz, G., Mattout, J., Soares, C., Gliga, T., Baillet, S., 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.

Buccino, G., Binkofski, F., Fink, G.R., Fadiga, L., Fogassi, L., Gallese, V., Seitz, R.J., Zilles, K., Rizzolatti, G., Freund, H.J., 2001. Action observation activates premotor and parietal areas in a somatotopic manner: an fMRI study. Eur. J. Neurosci. 13, 400–404.

Buxhoeveden, D.P., Switala, A.E., Litaker, M., Roy, E., Casanova, M.F., 2001. Lateralization of minicolumns in human planum temporale is absent in nonhuman primate cortex. Brain Behav. Evol. 57, 349–358.

Campbell, R., MacSweeney, M., Waters, D., 2008. Sign language and the brain: a review. J. Deaf Stud. Deaf Educ. 13, 3–20.

Carreiras, M., Lopez, J., Rivero, F., Corina, D., 2005. Linguistic perception: neural processing of a whistled language. Nature 433, 31–32.

Chomsky, N., Halle, M., 1968. The sound pattern of English.

Chomsky, N., Lasnik, H., 1993. The theory of principles and parameters. In: Syntax: An International Handbook of Contemporary Research, pp. 506–569.

Coleman, J., Pierrehumbert, J.B., 1997. Stochastic phonological grammars and acceptability. In: Computational Phonology: Third Meeting of the ACL Special Interest Group in Computational Phonology. Association for Computational Linguistics, Somerset, NJ, pp. 49–56.

Cristia, A., Egorova, N., Gervain, J., Cabrol, C., Minagawa-Kawai, Y., Dupoux, E. Socially relevant language in the infant brain, submitted for publication. Q6

Cristia, A., McGuire, G., Seidl, A., Francis, A.L. Effects of the distribution of cues in infants' perception of speech sounds. J. Phon., in press. Q7

Cristia, A., Seidl, A., Francis, A.L., 2011. Where do phonological features come from? Cognitive, physical and developmental bases of distinctive speech categories.

Curby, K.M., Hayward, G., Gauthier, I., 2004. Laterality effects in the recognition of depth-rotated novel objects. Cogn. Affect. Behav. Neurosci. 4, 100–111.

Davis, T., Love, B.C., Maddox, W.T., 2009. Two pathways to stimulus encoding in category learning? Mem. Cognit. 37, 394–413.

DeCasper, A.J., Lecanuet, J.-P., Busnel, M.-C., Granier-Deferre, C., Maugeais, R., 1994. Fetal reactions to recurrent maternal speech. Infant Behav. Dev. 17, 159–164.

Dehaene, S., Dupoux, E., Mehler, J., Cohen, L., Paulesu, E., Perani, D., van de Moortele, P.F., Lehericy, S., Le Bihan, D., 1997. Anatomical variability in the cortical representation of first and second language. Neuroreport 8, 3809–3815.

Dehaene-Lambertz, G., 1997. Electrophysiological correlates of categorical phoneme perception in adults. Neuroreport 8, 919–924.

Dehaene-Lambertz, G., Baillet, S., 1998. A phonological representation in the infant brain. Neuroreport 9, 1885–1888.

Dehaene-Lambertz, G., Dehaene, S., Hertz-Pannier, L., 2002. Functional neuroimaging of speech perception in infants. Science 298, 2013–2015.

Dehaene-Lambertz, G., Gliga, T., 2004. Common neural basis for phoneme processing in infants and adults. J. Cogn. Neurosci. 16, 1375–1387.

Dehaene-Lambertz, G., Hertz-Pannier, L., 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., Allirol, L., Dubois, J., Hertz-Pannier, L., Dehaene, S., 2010. Language or music, mother or

Mozart? Structural and environmental influences on infants' language networks. Brain Lang. 114, 53–65.

Dehaene-Lambertz, G., Pallier, C., Serniclaes, W., Sprenger-Charolles, L., Jobert, A., Dehaene, S., 2005. Neural correlates of switching from auditory to speech perception. Neuroimage 24, 21–33.

Devlin, J.T., Raley, J., Tunbridge, E., Lanary, K., Floyer-Lea, A., Narain, C., Cohen, I., Behrens, T., Jezzard, P., Matthews, P.M., Moore, D.R., 2003. Functional asymmetry for auditory processing in human primary auditory cortex. J. Neurosci. 23, 11516–11522.

Duclaux, R., Challamel, M.J., Collet, L., Roullet-Solignac, I., Revol, M., 1991. Hemispheric asymmetry of late auditory evoked response induced by pitch changes in infants: influence of sleep stages. Brain Res. 566, 152–158.

Duez, D., 2007. Consonant and vowel duration in Parkinsonian French speech. Travaux Interdisciplinaires du Labortoire Parole et Langage d'Aix-en-Provence 26, 15–31.

Duffau, H., 2008. The anatomo-functional connectivity of language revisited. New insights provided by electrostimulation and tractography. Neuropsychologia 46, 927–934.

Ehret, G., 1987. Left hemisphere advantage in the mouse brain for recognizing ultrasonic communication calls. Nature 325, 249–251.

Endress, A.D., Nespor, M., Mehler, J., 2009. Perceptual and memory constraints on language acquisition. Trends Cogn. Sci. 13, 348–353.

Filoteo, J.V., Maddox, W.T., Simmons, A.N., Ing, A.D., Cagigas, X.E., Matthews, S., Paulus, M.P., 2005. Cortical and subcortical brain regions involved in rule-based category learning. Neuroreport 16, 111–115.

Fitch, R.H., Brown, C.P., O'Connor, K., Tallal, P., 1993. Functional lateralization for auditory temporal processing in male and female rats. Behav. Neurosci. 107, 844–850.

Fitch, R.H., Tallal, P., Brown, C.P., Galaburda, A.M., Rosen, G.D., 1994. Induced microgyria and auditory temporal processing in rats: a model for language impairment? Cereb. Cortex 4, 260–270.

Fletcher, P., Buchel, C., Josephs, O., Friston, K., Dolan, R., 1999. Learning-related neuronal responses in prefrontal cortex studied with functional neuroimaging. Cereb. Cortex 9, 168–178.

Fodor, J.A., 1985. Precis of the modularity of mind. Behav. Brain Sci. 8, 1–5.

Forkstam, C., Hagoort, P., Fernandez, G., Ingvar, M., Petersson, K.M., 2006. Neural correlates of artificial syntactic structure classification. Neuroimage 32, 956–967.

Francis, A.L., Driscoll, C., 2006. Training to use voice onset time as a cue to talker identification induces a left-ear/right-hemisphere processing advantage. Brain Lang. 98, 310–318.

Friederici, A.D., 2002. Towards a neural basis of auditory sentence processing. Trends Cogn. Sci. 6, 78–84.

Friederici, A.D., 2009. Pathways to language: fiber tracts in the human brain. Trends Cogn. Sci. 13, 175–181.

Friederici, A.D., Alter, K., 2004. Lateralization of auditory language functions: a dynamic dual pathway model. Brain Lang. 89, 267–276.

Friederici, A.D., Bahlmann, J., Heim, S., Schubotz, R.I., 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.

Furuya, I., Mori, K., 2003. Cerebral lateralization in spoken language processing measured by multi-channel near-infrared spectroscopy (NIRS). No To Shinkei 55, 226–231.

Furuya, I., Mori, K., Minagawa-Kawai, Y., Hayashi, R., 2001. Cerebral Lateralization of Speech Processing in Infants Measured by Near-Infrared Spectroscopy. IEIC Technical Report (Institute of Electronics, Information and Communication Engineers) 100, 15–20.

Gallistel, C.R., 2000. The replacement of general-purpose learning models with adaptively specialized learning modules. In: Gazzaniga, M.S. (Ed.), The Cognitive Neurosciences. , 2d ed. MIT Press, Cambridge, MA, pp. 1179–1191.

Galuske, R.A., Schlote, W., Bratzke, H., Singer, W., 2000. Interhemispheric asymmetries of the modular structure in human temporal cortex. Science 289, 1946–1949.

Gandour, J., Tong, Y., Wong, D., Talavage, T., Dzemidzic, M., Xu, Y., Li, X., Lowe, M., 2004. Hemispheric roles in the perception of speech prosody. Neuroimage 23, 344–357.

Gandour, J., Wong, D., Lowe, M., Dzemidzic, M., Satthamnuwong, N., Tong, Y., Lurito, J., 2002. Neural circuitry underlying perception of duration depends on language experience. Brain Lang. 83, 268–290.

Gervain, J., Macagno, F., Cogoi, S., Pena, M., 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., Aslin, R.N., 2011. Near-infrared spectroscopy in cognitive developmental research. Dev. Cogn. Neurosci. 1, 22–46.

Gil-da-Costa, R., Braun, A., Lopes, M., Hauser, M.D., Carson, R.E., Herscovitch, P., Martin, A., 2004. Toward an evolutionary perspective on conceptual representation: species-specific calls activate visual and affective processing systems in the macaque. Proc. Natl. Acad. Sci. U. S. A. 101, 17516–17521.

Giraud, A.L., Kleinschmidt, A., Poeppel, D., Lund, T.E., Frackowiak, R.S., Laufs, H., 2007. Endogenous cortical rhythms determine cerebral specialization for speech perception and production. Neuron 56, 1127–1134.

Glanville, B.B., Best, C.T., Levenson, R., 1977. A cardiac measure of cerebral asymmetries in infant auditory perception. Dev. Psychol. 13, 54–59.

Goldstein, M.H., Schwade, J.A., 2008. Social feedback to infants' babbling facilitates rapid phonological learning. Psychol. Sci. 19, 515–523.

Golestani, N., Molko, N., Dehaene, S., LeBihan, D., Pallier, C., 2007. Brain structure predicts the learning of foreign speech sounds. Cereb. Cortex 17, 575–582.

Golestani, N., Paus, T., Zatorre, R.J., 2002. Anatomical correlates of learning novel speech sounds. Neuron 35, 997–1010.

Gonzalez, J., McLennan, C.T., 2009. Hemispheric differences in the recognition of environmental sounds. Psychol. Sci. 20, 887–894.

Graf Estes, K., Edwards, J., Saffran, J.R., 2011. Phonotactic constraints on infant word learning. Infancy 16, 180–197.

Granier-Deferre, C., Ribeiro, A., Jacquet, A.-Y., Bassereau, S., 2011. Near-term fetuses process temporal features of speech. Dev. Sci. 14, 336–352.

Grossmann, T., Oberecker, R., Koch, S.P., Friederici, A.D., 2010. The developmental origins of voice processing in the human brain. Neuron 65, 852–858.

Grossmann, T., Oberecker, R., Koch, S.P., Friederici, A.D. The developmental origins of voice processing in the human brain. Neuron 65, 852–858, in press. [Q8]

Haggard, M.P., Parkinson, A.M., 1971. Stimulus and task factors as determinants of ear advantages. Q. J. Exp. Psychol. 23, 168–177.

Hall, T.A., 2001. Distinctive feature theory. Mouton de Gruyter, Berlin.

Hauser, M.D., Andersson, K., 1994. Left hemisphere dominance for processing vocalizations in adult, but not infant, rhesus monkeys: field experiments. Proc. Natl. Acad. Sci. U. S. A. 91, 3946–3948.

Hayes, T.L., Lewis, D.A., 1993. Hemispheric differences in layer III pyramidal neurons of the anterior language area. Arch. Neurol. 50, 501–505.

Heinz, J. in press-a. Computational phonology part I: Foundations. Lang. Ling. Compass. [Q9]

Heinz, J. in press-b. Computational phonology part II: Grammars, learning, and the future. Lang. Ling. Compass. [Q10]

Hickok, G., Holt, L.L., Lotto, A.J., 2009. Response to Wilson: What does motor cortex contribute to speech perception? Trends Cogn. Sci. 13, 330–331.

Hickok, G., Poeppel, D., 2004. Dorsal and ventral streams: a framework for understanding aspects of the functional anatomy of language. Cognition 92, 67–99.

Hickok, G., Poeppel, D., 2007. The cortical organization of speech processing. Nat. Rev. Neurosci. 8, 393–402.

Holt, L.L., Lotto, A.J., 2010. Speech perception as categorization. Atten. Percept. Psychophys. 72, 1218–1227.

Homae, F., Watanabe, H., Nakano, T., Asakawa, K., Taga, G., 2006. The right hemisphere of sleeping infant perceives sentential prosody. Neurosci. Res. 54, 276–280.

Homae, F., Watanabe, H., Nakano, T., Taga, G., 2007. Prosodic processing in the developing brain. Neurosci. Res. 59, 29–39.

Homae, F., Watanabe, H., Nakano, T., Taga, G. Large-scale brain networks underlying language acquisition in early infancy. Front. Psychol. 2, in press. [Q11]

Horvath, J., Czigler, I., Sussman, E., Winkler, I., 2001. Simultaneously active pre-attentive representations of local and global rules for sound sequences in the human brain. Brain Res. Cogn. Brain Res. 12, 131–144.

Iacoboni, M., 2009. Imitation, empathy, and mirror neurons. Annu. Rev. Psychol. 60, 653–670.

Imada, T., Zhang, Y., Cheour, M., Taulu, S., Ahonen, A., Kuhl, P.K., 2006. Infant speech perception activates Broca's area: a developmental magnetoencephalography study. Neuroreport 17, 957–962.

Ivry, R., Robertson, L., 1998. The Two Sides of Perception. The MIT Press, Cambridge, MA.

Jacquemot, C., Pallier, C., LeBihan, D., Dehaene, S., 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., Matthews, P.M., 2006. Hemispheric specialization for processing auditory nonspeech stimuli. Cereb. Cortex 16, 1266–1275.

Jancke, L., Wustenberg, T., Scheich, H., Heinze, H.J., 2002. Phonetic perception and the temporal cortex. Neuroimage 15, 733–746.

Joanisse, M.F., Gati, J.S., 2003. Overlapping neural regions for processing rapid temporal cues in speech and nonspeech signals. Neuroimage 19, 64–79.

Jusczyk, P.W., Cutler, A., Redanz, N.J., 1993a. Infants' preference for the predominant stress patterns of English words. Child Dev. 64, 675–687.

Jusczyk, P.W., Friederici, A.D., Wessels, J.M.I., Svenkerud, V.Y., Jusczyk, A.M., 1993b. Infants' sensitivity to the sound patterns of native language words. J. Mem. Lang. 32, 402–420.

Kell, C.A., Morillon, B., Kouneiher, F., Giraud, A.L., 2011. Lateralization of speech production starts in sensory cortices—A possible sensory origin of cerebral left dominance for speech. Cereb. Cortex, 932–937.

Kenstowicz, M.J., 1994. Phonology in Generative Grammar. Blackwell, Cambridge, MA.

Kenstowicz, M.J., Kisseberth, C.W., 1979. Generative Phonology: Description and Theory. Academic Press, San Diego, California.

Koivisto, M., Revonsuo, A., 2004. Preconscious analysis of global structure: Evidence from masked priming. Visual Cogn. 11, 105–127.

Kotilahti, K., Nissila, I., Nasi, T., Lipiainen, L., Noponen, T., Merilainen, P., Huotilainen, M., Fellman, V., 2010. Hemodynamic responses to speech and music in newborn infants. Hum. Brain Mapp. 31, 595–603.

Kuhl, P.K., Andruski, J.E., Chistovich, I.A., Chistovich, L.A., Kozhevnikova, E.V., Ryskina, V.L., Stolyarova, E.I., Sundberg, U., Lacerda, F., 1997. Cross-language analysis of phonetic units in language addressed to infants. Science 277, 684–686.

Kuhl, P.K., Williams, K.A., Lacerda, F., Stevens, K.N., Lindblom, B., 1992. Linguistic experience alters phonetic perception in infants by 6 months of age. Science 255, 606–608.

Leech, R., Holt, L.L., Devlin, J.T., Dick, F., 2009. Expertise with artificial nonspeech sounds recruits speech-sensitive cortical regions. J. Neurosci. 29, 5234–5239.

Lenneberg, E.H., 1966. Biological Foundations of Language. Willry, New York.

Ley, R.G., Bryden, M.P., 1982. A dissociation of right and left hemispheric effects for recognizing emotional tone and verbal content. Brain Cogn. 1, 3–9.

Liebenthal, E., Desai, R., Ellingson, M.M., Ramachandran, B., Desai, A., Binder, J.R., 2010. Specialization along the left superior temporal sulcus for auditory categorization. Cereb. Cortex 20, 2958–2970.

Liegeois, F., Connelly, A., Cross, J.H., Boyd, S.G., Gadian, D.G., Vargha-Khadem, F., Baldeweg, T., 2004. Language reorganization in children with early-onset lesions of the left hemisphere: an fMRI study. Brain 127, 1229–1236.

MacSweeney, M., Campbell, R., Woll, B., Giampietro, V., David, A.S., McGuire, P.K., Calvert, G.A., Brammer, M.J., 2004. Dissociating linguistic and nonlinguistic gestural communication in the brain. Neuroimage 22, 1605–1618.

MacSweeney, M., Woll, B., Campbell, R., McGuire, P.K., David, A.S., Williams, S.C., Suckling, J., Calvert, G.A., Brammer, M.J., 2002. Neural systems underlying British Sign Language and audio-visual English processing in native users. Brain 125, 1583–1593.

Marsolek, C.J., Burgund, E.D., 2008. Dissociable neural subsystems underlie visual working memory for abstract categories and specific exemplars. Cogn. Affect. Behav. Neurosci. 8, 17–24.

Mattock, K., Molnar, M., Polka, L., Burnham, D., 2008. The developmental course of lexical tone perception in the first year of life. Cognition 106, 1367–1381.

Mattys, S.L., Jusczyk, P.W., Luce, P.A., Morgan, J.L., 1999. Phonotactic and prosodic effects on word segmentation in infants. Cogn. Psychol. 38, 465–494.

Maye, J., Weiss, D.J., Aslin, R.N., 2008. Statistical phonetic learning in infants: facilitation and feature generalization. Dev. Sci. 11, 122–134.

Mazoyer, S., Lalle, P., Narod, S.A., Bignon, Y.J., Courjal, F., Jamot, B., Dutrillaux, B., Stoppa-Lyonnett, D., Sobol, H., 1993. Linkage analysis of 19 French breast cancer families, with five chromosome 17q markers. Am. J. Hum. Genet. 52, 754–760.

McNealy, K., Mazziotta, J.C., Dapretto, M., 2006. Cracking the language code: neural mechanisms underlying speech parsing. J. Neurosci. 26, 7629–7639.

Mehler, J., Bertoncini, J., Barriere, M., 1978. Infant recognition of mother's voice. Perception 7, 491–497.

Mehler, J., Endress, A.D., Gervain, J., Nespor, M., 2008. From perception to grammar. Early language development: Bridging brain and behaviour. Trends Lang. Acquisition Res. (TiLAR) 5, 191–213.

Meyer, M., Alter, K., Friederici, A.D., Lohmann, G., von Cramon, D.Y., 2002. FMRI reveals brain regions mediating slow prosodic modulations in spoken sentences. Hum. Brain Mapp. 17, 73–88.

Miller, J.L., Liberman, A.M., 1979. Some effects of later-occurring information on the perception of stop consonant and semivowel. Percept. Psychophys. 25, 457–465.

Minagawa-Kawai, Y., Mori, K., Hebden, J.C., 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., 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., 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., Kojima, S., 2009b. Cerebral laterality for phonemic and prosodic cue decoding in children with autism. Neuroreport 20, 1219–1224.

Minagawa-Kawai, Y., Naoi, N., Kojima, S., 2009a. New approach to functional neuroimaging: Near Infrared Spectroscopy. Keio University Press.

Minagawa-Kawai, Y., van der Lely, H., Ramus, F., Sato, Y., Mazuka, R., Dupoux, E., 2011. Optical brain imaging reveals general auditory and language-specific processing in early infant development. Cereb. Cortex 21, 254–261.

Molfese, D.L., 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.

Morillon, B., Lehongre, K., Frackowiak, R.S., Ducorps, A., Kleinschmidt, A., Poeppel, D., Giraud, A.L., 2010. Neurophysiological origin of human brain asymmetry for speech and language. Proc. Natl. Acad. Sci. U. S. A. 107, 18688–18693.

Morton, J., Johnson, M.H., 1991. CONSPEC and CONLERN: a two-process theory of infant face recognition. Psychol. Rev. 98, 164–181.

Mottonen, R., Calvert, G.A., Jaaskelainen, I.P., Matthews, P.M., Thesen, T., Tuomainen, J., 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., Pons, F., Fais, L., Dietrich, C., Werker, J.F., Amano, S., 2009. Perception of vowel length by Japanese- and English-learning infants. Dev. Psychol. 45, 236–247.

Naatanen, R., Lehtokoski, A., Lennes, M., Cheour, M., Huotilainen, M., Iivonen, A., Vainio, M., Alku, P., Ilmoniemi, R.J., Luuk, A., Allik, J., Sinkkonen, J., Alho, K., 1997. Language-specific phoneme representations revealed by electric and magnetic brain responses. Nature 385, 432–434.

Nazzi, T., Jusczyk, P.W., Johnson, E.K., 2000. Language discrimination by English-learning 5-month-olds: effects of rhythm and familiarity. J. Mem. Lang. 43, 1–19.

Nespor, Vogel, I., 1986. Prosodic phonology. Foris, Dordrecht. Q12

Neville, H.J., Bavelier, D., Corina, D., Rauschecker, J., Karni, A., Lalwani, A., Braun, A., Clark, V., Jezzard, P., Turner, R., 1998. Cerebral organization for language in deaf and hearing subjects: biological constraints and effects of experience. Proc. Natl. Acad. Sci. U. S. A. 95, 922–929.

Newman, A.J., Bavelier, D., Corina, D., Jezzard, P., Neville, H.J., 2002. A critical period for right hemisphere recruitment in American Sign Language processing. Nat. Neurosci. 5, 76–80.

Novak, G.P., Kurtzberg, D., Kreuzer, J.A., Vaughan Jr., H.G., 1989. Cortical responses to speech sounds and their formants in normal infants: maturational sequence and spatiotemporal analysis. Electroencephalogr. Clin. Neurophysiol. 73, 295–305.

Nowak, P.M., 2006. The role of vowel transitions and frication noise in the perception of Polish sibilants. J. Phonetics 34, 139–152.

Obrig, H., Rossi, S., Telkemeyer, S., Wartenburger, I., 2010. From acoustic segmentation to language processing: evidence from optical imaging. Front. Neuroenergetics, 2.

Opitz, B., Friederici, A.D., 2003. Interactions of the hippocampal system and the prefrontal cortex in learning language-like rules. Neuroimage 19, 1730–1737.

Opitz, B., Friederici, A.D., 2004. Brain correlates of language learning: the neuronal dissociation of rule-based versus similarity-based learning. J. Neurosci. 24, 8436–8440.

Peña, M., Maki, A., Kovacic, D., Dehaene-Lambertz, G., Koizumi, H., Bou-quet, F., 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.

Penhune, V.B., Zatorre, R.J., MacDonald, J.D., Evans, A.C., 1996. Interhemispheric anatomical differences in human primary auditory cortex: probabilistic mapping and volume measurement from magnetic resonance scans. Cereb. Cortex 6, 661–672.

Perani, D., Dehaene, S., Grassi, F., Cohen, L., Cappa, S.F., Dupoux, E., Fazio, F., Mehler, J., 1996. Brain processing of native and foreign languages. Neuroreport 7, 2439–2444.

Perani, D., Paulesu, E., Galles, N.S., Dupoux, E., Dehaene, S., Bettinardi, V., Cappa, S.F., Fazio, F., Mehler, J., 1998. The bilingual brain. Proficiency and age of acquisition of the second language. Brain 121 (10), 1841–1852.

Peretz, I., 1990. Processing of local and global musical information by unilateral brain-damaged patients. Brain 113 (4), 1185–1205.

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., van Wassenhove, V., 2008. Speech perception at the interface of neurobiology and linguistics. Philos. Trans. R. Soc. Lond. B Biol. Sci. 363, 1071–1086.

Poizner, H., Klima, E.S., Bellugi, U., 1987. What the Hands Reveal about the Brain. The MIT Press, Cambridge, MA.

Poremba, A., Malloy, M., Saunders, R.C., Carson, R.E., Herscovitch, P., Mishkin, M., 2004. Species-specific calls evoke asymmetric activity in the monkey's temporal poles. Nature 427, 448–451.

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.

Querleu, D., Renard, X., Versyp, F., Paris-Delrue, L., Crepin, G., 1988. Fetal hearing. Eur. J. Obstet. Gynecol. Reprod. Biol. 28, 191–212.

Rivera-Gaxiola, M., Csibra, G., Johnson, M.H., Karmiloff-Smith, A., 2000. Electrophysiological correlates of cross-linguistic speech perception in native English speakers. Behav. Brain Res. 111, 13–23.

Rogers, J., Hauser, M. The use of formal languages in artificial language learning: a proposal for distinguishing the differences between human and nonhuman animal learners. In Recursion and Human Language, edited by Harry van der Hulst, pp. 213–232. Mouton De Gruyter, Berlin, Q13 Germany.

Rogers, J., Hauser, M., 2010. The use of formal languages in artificial language learning: a proposal for distinguishing the differences between human and nonhuman animal learners. Recursion Hum. Lang., 213–232.

Rosen, S., 1992. Temporal information in speech: acoustic, auditory and linguistic aspects. Philos. Trans. R. Soc. Lond. B Biol. Sci. 336, 367–373.

Rybalko, N., Suta, D., Nwabueze-Ogbo, F., 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., 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., Mori, K., Furuya, I., Hayashi, R., Minagawa-Kawai, Y., Koizumi, T., 2003. Developmental changes in cerebral lateralization to spoken language in infants: measured by near-infrared spectroscopy. Jpn. J. Logopedics Phoniatric 44, 165–171.

Sato, Y., Sogabe, Y., Mazuka, R., 2007. Brain responses in the processing of lexical pitch-accent by Japanese speakers. Neuroreport 18, 2001–2004.

Sato, Y., Sogabe, Y., Mazuka, R., 2010. Development of hemispheric specialization for lexical pitch-accent in Japanese infants. J. Cogn. Neurosci. 22, 2503–2513.

Schonwiesner, M., Rubsamen, R., von Cramon, D.Y., 2005. Hemispheric asymmetry for spectral and temporal processing in the human anterolateral auditory belt cortex. Eur. J. Neurosci. 22, 1521–1528.

Scott, S.K., Blank, C.C., Rosen, S., Wise, R.J., 2000. Identification of a pathway for intelligible speech in the left temporal lobe. Brain 123 (Pt 12), 2400–2406.

Scott, S.K., Johnsrude, I.S., 2003. The neuroanatomical and functional organization of speech perception. Trends Neurosci. 26, 100–107.

Seldon, H.L., 1981. Structure of human auditory cortex. II: Axon distributions and morphological correlates of speech perception. Brain Res. 229, 295–310.

Shankweiler, D., Studdert-Kennedy, M., 1967. Identification of consonants and vowels presented to left and right ears. Q. J. Exp. Psychol. 19, 59–63.

Simos, P.G., Molfese, D.L., 1997. Electrophysiological responses from a temporal order continuum in the newborn infant. Neuropsychologia 35, 89–98.

Simos, P.G., Molfese, D.L., Brenden, R.A., 1997. Behavioral and electrophysiological indices of voicing-cue discrimination: laterality patterns and development. Brain Lang. 57, 122–150.

Stephan, K.E., Fink, G.R., Marshall, J.C., 2007. Mechanisms of hemispheric specialization: insights from analyses of connectivity. Neuropsychologia 45, 209–228.

Strange, W., Bohn, O.S., 1998. Dynamic specification of coarticulated German vowels: perceptual and acoustical studies. J. Acoust. Soc. Am. 104, 488–504.

Szaflarski, J.P., Binder, J.R., Possing, E.T., McKiernan, K.A., Ward, B.D., Hammeke, T.A., 2002. Language lateralization in left-handed and ambidextrous people: fMRI data. Neurology 59, 238–244.

Telkemeyer, S., Rossi, S., Koch, S.P., Nierhaus, T., Steinbrink, J., Poeppel, D., Obrig, H., Wartenburger, I., 2009. Sensitivity of newborn auditory cortex to the temporal structure of sounds. J. Neurosci. 29, 14726–14733.

Tervaniemi, M., Hugdahl, K., 2003. Lateralization of auditory-cortex functions. Brain Res. Brain Res. Rev. 43, 231–246.

Turkeltaub, P.E., Coslett, H.B., 2010. Localization of sublexical speech perception components. Brain Lang. 114, 1–15.

Vargha-Khadem, F., Corballis, M.C., 1979. Cerebral asymmetry in infants. Brain Lang. 8, 1–9.

Vihman, M., 1996. Phonological Development: The Origins of Language in the Child.

Vouloumanos, A., Kiehl, K.A., Werker, J.F., Liddle, P.F., 2001. Detection of sounds in the auditory stream: event-related fMRI evidence for differential activation to speech and nonspeech. J. Cogn. Neurosci. 13, 994–1005.

Werker, J.F., Tees, R.C., 1984. Phonemic and phonetic factors in adult cross-language speech perception. J. Acoust. Soc. Am. 75, 1866–1878.

Werker, J.F., Yeung, H.H., 2005. Infant speech perception bootstraps word learning. Trends Cogn. Sci. 9, 519–527.

Wetzel, W., Ohl, F.W., Wagner, T., Scheich, H., 1998. Right auditory cortex lesion in Mongolian gerbils impairs discrimination of rising and falling frequency-modulated tones. Neurosci. Lett. 252, 115–118.

White, K.S., Morgan, J.L., 2008. Sub-segmental detail in early lexical representations. J. Mem. Lang. 59, 114–132.

Whitehouse, A.J., Bishop, D.V., 2009. Hemispheric division of function is the result of independent probabilistic biases. Neuropsychologia 47, 1938–1943.

Xu, Y., Gandour, J., Talavage, T., Wong, D., Dzemidzic, M., Tong, Y., Li, X., Lowe, M., 2006. Activation of the left planum temporale in pitch processing is shaped by language experience. Hum. Brain Mapp. 27, 173–183.

Yamazaki, Y., Aust, U., Huber, L., Hausmann, M., Gunturkun, O., 2007. Lateralized cognition: asymmetrical and complementary strategies of pigeons during discrimination of the "human concept". Cognition 104, 315–344.

Yuen, I., Davis, M.H., Brysbaert, M., Rastle, K., 2009. Activation of articulatory information in speech perception. Proc. Natl. Acad. Sci. U. S. A. 107, 592–597.

Zaehle, T., Geiser, E., Alter, K., Jancke, L., Meyer, M., 2008. Segmental processing in the human auditory dorsal stream. Brain Res. 1220, 179–190.

Zaehle, T., Wustenberg, T., Meyer, M., 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., Evans, A.C., Meyer, E., Gjedde, A., 1992. Lateralization of phonetic and pitch processing in speech perception. Science 256, 846–849.

Zatorre, R.J., 1988. Pitch perception of complex tones and human temporal-lobe function. J. Acoust. Soc. Am. 84, 566–572.

Zatorre, R.J., Belin, P., 2001. Spectral and temporal processing in human auditory cortex. Cereb. Cortex 11, 946–953.

Zatorre, R.J., 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.

Zatorre, R.J., Meyer, E., Gjedde, A., Evans, A.C., 1996. PET studies of phonetic processing of speech: review, replication, and reanalysis. Cereb. Cortex 6, 21–30.

Zeithamova, D., Maddox, W.T., Schnyer, D.M., 2008. Dissociable prototype learning systems: evidence from brain imaging and behavior. J. Neurosci. 28, 13194–13201.

Zhang, Y., Kuhl, P.K., Imada, T., Iverson, P., Pruitt, J., Stevens, E.B., Kawakatsu, M., Tohkura, Y., Nemoto, I., 2009. Neural signatures of phonetic learning in adulthood: a magnetoencephalography study. Neuroimage 46, 226–240.