Functional asymmetry between consonants and vowels from birth to 6 months of age: Cerebral imaging and behavioral data.

Camillia Bouchon

UNIVERSITÉ PARIS DESCARTES

SORBONNE PARIS CITÉ

ECOLE DOCTORALE ED261:
« COGNITION, COMPORTEMENT, CONDUITES HUMAINES »

Asymétrie fonctionnelle entre consonnes et voyelles de la naissance à l'âge de 6 mois
Données d'imagerie cérébrale et de comportement

Functional asymmetry between consonants and vowels from birth to 6 months of age
Cerebral imaging and behavioral data

Camillia BOUCHON

Thèse de doctorat de Sciences Cognitives

Dirigée par Thierry NAZZI et Judit GERVAIN

Présentée et soutenue publiquement le lundi 24 novembre 2014 devant un jury composé de

TORO SOTO Juan Manuel	Professeur	Rapporteur
ZESIGER Pascal	Professeur	Rapporteur
PEPERKAMP Sharon	Directrice de recherche	Examinateur
WARTENBURGER Isabell	Professeure	Examinateur
NAZZI Thierry	Directeur de recherche	Directeur de thèse
GERVAIN Judit	Chargée de recherche	Co-Directrice de thèse

Résumé

Consonnes et voyelles sont les deux catégories de sons qui composent la parole. Elles se distinguent à divers niveaux et notamment servent des fonctions linguistiques différentes. Cette asymétrie consonne/voyelle établie chez les adultes, a conduit Nespor, Peña et Mehler (2003) à suggérer un partage du travail dès la naissance, les consonnes facilitant l'acquisition des mots tandis que les voyelles aideraient à apprendre les règles de grammaire. La validité développementale de cette hypothèse est explorée par l'étude de ses origines chez les bébés français.

Premièrement, nos études d'imagerie cérébrale optique montrent que consonnes et voyelles sont également traitées par les mécanismes précurseurs de l'apprentissage syntaxique à la naissance (Exp. 1 - 3). Deuxièmement, nos études sur la reconnaissance du prénom chez les enfants de 5 mois montrent une sensibilité à une modification vocalique (Alix/Elix) chez les bébés monolingues, mais pas à une modification consonantique en position initiale (Victor/Zictor) chez les bébés monolingues et bilingues, ou finale chez les monolingues (Luca/Luga; Exp. 4 - 9). Au stade des premiers mots, le traitement lexical privilégie donc les voyelles.

Nos résultats contribuent à la compréhension des origines développementales de l'asymétrie fonctionnelle consonne/voyelle, et du rôle spécifique de la langue native dans son émergence.

Mots clés: Consonnes et voyelles, traitement lexical et syntaxique, NIRS, nouveau-nés, reconnaissance du prénom.

Abstract

Speech is composed of two categories of sound, i.e. consonants and vowels, which have different properties and serve different linguistic functions. This consonant/vowel asymmetry, which is established in adults, has led Nespor, Peña and Mehler (2003) to suggest a division of labor present from birth, whereby consonants would facilitate lexical acquisition while vowels would help to learn grammatical rules of language. We have explored the developmental validity of this hypothesis by studying its origins in French-learning infants.

Firstly, our optical brain imaging studies show that both consonants and vowels provide input for precursory mechanisms of syntax processing (Exp. 1 - 3). Secondly, our studies on own-name recognition at 5 months demonstrate sensitivity to a vowel mispronunciation in monolingual infants (Alix/Elix), but fail to show a reaction to a consonant mispronunciation in initial position (Victor/Zictor) for monolinguals and bilinguals, or in final position (Luca/Luga) for monolinguals (Exp. 4 - 9). Thus, vowels are a better input for lexical processing in first familiar words.

Our results contribute to the understanding of the developmental origin of consonant/vowel functional asymmetry, hence the influence of the native input on its emergence.

Keywords: Consonants and vowels, lexical and syntactic processing, NIRS, neonates, own-name recognition.

Acknowledgments – Remerciements

First and foremost, I am deeply grateful to my supervisors Thierry Nazzi and Judit Gervain. Thank you both for your benevolent and watchful guidance during these four years. I feel incredibly lucky to have benefited from two such role models. Judit, ton optimisme à toute épreuve et ta confiance ont été très importants. Thierry, merci d'être à la fois exigeant sans jamais te départir de ta gentillesse.

I am very grateful to Juan Manuel Toro Soto, Pascal Zesiger, Isabell Wartenburger and Sharon Peperkamp for doing me the honor of being part of my thesis committee. Thank you for sparing the time to read and comment my thesis and participating to my defense.

C'est à Caroline Floccia que je dois l'opportunité imprévue d'avoir commencé cette thèse sur le langage chez les bébés. J'en suis très reconnaissante. Merci à toi, ainsi qu'à Claire pour la fructueuse collaboration qui s'en est suivie.

Je tiens à remercier Martine Adda-Decker et Thibaut Fux pour leur disponibilité lors de notre collaboration et de leur intérêt pour nos bébés et leurs prénoms. Merci à Lionel Granjon d'assurer la traduction.

Je dois à Christian Lorenzi mon premier contact révélateur avec la recherche et d'avoir fait le Cogmaster. Merci à Christian, à l'ensemble des chercheurs du Cogmaster et en particulier à Franck Ramus et Sharon Peperkamp d'avoir encadré mes stages en Master. Je vous suis très reconnaissante de tout ce que j'ai appris avec vous et qui m'a amenée à continuer.

Après 4 ans au LPP, ce ne sera pas facile de quitter un environnement aussi riche scientifiquement qu'humainement. Merci aux membres de l'équipe Parole du LPP : Josette, Viviane, Lionel, Lucie, Delphine, Josiane, Ranka, Pia, Louise, Katie, Henny, Silvana (the best conference roomate ever!) et Scania (merci pour ses encouragements et précieux conseils).

Merci aux nombreux bébés qui ont participé à mes expériences et à leurs parents pour leur intérêt et disponibilité pour la recherche.

Merci aux sages-femmes et à toute l'équipe soignante de la maternité Robert Debré d'avoir facilité les passations avec les nouveau-nés et le contact avec les parents.

Merci aux *officemates* de la 608 pour l'entre-aide et les rires indispensables a tout bon travail, et bravo de supporter mes manies/humeurs/vols de gâteaux compulsifs etc. Aux anciennes qui nous manquent fort: Laurianne, Lauriane; et les actuels: Nawal, Léo, Louah, Carline et Cécile.

Un grand merci au Docteur Cohen-Zardi.

Les jongleurs historiques: Jonathan, Lila, Kim, Sylvain, Pascal et Lia, Merci !

Yann, merci pour je ne sais pas combien de kilos de Beaufort d'été !

Pierre, merci pour le Tartuffe, le gîte, le couvert et bien d'autres choses encore

Romain, merci pour ta patience et ta présence de dernière minute :)

Les arméniens, Audrey, Alexe, Lucie, Lyre, Julie et Paul, merci pour les raclettes endiablées, vivement la prochaine.

Guillaume, allé rien que pour toi : « Consonnes… voyelles… consonnes… » Laurent Romejko.

Anne, (+ Thomas): Thank you for the most inspirational Amsterdam break !

Marie, combien d'heures au téléphone, de mails, de sms? Cette année c'est la bonne.

Mille fois merci à Sarah, ma sarachou, et à Souad et Christian, mes parents, pour leur soutien et leur confiance. Et maintenant, des bulles !

List of papers presented in this dissertation

Bouchon, C., Nazzi, T., & Gervain, J. (in revision). Hemispheric asymmetries in repetition enhancement and suppression effects in the newborn brain.

Bouchon, C., Floccia, C., Fux, T., Adda-Decker, M., & Nazzi, T. (in press). Call me Alix, not Elix: Vowels are more important than consonants in own name recognition at 5 months. *Developmental Science*.

Contents

List of Figures

List of Tables

CHAPTER 1 – GENERAL INTRODUCTION

For more than 50 years now, researchers in language acquisition have been looking for and defining the innate perceptual constraints and powerful learning mechanisms that conspire to ensure the phonological, lexical and syntactic achievements infants have to attain by twelve months to start producing language. Originally, the nativist theory of language defined by Chomsky (1959) encouraged investigations in early language acquisition whereby innate linguistic competence would be unmasked. Following Liberman's work at the Haskins laboratories revealing the categorical perception of speech sounds in adults, developmental psycholinguists studied infants' speech perception and showed that infants start with universal discrimination of a broad range of phonetic contrasts (Eimas, Siqueland, Jusczyk, & Vigorito, 1971; Kuhl, Williams, Lacerda, Stevens, & Lindblom, 1992; Werker & Tees, 1984) and attune to the phonology of their native language before the end the first year of life, approximately at the age they start producing their first words. The linguistic functions at multiple levels (phonology, lexicon, syntax) do not appear sequentially in development, but rather emerge interactively through powerful learning mechanisms, whereby existing knowledge at a given level of organization facilitates the acquisition of other linguistic functions (Morgan & Demuth, 1996; Werker & Yeung, 2005).

Nespor, Peña, and Mehler (2003) proposed a bootstrapping mechanism based on the functional asymmetry in the processing of consonants and vowels in language, which could help infants learn the lexical and structural (prosodic/syntactic) aspects of their native language. This hypothesis proposes that: *"Based on evidence coming from the nature of grammar and the lexicon as well as language acquisition and language loss, we have proposed that there is a division of labour between vowels and consonants: vowels are specialised for conveying information about grammar and consonants about the lexicon. This is a plausible scenario if the different role of vowels and consonants is part of UG, so that human beings come into the world knowing that languages are structured in such a way that consonants, rather than vowels, are most relevant to build the lexicon, and vowels, rather than consonants, are most relevant for grammatical information." (page 224-225).* Nespor et al. (2003) thus suggest that this functional asymmetry could originate in two initial perceptual biases present before any exposure to language: a consonant bias, according to which consonants would intrinsically be used in lexically-relevant mechanisms and a vowel bias leading vowels to trigger the detection of structures.

This division of labor between consonants and vowels is illustrated in the world's languages, particularly in Semitic languages, in which semantic roots are mostly consonantal and varying vowels derive the different morphological forms. This functional asymmetry is observed in non-Semitic

languages as well, for instance in the ablaut phenomena in Germanic languages, whereby vowel variations in a verbal consonantal root signify different grammatical variants, e.g. *sing, sang, sung* in English; *singen, sang, gesungen* in German; *stelen, stal, gestolen* in Dutch (i.e. to steal). Another example comes from writing abbreviations, in which we often remove vowels and keep consonants to avoid compromising meaning (e.g. *txt msgng* for text-messaging). A large number of empirical findings from several disciplines, and from many of the world's languages show that consonants play a more important role in carrying lexical information, while vowels play a more important role in conveying structures (prosodic and morphosyntactic).

This proposal has launched a large number of investigations in adults and infants over the last ten years. Among them, this thesis aims at investigating whether consonants and vowels are two functionally distinct categories intrinsically providing the means of analyzing the speech stream into words and rules at birth and in the first 6 months of life, or whether their respective linguistic functions need to be learned through speech exposure. In other words, our general goals are:

- to determine the extent to which the consonant and vowel biases proposed by Nespor et al. (2003) belong to the category of initial perceptual constraints bootstrapping infants from a complex acoustic signal to a meaningful organized language system;
- to find out how early infants use consonants and vowels differentially to learn and recognize words, and extract and generalize rules, respectively.

This theoretical framework offers an ideal opportunity for investigating interactions in the early processing and acquisition of the acoustic/phonetic, phonological, lexical and syntactic levels of language. Specifically, the present dissertation will investigate the existence and role of the consonant/vowel asymmetry (later on, C/V asymmetry) in learning words vs. structural rules at two early stages of development: *at birth*, when infants are first exposed to speech in its full form, and *at 5 months,* when infants are building their first entries in their lexicon.

The present dissertation is thus organized into two experimental chapters bearing on French-learning infants. The neuroimaging experiments in the first experimental chapter explore with optical imaging the origins of the C/V division of labor following the work of Gervain et al. (2008) on the detection of repetition-based regularities of syllables. We explore the newborn brain's detection of repetition-based regularities of *consonants* or *vowels*, when these repetitions occur in two different contexts: one favoring *rule-learning* (for which Nespor et al., 2003 predict a vowel bias*)*, and the other favoring *item*-memorization (for which a consonant bias is predicted). We thus directly investigate the relative weight given to consonants and vowels in a *syntax-like ability* versus

a *lexically-related process* in the newborn brain. In the second experimental chapter, our main goal is to further investigate *the consonant bias* in lexical processing at the age of 5 months, i.e. at the *onset* of lexical acquisition. Using a behavioral familiar word preference task testing infants on the recognition of their own name (Mandel et al., 1995), reactions to mispronunciations will be measured to assess the sensitivity of 5-month-olds to *a consonant* versus a *vowel change* in one of their *first familiar word-form*.

In the rest of the Introduction, we will review the arguments of Nespor et al. (2003) and update them with more recent reports suggesting that at various levels of observation including linguistic function, consonants and vowels are to be considered categorically. Then we will present the results of recent infant studies investigating the origins of the C/V asymmetry, which according to Nespor et al.'s (2003) original hypothesis should be innate. The results of these recent studies gave rise to alternative hypotheses, which the experimental work conducted in this dissertation will help evaluate.

Nespor et al. (2003) thus presented data from many disciplines supporting a functional categorical distinction between consonants and vowels, from which they conclude that consonants are intrinsically more relevant to represent and process *words*, while vowels are more suited to process structures and patterns, as found at the *prosodic and morphosyntactic* levels. In the following we present and complement these data with results from more recent psychological data on the processing of consonants and vowels and explore whether they lead to the same conclusion as in Nespor et al. (2003) or whether the original proposal needs to be modified.

I. 1. Acoustic, articulatory and basic speech perception differences between consonants and vowels

Consonants and vowels are the two segmental categories that compose speech, and together form the syllables that make up utterances (Ladefoged, 1993). Acoustically, vowels are more salient than consonants. Indeed, in the speech stream they are typically longer, louder and constitute the most periodic portions of speech (Cutler & Mehler, 1993). Consonants and vowels also differ acoustically in their level of sonority and continuity (Ladefoged, 1993, 2001). When forming vowels the airstream passes through the vocal tract uninterrupted by constrictions. Vowels are therefore continuous and sonorant sounds. When forming consonants, the airstream is

obstructed in some way through the vocal tract, so that consonants tend to be less sonorant. Non-continuous consonants include plosives, which are realized by the complete closure of the articulators so that the airstream cannot escape through the mouth (/p t k b d g/). However, continuous consonant subclasses are more sonorant, e.g nasals (/m n/), liquids (/l R/) and fricatives (/f s ʃ v z ʒ/). As a result, although consonants and vowels are different classes of sounds, some consonants are acoustically more vowel-like than others, as captured by the sonority scale, on which vowels are at the top and plosive consonants at the bottom.

Moreover, due to the anatomy of the speech tract when producing consonants, consonant categories are acoustically more distinct from one another than vowel categories. However, the articulatory discontinuity between continuous consonants (e.g. between /s/ and / ʃ/ or between /R/ and /l/) is less clear than between plosives (e.g./b/ and /d/, Ladefoged, 1993, 2001). Thus, phonetic contrasts might increase in distinctiveness as contrasted sounds are lower on the sonority scale

In the natural speech stream, the effects of coarticulation modulate the production and consequently the acoustic characteristics of a given sound depending on its surrounding vowels and consonants (Liberman et al., 1954). Vowels tend to be more affected than consonants, as more factors modulate their duration, formant structure etc. These factors are voicing of the surrounding consonants, emphatic stress, focus, position in an utterance, and affect (for an overview see Erickson, 2000).

Do the differences listed so far affect how adults perceive consonants and vowels and discriminate consonant and vowel contrasts? Languages differ in the set of consonants and vowels they use, and in the set of lexically meaningful, i.e. phonemic, contrasts. Adults typically have difficulties discriminating non-native phonetic differences that are not contrasting meaning in their language (e.g., Werker & Tees, 1984), although they are able to make some such discrimination (e.g., Best, McRoberts, & Sithole, 1988). Liberman and colleagues revealed that with similarly distant pairs of sounds along an acoustic continuum, pairs belonging to different phonetic categories were better discriminated than pairs belonging to the same phonetic category. Therefore, subjects are more sensitive in the regions of the phoneme boundary than at other points of the acoustic continuum, suggesting that speech sounds are perceived categorically. Moreover, along the same acoustic continuum of speech sounds, adults use a similar categorical label for the tokens falling on one side of the discrimination boundary, and another label for the tokens falling on the other side of the boundary (Liberman, Harris, Hoffman, & Griffith, 1957; Liberman, Cooper, Shankweiler, & Studdert-Kennedy, 1967). This pattern of perception of speech sounds has been called categorical perception, and it is relies on specialized neural processing mechanisms involved (Näätänen, 2001).

Interestingly, although adult categorical perception is readily and robustly demonstrated with stop consonants, within-category discrimination is often quite good for vowels, in particular when presented in isolation (Fry, Abramson, Eimas & Liberman, 1962; Liberman et al., 1967; Stevens & Blumstein, 1978; Repp, 1984). For instance, adult discrimination abilities were compared, for synthetic vowel sounds varying on a vowel continuum (/I - ɛ - œ/) and for consonant sounds in a consonant continuum (/b - d - g/; Fry et al., 1962). Results revealed that, as for consonants, vowel categories were discriminated and labeled correspondingly to their phonemic category. However, unlike for the consonant continuum, adults could also distinguish small acoustic/phonetic differences between tokens of the same vowel category. Therefore, although categorical perception is found for both native vowel and consonant categories in adults, vowel perception seems to be more continuous than consonant perception (Pisoni, 1973).

Moreover, some models of native and non-native phoneme categorization introduced the notion of perceptual magnet effect to characterize vowel perception (Kuhl, 1991). Indeed, adults show greater generalization of neighboring tokens when the prototype vowel (e.g. /i/) serves as a referent than when a non-prototypical sound serves as referent. Thus, prototypical vowels act as perceptual magnets for other category members (***Fig. I. 1***. Adults).

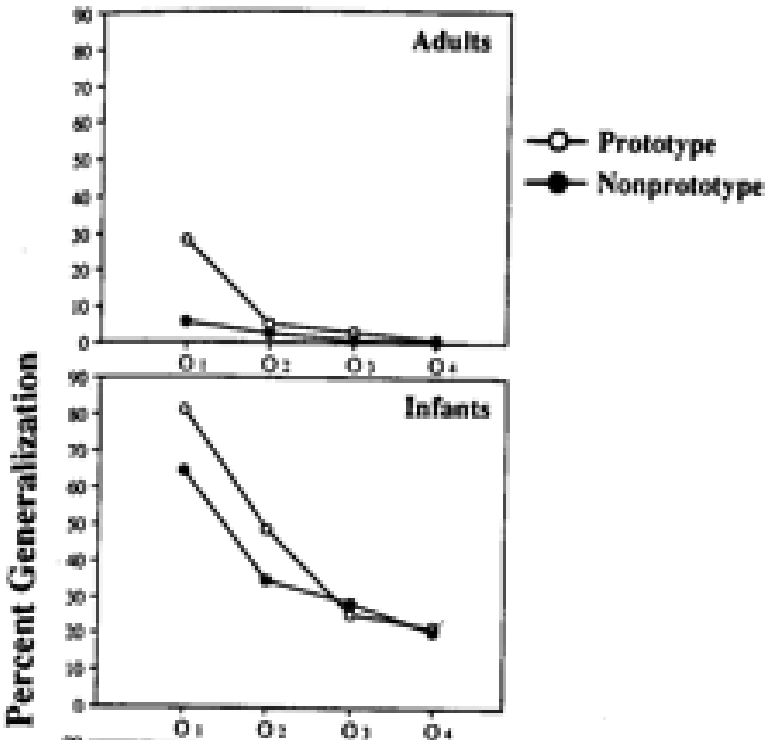


Figure I. 1. *Average generalization scores (percentage of trials in which subjects indicated that they perceived the two stimuli as the same) in subjects in the prototype condition (white dots) and in the non-prototype condition (black dots), as a function of distance between the two stimuli (from O1 to O4, on the x axis, From Kuhl, 1991).*

Thus as for consonant categories, adults have vowel categories; however, contrary to consonants, they rather easily perceive differences within vowel categories, and the prototype magnet effect introduces directional effects in sensitivity to vowel changes. This suggests two particular modes of perception differentiating vowels from consonants.

Interestingly, categorical perception for consonants and vowels, and prototype-based perception for vowels is present in animals too (for consonants: in monkeys, Morse & Snowdon, 1975; Waters & Wilson, 1976; Kuhl & Padden, 1982, 1983; Sinnott, 1989; in rodents: Kuhl & Miller, 1978; Kuhl, 1981; Sinnott & Kreiter, 1991; in birds, Dooling, Best, & Brown, 1995; for vowels: in cats, Dewson, 1964), and after brief exposures to vowel categories, rats and birds show the prototype magnet effect (Kluender, Lotto, Holt, & Bloedel, 1998; Pons, 2006). The fact that categorical perception of consonants and prototype-based perception of vowels are not specific to the human perceptual system suggests that this may be present in primates' common ancestors, and perhaps even beyond. Moreover, it indicates that these perceptual effects are grounded in the different acoustic and articulatory properties of the two speech classes.

Overall, adults categorically perceive the consonants of their native language, while they perceive vowels more continuously. These distinct perceptual sensitivities to consonants and vowels may be based on their acoustic and articulatory characteristics. How do infants develop speech perception abilities to yield the more categorical perception for consonants and the prototype-based perception for vowels observed in adults?

I. 2. Acquisition differences: timing differences – different modes (consonant categories, vowel prototypes)

Initially, infants' discrimination of speech sounds covers a broader range of phonetic contrasts than in adults. Indeed, infant speech perception studies demonstrated that almost all phonetic distinctions used across languages are initially discriminated, thus that infants start out with language-general perceptual abilities (Eimas et al., 1971; Streeter, 1976; Jusczyk, 1997; Cheour-Luhtanen et al., 1995; Dehaene-Lambertz & Baillet, 1998). Eimas et al. (1971) first tested phonetic discrimination abilities in 1- and 4-month-old infants using a High Amplitude Sucking Procedure (HAS). This procedure measures the rate of high-amplitude sucks produced by infants during the

presentation of stimuli. Eimas et al. (1971) presented infants with syllables in which the consonant varied continuously in VOT (Voice Onset Time). Infants were habituated to one stimulus until their sucking responses decreased to a habituation criterion. One group of infants was exposed to a VOT change corresponding to a change in consonant category (+20 ms versus +40 ms, corresponding respectively to /ba/ and /pa/ as identified by adults). A second group was exposed to a VOT change of similar magnitude but that did not correspond to a phonetic category difference (both sounds perceived as /ba/ or /pa/ by adults). A last, control group was presented with the same sound throughout the experiment. Results show that sucking rates increase for a between-category change but not for a within-category change (nor in the absence of change), indicating categorical perception for consonants in 1- and 4-month-old infants.

Dehaene-Lambertz and Baillet (1998) later explored the neural basis of this categorical perception of consonants in 3-month-olds and measured event-related potentials (ERPs). They found that the electrophysiological correlate for similar acoustic changes (on VOT) was larger between phonetic categories than within phonetic category, thus revealing some sensitivity at the brain level to within-category changes while confirming early categorical perception of consonant sounds.

Vowel perception was explored at birth using ERPs (Cheour-Luhtanen et al., 1995). Newborns were exposed to repetitions of /y/ vowels with occurrences of two possible deviant stimuli: either an /i/ vowel, or an intermediate /y-i/ vowel on the /i-y/ continuum. Results showed that only the phonetic change from /y/ to /i/ elicited a significant Mismatch Negativity response (i.e. electro-physiological correlate typically signaling the detection of a deviant stimulus) similar to the adult response, with a peak 200ms after the stimulus. These studies thus demonstrate accurate consonant and vowel categorical discrimination in the first months of life.

Later in development phonetic perception undergoes perceptual attunement: as infants are exposed to their native language, their discrimination abilities narrow down to native contrasts during their first year. Interestingly, infants' perception becomes language-specific for vowels before consonants (Werker & Tees, 1984; Kuhl, et al., 1992; Kuhl, 2004). The first native vowel prototypes are found at 6 months (Kuhl, et al., 1992; Polka & Werker, 1994; Cheour et al., 1998), and continue to develop until at least 8–12 months of age (Bosch & Sebastián-Gallés, 2003b; Polka & Bohn, 1996). Native consonant categories are first built by 8.5 months of age (Best et al., 1988; Werker & Lalonde, 1988; Werker & Tees, 1984; Anderson, Morgan, & White, 2003), and are still narrowing down at 10–12 months of age (see Saffran, Werker, & Werner, 2007 for a review). Therefore, the perception of

vowels seems to be affected by the native language earlier in development than the perception of consonants.

In addition, by 6 months, infants' vowel discrimination shows the same directional effect as in adults, characteristic of a prototypical magnet effect (Kuhl, 1991; Kuhl et al., 1992, ***Fig. I. 1.*** Infants). Vowel instances that are prototypical in the language exhibit a strong perceptual pull on less prototypical vowel instances, making discrimination more difficult around the prototypes.

In sum, we have reviewed data showing that perceptual attunement differs between consonants and vowels. First, different perceptual modes were found: consonants and vowels are perceived as categories, but consonant perception is more categorical and vowel perception more continuous with a magnet effect. Second, perceptual attunement to consonants and vowels are shifted in time, consonants being acquired later than vowels. These differences add to our previous description of how consonants and vowels differ acoustically. In this context, do consonants and vowels also differ functionally in language?

I. 3. Typological evidence for consonants and vowels carrying different linguistic information in the world's languages

Other arguments used by Nespor et al. (2003) relate to typological evidence. Indeed, in most language systems, consonants are more numerous than vowels, partly for articulatory reasons, and consequently they provide more opportunities to distinguish words. The average consonant/vowel ratio (number of consonants divided by number of vowels) calculated for 563 languages of the world reach on average 4.25, and 50% of these languages have 3.5 times more consonants than vowels (Maddieson, 2011). An extreme case of consonant-rich languages is the Australian language Diyari (Pama-Nyungan; South Australian) with 22 consonants but only three vowel qualities. There are also low CV ratio exceptions with languages having less than 10 consonants (e.g. Hawaiian) but the number of vowels is usually even smaller. Cases where the number of vowels and consonants is balanced are rare (Andoke: 10 consonants and 9 vowels; Swedish and French: 16 consonants and 17 vowels).

In this context, Danish is an exception, as it has a much richer vowel than consonant inventory, with 31 vowel contrasts and 15 consonant contrasts (Grønnum, 1998). Moreover, contrary to most other languages, it undergoes extensive consonant reduction, possibly making consonants harder to distinguish than in other languages, and vowels perceptually more distinctive than consonants (Bleses, Basbøll & Vach, 2011; Grønnum, 1998; Pharao, 2011). Therefore, as will be later described in adult and infant studies, Danish is a critical case for the investigation of the C/V functional hypothesis, given that the relatively reduced informativeness of consonants compared to vowels could disfavor a consonant bias in lexical processing.

Linguistic evidence also shows that consonants carry more lexical information while vowels convey more morphosyntactic cues. For instance, consonants tend to disharmonize within a word, i.e. to alternate in quality, while vowels tend to harmonize within words, yielding words like *rotolo* 'roll' in Italian. Further, many languages such as Turkish, Hungarian, Finnish etc., show vowel harmony in their morphology, such that affixes harmonize with stems. Vowel harmony can thus be informative about boundaries of domains larger than words, e.g. morphosyntactic domains (Nespor & Vogel, 1986). Vowels are also more often reduced by prosodic phenomena at the word level, e.g. vowels in unstressed positions often loose their quality (e.g. in English, Cutler, Mehler, Norris, & Segui, 1986). On the contrary, consonants rarely lose their lexical distinctiveness (although some consonants may undergo weakening through consonant reduction/lenition, e.g. in Danish, as will be later described). In Semitic languages, lexical roots are mostly consonantal and morphology is provided by the insertion of vowels. For example, several Hebrew words deriving from the same meaning "round" share the same *svv* root and vary in their vowels (***Table I. 1***).

Word	Hebrew Transcription	Gloss
sivúv	סיבוב	A turn
sévev	סבב	A round
sovév	סובב	He turned
sevivón	סביבון	A dradle
svivá	סביבה	Surroundings

Table I. 1. Illustration of Hebrew words generated from the lexical root svv (From Berent et al., 2007)

However, the C/V functional distinction in languages is relative. For instance, consonants may be involved in syntax (e.g. the liaison in French whereby consonants signal syntactic constituents, Nespor & Vogel, 1986; Selkirk, 1974) and vowels can also carry lexical information (e.g.

the tense/lax /I/-/i/ contrast in English distinguishes many minimal pairs of words: *sheep-ship; peek-pick; feet-fit; leek-lick; seek-sick; beat-bit…* see Toro, Shukla et al., 2008). Therefore, the division of labor in the worlds' languages is described as "partial" by Nespor et al. (2003), i.e. probabilistic rather than deterministic.

To conclude, in most languages consonants outnumber vowels (Maddieson, 2011) and in general, they appear to carry more lexical information than vowels. The frequent phenomena of vowel harmony and consonant disharmony within words result in consonants alternating in quality, which reinforces their lexical informativeness. Vowels on the other hand, tend to vary in quantity, thus providing prosodic cues that signal morphosyntactic properties across and within words.

The predominant role of vowels over consonants in carrying prosodic information (Nespor & Vogel, 1986) is instantiated in lexical stress in English (whereby the stress pattern of a word has a lexically distinctive function) and its correlational link with the quality of the vowel in unstressed syllables (Cutler et al., 1986). Indeed, within a pair of homophone words with opposite stress patterns, the vowel of the initial syllable of the strong-weak word is full (e.g. CONtent), while in the weak-strong word (e.g conTENT), it becomes a schwa, i.e. it loses its vowel quality. This type of phenomenon may be of great importance for the C/V functional asymmetry because through prosodic processing infants may have access to segmental information, and in particular to vowel information. Indeed, infants are highly sensitive to prosodic information very early on as evidenced by the role of prosody in early speech perception (Mehler et al., 1988; Jusczyk, Friederici, Wessels, Svenkerud, & Jusczyk, 1993; Jusczyk, Houston, & Newsome, 1999; Nazzi, Bertoncini, & Mehler, 1998; Friederici, Steinhauer, & Pfeifer, 2002; Thiessen & Saffran, 2003; Weissenborn & Höhle, 2001; Kuhl, 2004; Friederici, 2005; Sansavini, Bertoncini, & Giovanelli, 1997; Nazzi, Floccia & Bertoncini, 1998). Acoustically, prosodic cues are variations of pitch contour, intensity, and duration, and they occur at different levels: segments, syllables, phrases and sentences. They are used to convey accentuation, intonation (melodic contour carried by the fundamental frequency or F0), rhythm… The womb filters out most of the fine details of speech but preserves prosodic cues such as intonational contours and rhythm; therefore exposure to prosody begins in utero. Prosodic cues convey syntactic information, e.g. they contribute to the interpretation of ambiguous sentence (e..g. "*le couplet complet - le couple est complet"*, for a review, see Cutler, Dahan & Donselaar, 1997). Prosodic cues also signal word boundaries and newborns are able to rely on prosodic cues to discriminate between two phonemically identical bisyllables (e.g. *mati)*, when one is the internal part of a word (*mathematician*) and the other overlaps a prosodic boundary (*pyjama tigré*) (Christophe, Dupoux, Bertoncini, & Mehler, 1994), even in the case of non-native prosodic cues (Christophe, Mehler, &

Sebastiàn-Gallés, 2001). Overall this suggests that the early ability to use prosodic cues to process syntactic structures (Morgan & Demuth, 1996) and track lexical boundaries is potentially interacting with their knowledge of the consonants and vowels.

An example of such an interaction is provided by one of the properties of prosody, rhythm. According to an early proposal, languages were divided into rhythmic classes depending on their predominant rhythmic unit (Abercrombie, 1967; Pike, 1945; Ladefoged, 1993, 2001): most Germanic languages (and Arabic) have a rhythm based on the inter-stress interval (i.e. the alternation of strong and weak syllables); most Romance languages have a rhythm is based on the syllable (e.g., French, Italian, Spanish); other languages such as Japanese have a mora-based rhythm (the mora is a metrical unit that can either be a V or CV syllable, or the subsyllabic second half of a long vowel, post-vocalic nasal consonant, or post-vocal portion of a geminate consonant). Later on, several metrics have been proposed to define linguistic rhythm, which depend on the alternations of consonants and vowels in the signal (e.g. %V, the average proportion of vocalic intervals, ΔV the average standard deviations of and vocalic intervals in Ramus, Nespor, & Mehler, 1999; but see also Grabe & Low, 2002 and White & Mattys, 2007, for a comparative evaluation of different metrics for the quantification of speech rhythm).

Newborns are able to classify different languages based on their rhythm (Nazzi et al., 1998). Could they start learning about the proportions of consonants and vowels in their native language through their sensitivity to rhythm? Indeed the rhythmicity of a language is reflected in the relative timings of vocalic and consonantal spaces to such an extent that when speech is resynthesized so that all consonants are replaced by /s/ and all vowels by /a/ (the *sasasa* manipulation Ramus, Hauser, Miller, Morris, & Mehler, 2000), adults are still able to discriminate rhythmically different languages, but fail if all consonants are replaced by /a/ (Ramus & Mehler, 1999). Crucially, newborns can also discriminate *sasasa* languages differing in rhythmicity (Ramus et al., 2000), suggesting that newborns as well as adults might process rhythm based on the relative intervals occupied by consonants and vowels in the speech signal.

To conclude, linguistic evidence shows a relative division of labor between consonants and vowels. So far, we reviewed how consonants and vowels differ on their physical and perceptual properties, as well as the linguistic roles they seem to play in languages of the world. Do adults process the two categories of speech sounds according to their functional assignments? Are consonants better cues for lexical processing and vowels for morphosyntactic processing?

I. 4. Dissociations between consonant and vowel processing: Psychological evidence of in adults

We will now review evidence that consonants and vowels are processed differently at a functional level, starting by some neuropsychological evidence indicating that distinct neural networks support the processing of consonants and vowels in adults.

I. 4. a. Neuropsychological evidence of C/V double dissociation in the brain

Contrasting patterns of errors when producing vowels and consonants have been reported in evidence of patients with brain lesions. Cubelli (1991) first described two patients with dysgraphia (i.e. deficit in written language production) who showed a selective deficit in writing vowels. Caramazza, Chialant, Capasso and Miceli (2000) later reported data from two patients with aphasia (i.e. deficit in language comprehension and production) showing contrasting, selective difficulties in producing vowels or consonants, demonstrating a case of double dissociation for consonant and vowel processing in the adult brain. Indeed one patient with vowel processing errors in speech production had a lesion in the left parietal and temporal lobes and a small lesion in the right parietal lobe, while a second patient with consonant processing errors in speech production had a lesion in left supramarginal, angular and superior temporal gyri. A number of control studies showed that the relative sonority of the consonant is not a parameter that plays a role in the amount of difficulty the first patient presents with the processing of consonants, leading the authors to conclude that different neural mechanisms are responsible for the processing of vowels and consonants. Finally Ferreres, López and China (2003) describe a patient with alexia (i.e. deficit in written word comprehension) exhibiting more errors on vowels than on consonants when reading non-words. Overall, these results in brain-damaged patients are indications of the processing of consonants and vowels in the adult brain.

I. 4. b. Consonants and vowels in lexical and syntactic processing

We will now review the main behavioral and neuroimaging investigations of the C/V functional asymmetry proposed by Nespor et al. (2003) in adults.

I. 4. b. i. The consonant bias in lexical processing

The lexically-related consonant bias is evidenced by a greater reliance on consonants over vowels in lexical learning and lexical access. At present, the consonant bias in adulthood has been found in the written and oral modalities, in many different tasks and for languages with different C/V ratios: English, Dutch, Spanish, French, Portuguese and Italian.

Word intelligibility studies suggest a more important role for consonants than for vowels (Owren & Cardillo, 2006; Kewley-Port, Burkle & Lee, 2007). In a lexical judgment task, words with only consonants present (vowels excised and replaced by noise) were intelligible, while in words with only the vowels preserved (consonants replaced by noise) listeners better recognized the indexical information about the speaker (Owren & Cardillo, 2006).

Nespor et al. (2003) provided evidence coming from *word reconstruction studies* to support their hypothesis of a consonant bias in speech processing. In these studies, participants had to transform a pseudoword into a real word by changing only one phoneme. Results showed that English listeners would change *ke*bra into *co*bra rather than to *z*ebra, thus preferring to keep the consonant tier constant rather than the vocalic tier (in English: van Ooijen, 1996; in Dutch and Spanish: Cutler, Sebastiàn-Gallés, SolerVilageliu, & van Ooijen, 2000). However, these were offline tasks and *priming studies* were later used to confirm this bias in online speech processing.

For instance, in *visual priming experiments* involving replaced-letter experiments, the primes consist in whole consonant tier or whole vowel tier replacement from the target word. Results show that adults rely more on consonants than on vowels, thus an advantage for consonant- over vowel-related primes (e.g. *duvo* primes *diva* more than *rifa* does, Duñabeitia & Carreiras, 2011; Vergara-Martínez et al., 2011; in French: New & Nazzi, 2014; New, Araújo, and Nazzi, 2008; in French: for European Portuguese: Acha & Perea, 2010; Soares, Perea & Comesaña, 2014; for English: Lee, Rayner & Pollatsek, 2002; in Spanish: Carreiras, Gillon-Dowens, Vergara & Perea, 2009; Carreiras, Dunabeitia & Molinaro, 2009). In one *auditory priming study,* English and French adults were tested in a lexical decision task in which auditory targets (e.g. /*sinema*/, 'cinema') were preceded by auditory prime non-words (Delle Luche, Poltrock, Goslin, New, Floccia & Nazzi, 2014). The results in this auditory modality established an overall facilitator effect of a prime sharing the consonant tier (/*synøma*/) with the target but not for a prime sharing the vowel tier (/*timema*/). Thus, priming studies consistently converge towards larger consonantal than vocalic priming effects, hence towards a consonant bias.

In word learning studies, French adults identified a newly learned word faster when it differed from a distractor by one consonant (/*pyv/ - /tyv/* respectively) than when it differs by a vowel (/*pos/-/poes/*, Havy, Serres, & Nazzi, 2014). Similarly, in a different word-learning paradigm with English adults, newly learned words sharing their consonants were more often confused (/*suba/-/sabo/*) than those sharing their vowels (/*nasi/-/tagi/,* Creel, Aslin & Tanenhaus, 2006). Both results show that adults attend more to consonants when learning new words, an event that occurs very frequently, as adults are estimated to learn an average of three word forms per day, every day of the year (Nation & Waring, 1997).

A direct evaluation of the relative role of consonants and vowels has also been conducted using segmentation studies relying on *auditory artificial language paradigms,* implementing syllable transitional probability (TPs) differences within words and between words. Indeed, adults (and infants) can perform statistical computations to detect the boundaries of words in continuous speech, by keeping track of TPs between syllables (Saffran, Aslin, & Newport, 1996b; Saffran, Newport & Aslin, 1996a). Thus, words can be distinguished from part-words based on the probability of syllables to appear adjacently within a continuous artificial stream of speech. In a study based on this ability, Bonatti, Peña, Nespor and Mehler (2005) tested the hypothesis that vowels give cues mainly about syntax while consonants give cues mainly about the lexicon. French-speaking adults were familiarized to a continuous stream of speech in which either the TPs of the consonant tier or the TPs of the vowel tier could serve to spot words. Results shows that at test, adults identified, 'families' of words with TPs based on the consonant tier (/*puragi/-/puregy/*), and not families of words with TPs based on the vowel tier (/*pokima/-/porila/*). This result was confirmed in another language, Italian (Toro, Nespor, Mehler & Bonatti, 2008; Mehler, Peña, Nespor, & Bonatti, 2006). Thus, French- and Italian-speaking adults use consonants preferentially when identifying word-like units from continuous speech.

However, some preliminary unpublished data in Danish-speaking adults using a similar paradigm suggest that this consonant bias may not be universal (Trecca, Bleses, Christiansen, Basbøll, Højen, Madsen & Andersen, 2014). As previously described, the relatively reduced informativeness of consonants compared to vowels in Danish should disfavor a consonant bias in lexical processing, so that Danish is a critical case for the investigation of the C/V functional hypothesis. In a similar artificial language segmentation paradigm, adults with different native languages, crucially involving Danish, as well as Norwegian and American English, were tested on their ability to segment words from strings of concatenated CV-syllables. Two artificial language conditions were used. In the "contoid language" condition, plosive consonants and vowels were

alternating, whereas in the "vocoid condition", consonants were replaced by vocoids (semivowels). The authors hypothesized that in the vocoid condition; Danish adults would perform better than the other groups. Subjects' performance in distinguishing words from part-words at test failed to reveal an effect of condition, indicating that unlike the authors' prediction, the contoid language was not easier to segment than the vocoid language. Moreover, subjects' performance also failed to show an effect of native language. However, Danish-speaking adults were significantly faster than other groups. Additionally, they were significantly faster in the vocoid- than contoid condition, unlike Norwegian- and American English-speaking adults. These results indicate that adults with a native language with vowel-rich phonetic structure may have better vowel-based segmentation abilities than others (Bonatti et al., 2005; Mehler et al., 2006). However, the precise extent of the vocalic bias for segmentation suggested by this unpublished study remains to be confirmed.

Neuroimaging and neurophysiological evidence also support the view that different underlying cortical networks are involved in consonant vs. vowel processing in lexically related tasks (Carreiras & Price, 2008; Carreiras, Vergara, & Perea, 2009; Vergara-Martínez, Perea, Marín, & Carreiras, 2011). For instance, Sharp, Scott, Cutler and Wise (2005) used PET (Positron emission tomography, a hemodynamic neuroimaging technique) to measure English participants' brain responses when performing the *kebra-cobra-zebra* task (Cutler et al., 2000). They confirmed the results from Van Ooijen (1996) and Cutler et al. (2000) that a consonant substitution (*kebra-zebra*) is more difficult than a vowel substitution (*kebra-cobra*). Moreover, they found a higher left-lateralized cortical activation in the inferior frontal gyrus (IFG, Broadmann's area BA44, BA45, BA47) and in the premotor cortex (PMC, BA6) for consonant over vowel word reconstruction (***Fig. I.2.***), which are areas known for their role in working memory retrieval of words through the activation of, and selection among, alternative candidates from the mental lexicon (Cutler & Clifton, 1999).

Overall various language groups and methods have yielded converging evidence of a consonant bias in adult lexical processing. Is the vowel bias for syntactic processing as well established in adults as the lexical part of the C/V functional asymmetry appears to be?

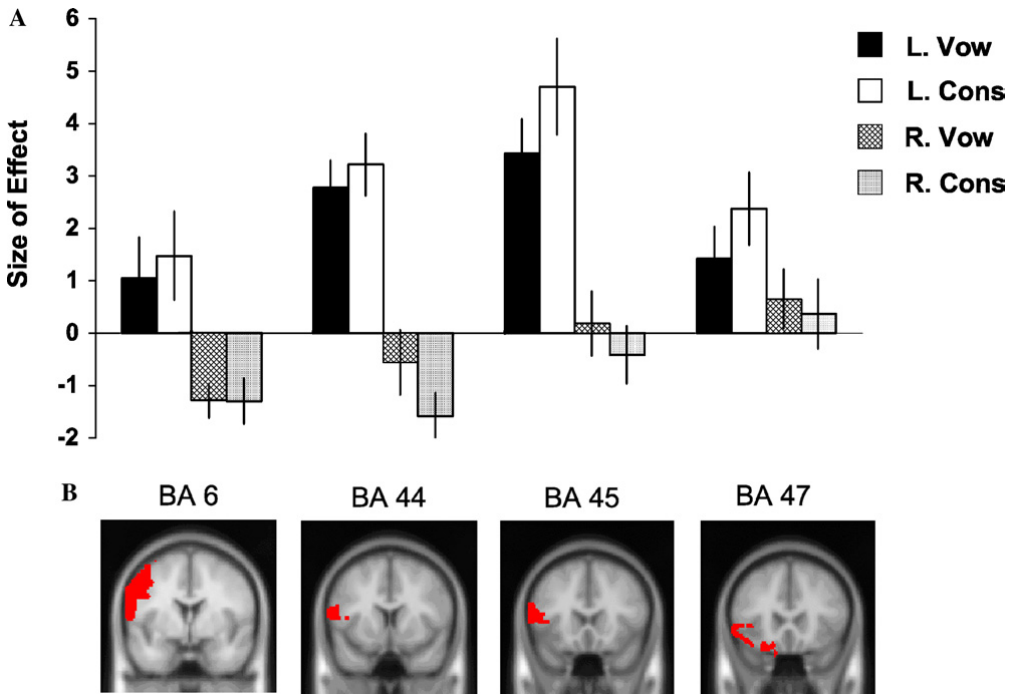


Figure I. 2. A. *Mean activity (and SEs) in the Left (L) and the Right (R) hemispheres' ROIs, in response to vowel (Vow) and consonant (Cons) word reconstructions.* ***B.*** *Localizations of the 4 ROIs (on structural MRI coronal slices): BA 6 = PMC; BA 44 = posterior IFG and BA 45 & 47 = anterior IFG (from Sharp et al., 2005).*

I. 4. b. ii. The vowel bias in morphosyntactic processing

To explore whether adults' lexical processing and structural generalization could occur simultaneously on separate levels of representation in speech (the consonantal and the vocalic tiers, respectively), Toro, Nespor, Mehler and Bonatti (2008) used an artificial-language learning paradigm. In four experiments, they tested whether statistical computations allowing for word segmentation preferentially rely on consonants, and whether vowels are better cues for structural processing. For 10 min, Italian participants were exposed to a continuous stream of syllables in which CVCVCV words could be segmented in one tier (i.e. the sequence of consonants in Exp. 1 or vowels in Exp. 2), and repetition-based regularities (ABA) could be extracted from the other tier (vice versa). Indeed in Experiment 1 the familiarization words could only be reliably segmented by computing consonant TPs, while the vowel tier followed an ABA pattern (the first and last vowels of words were identical while consonants were all different, e.g. …*tapena…bedoke…*). After familiarization participants were tested on their ability to segment words and on their ability to generalize rule-words to novel items. Generalization test sequences that had not appeared in the stream included rule-words for which the familiarized structure was implemented (e.g. *biduki*) whereas non-rule words followed an ABB structure (e.g. *buduki*). Segmentation test sequences included words isolated from the familiarization stream (e.g. *tapena*), and part-words overlapping familiarization words and complying with the vocalic ABA pattern (e.g. *penabe*).

In Experiment 1, participants considered that rule words and words were more likely to belong to the artificial language than non-rule-words and part-words, thus showing that they had extracted and generalized the structure implemented by the vowel tier and segmented the sequences by computing consonant TPs. In Experiment 2, when words could be segmented from vowels, and repetition-based regularities could be extracted from consonants, participants failed in both tasks. When 25-ms pauses were introduced at word edges (Exp. 3) or when words implemented a simpler repetition AAA rule (e.g. *babube, nanune,* Exp. 4), word segmentation was facilitated but participants still failed at distinguishing rule from non-rule words based on consonant repetition. Overall, this study demonstrated for the first time on a single speech input the division of labor between consonants and vowels in lexical segmentation vs. structures, thus emphasizing the specificity of vowels in generalizing repetition-based regularities.

However, as previously described, vowels and consonants differ acoustically, thus structural extraction may be targeting vowels preferentially because they are more salient acoustically and carry more energy than consonants. Toro, Shukla, Nespor and Endress (2008) investigated the contribution of the higher superior salience of vowels compared to consonants in adults' preferential use of vowel information in generalizing repetition-based regularities. When the rule was embedded in barely audible vowels, adults readily demonstrated their ability to extract and generalize the rule. In two experiments the authors then tried to promote generalization of an ABA structure embedded in consonants by using sonorant consonants (e.g. *raneru, linole…)* or by reducing vowel duration to a third of that of consonants. Adults recognized the words, but failed to generalize the rule. Even when vowels were totally eliminated from the stream, they only succeeded marginally in generalizing the rule. These results confirm that adults process vowels and consonants as different functional categories, and especially that vowels are more fitted than consonants for learning grammatical regularities, independently of their superior acoustic saliency in the signal presented to adults (although a possible influence of this factor earlier in development needs to be investigated).

To conclude on adult consonant and vowel processing, the cross-linguistic body of literature available since 2003 has confirmed the presence of the two biases in adult processing. They are now well established in many languages (especially the consonant bias for lexical processing), thus supporting the language-general hypothesis for the division of labor predicted by Nespor et al. (2003), although the case of Danish suggests some possible crosslinguistic modulation. The consonant bias in particular, is found in the written and in the oral modality, in behavioral and neuroimaging studies, and supported by neuropsychological data showing double dissociation for

consonant and vowel processing and differential underlying cerebral processing in the adult brain. Thus, consonants and vowels seem to differentially fit the specific requirements of lexical vs. structural processing, respectively, and adults are able to exploit this asymmetry functionally. What about the empirical data in toddlers and infants?

I. 5. New insights from infant studies

Nespor et al. (2003) proposed that the consonant and vowel biases may be innate and bootstrap infants' lexical and morphosyntactic acquisition. This has led to many studies with children and infants. As will be described below, the new empirical findings yield new predictions concerning the emergence of the consonant bias in lexical processing, and raise issues regarding the origins of the C/V functional distinction in development.

I. 5. a. The consonant bias in lexical processing

I .5. a. i. The 3 hypotheses on the origins of the consonant bias and their predictions

As previously described, the C/V asymmetry and especially the consonant bias in lexical processing seem quite stable in adulthood across languages. Nespor et al. (2003) predict that the consonant bias should be observed from birth and throughout development in lexically-related tasks. However, as we have described, during the very first stages of speech perception, vowels seem to be processed preferentially due to their high acoustic salience, and their prominent role in prosody. Infants may therefore learn that consonants are lexically more informative through linguistic experience, before they can exploit the consonant bias to learn new words. Two alternative proposals to Nespor et al.'s (2003) original innate bias hypothesis have been made regarding the developmental trajectory of the consonant bias.

According to the lexical hypothesis (Keidel, Jenison, Kluender & Seidenberg, 2007), the C/V functional asymmetry would arise from differences in the statistical distribution of consonants and vowels across languages and the degree to which they are informative to code the lexicon. Therefore, infants could not use the higher informativeness of consonants at the lexical level before they have reached a certain stage of lexical acquisition. This 'lexically-related emergence' hypothesis thus implies that the consonant bias would not emerge before 12 months at least, once a sizeable lexicon has been acquired. Moreover, it predicts that language-specific factors modulating the relative informativeness of consonants over vowels at the lexical level will yield crosslinguistic

differences in the timeline of emergence of the consonant bias. Danish-learning infants, for instance, are exposed to a language richer in vowels than in consonants and in which consonants are often reduced. Therefore, the lesser informativeness of consonants specific to Danish might not favor the emergence of a consonant bias in lexical processing. On the other hand, French exhibits a relatively atypical balanced C/V ratio, and consonant tiers are distributionally more informative about word identity than vowel tiers in French (according to calculations in Keidel et al., 2007, based on a search of *Lexique 3*; New, Pallier, Ferrand, & Matos, 2001). Given the early role of statistical learning in infancy (statistical learning plays an important role in language acquisition (Saffran, Aslin & Newport, 1996; Mersad & Nazzi, 2012, for French), French-learning infants should thus be able to discover the greater informativeness of consonants from experience with the words of their language.

Recently, Floccia, Nazzi, Delle Luche, Poltrock and Goslin (2014) proposed an intermediary "acoustic/phonetic hypothesis", whereby consonants and vowels would be processed categorically due to language-general acoustic differences (consonants often being shorter, softer, less periodic, less steady, and perceived more categorically than vowels, Repp, 1984, Pisoni, 1973). The C/V functional asymmetry would be modulated by language-specific properties of the input, thus leading to different trajectories depending on the native language. Floccia et al. (2014) have listed the following potential factors varying across languages and that could contribute to emphasize or attenuate the distinction between vowels and consonants: the realization of vowels and consonants may vary as a function of the number of vowels and consonants, the presence of vocalic or consonant reduction/lenition, phonological short/long vowel contrasts, ambisyllabicity… (i.e. consonants acting simultaneously as the coda of one syllable and the onset of the following syllable). Thus, the consonant bias would emerge as a result of acoustic/phonetic processing and phonological acquisition in the first months. Contrary to the lexical hypothesis (Keidel et al., 2007) the acoustic/phonetic origin of the consonant bias would not require very advanced lexical knowledge. Infants could distinguish consonant-sounds already from their particular distinctiveness when picking up their first consistent sound-patterns of words.

Several crosslinguistic comparisons of the role of consonants versus vowels in early lexical processing have been conducted at different ages using *word-form recognition tasks, word recognition tasks and novel word learning tasks* in order to determine infants' sensitivity to a mispronunciation. These two types of tasks involve different aspects of lexical knowledge. To learn a word, infants have to (1) extract the sound pattern of the word from the input and store it, (2) build a concept and (3) associate the word-form and its meaning. Therefore, to recognize a mispronunciation in a familiar word-form, infants only need to compare the auditory sound-pattern

to their encoded representation, while they also need to process meaning information in word recognition and novel word learning tasks.

Both word-form recognition and word recognition/learning have been evidenced as early as 6 months, for frequent words like "Mommy" and "Daddy", or "hand" and "feet" (Tincoff & Jusczyk, 1999; 2012 respectively). Moreover, 6- to 9-month-olds distinguish words referring to food and words related to body-parts (Bergelson & Swingley, 2012). The phonetic detail in these very first words has not been investigated in detail, but some data on phonetic detail in early familiar words showed that before the end of their first year, French-, English- and Dutch-learning infants are able to represent phonetic details in consonants (Hallé & Boysson-Bardies, 1996; Vihman, Nakai, de Paolis & Hallé, 2004; Swingley, 2005; Zesiger, Lozeron, Levy & Frauenfelder, 2012).

In the following, we review some of these studies and studies directly investigating the consonant bias in early lexical processing. We will distinguish language groups (French/Italian – English/German/Dutch/Danish), and different lexical tasks (word learning task, familiar word recognition, and word-form recognition), in order to try and highlight the factors influencing or not the observation of the consonant bias in early lexical processing.

I .5. a. ii. Evidence in French/Italian infants

Word learning studies

Nazzi (2005) explored the relative role of consonants vs. vowels in novel word-object pairings using the name-based categorization (NBC) task (Nazzi & Gopnik, 2001). In an interactive procedure children are presented with triads of unfamiliar objects. Two objects are labeled with the same name, the third object with a different name. At test the experimenter shows the infant one of the object while asking: "*Can you give me the object that goes with this one*?" The NBC task thus evaluates infants' ability to associate objects and names, a necessary step to word learning. After establishing that 20-month-old infants were able to distinguish and categorize objects with phonetically unrelated labels (e.g. /*pize*/ vs. /*moRa*/), further experiments contrasted minimal pairs of words differing by only a consonant or a vowel in several positions. Infants succeeded in learning stop consonant minimal pairs independently of the position (e.g. /*pize*/ vs. /*tize*/, /*pide*/ vs. /*pige*/). However they failed to learn pairs differing in a vowel contrast, both on a minimal (/*pize*/ vs. /*pyze*/; /*pize*/ vs. /*pizu*/) and on non-minimal changes (/*pize*/ vs. /*paze*/).

Nazzi, Floccia, Moquet and Butler (2009) compared the roles of consonants versus vowels in a conflicting situation, where infants were successively taught two object-name associations, e.g. /gɪb/ and /dɛb/. Then, the experimenter labeled a third target object, with a name which was a

minimal mispronunciation of both familiarized labels (e.g. /dɪb/), a consonant mispronunciation in one case (/gib/-/dib/), and a vowel mispronunciation in the other case (/dɛb/-/dib/). Similarly to the NBC task, at test, the infant is asked to give the experimenter the object that "goes with" the target object (without naming the object again). Infants' choice thus mainly depends on which pair of words sounds more similar to them than the other: if the vowel change yields more similar labels, then (/dɪb/-/dɛb/) should go together. If, on the other hand, infants rely more on vowels, they should choose the object whose label shares the same vowel with the target and differs in the consonants (/dɪb/ and /gɪb/). In this conflicting situation involving a choice between a consonantal change or a vocalic change, French-learning 30-month-olds choose to neglect the vocalic change rather than the consonantal change (Nazzi et al., 2009). In the original NBC task, 20-month-olds also showed the ability to process consonant information in detail when associating new objects to new names when changes occurred in coda and in onset position (e.g., /*bat*/ vs. /*bad*/; Nazzi & Bertoncini, 2009), suggesting that the consonant bias is independent of the position of the change. Moreover, 20-month-olds also succeeded when continuous, thus acoustically more vowel-like consonants were tested (nasals, liquids, fricatives: Nazzi & New, 2007), showing that the advantage for consonants over vowels is generalizable to more vowel-like categories of consonants.

This categorization task was then adapted to be a less demanding novel word-learning task and allow testing younger infants (Havy & Nazzi, 2009). In that adaptation, the third object was labeled with the same name as one of the previously taught label (e.g. /paʃ/-/baʃ/) but it was explicitly labeled by the experimenter at test (e.g. "*Look, this is a /paʃ/. I put this /paʃ/ in the cup. Can you put the other /paʃ/ in the cup*?"). French-learning 16-month-olds succeeded in learning pairs of words differing by an onset consonant (/paʃ/-/baʃ/) but not pairs of words differing by a vowel (/dɛr/-/dar/; Havy & Nazzi, 2009). In older children, using eye tracking, a consonant bias was found in 3-, 4- and 5-year-olds learning two novel word-object associations (Havy et al., 2014).

This consistent bias in French-learning infants was confirmed in a study using an audiovisual IPL paradigm, Jöhr (2014, chapter 5). Indeed 18-month-olds were not sensitive to a vowel mispronunciation (e.g. *mora-mura*) and reacted more to an initial consonant mispronunciation (e.g. *mora/bora*).

To sum up on word learning studies in French-learning toddlers, a consonant bias was robustly observed from the age of 16 months, and found to be independent of the consonant category tested, as well as the position of the word.

One study tested Italian-learning infants to determine the relative role of consonants and vowels in novel word learning (Hochmann, Benavides-Varela, Nespor, & Mehler, 2011). In an audio-visual novel word-learning task involving word-object pairings (***Fig. I. 3***.), infants were familiarized with an object always appearing on one side of the screen and always labeled with the same pseudowords (object *keke,* always on the right) and another labeled object appearing on the other side of the screen (another toy labeled *dudu* associated to the left). In the test phase, infants heard *kuku or dede,* and infants' anticipatory eye-movements were measured to infer whether infants relied more on the consonants (k_k_) or on the vowels (_u_u) to anticipate where the toy would appear. Twelve-month-olds infants looked more at the side predicted by *keke* than the side predicted by *dudu,* indicating that *kuku* was considered as more similar *to keke* than to *dudu*. The results show that 12-month-old Italian-learning infants consider two words sharing their consonants as more similar than two words sharing their vowels. Note that using exactly the same paradigm, Hochmann (2010) reports in his dissertation that 6-month-olds looked more at the side predicted by vowels. With the evidence of a consonant bias in 12-month-olds, the evidence of a vowel bias in 6-month-old Italian-learning infants will become relevant for our own results.

The study by Hochmann et al. (2011), however, has several methodological limitations that limit the scope of the conclusions that can be drawn from its results. First, only three-feature contrasts were tested. Thus, whether minimal consonant and vowel contrast would yield the same biases remain to be tested. Moreover, only one contrast was used for each phonological category (for consonants: /k/ - /d/; for vowels: /u/ - /e/), while previous studies have shown that the type of phonetic feature contrasted or even the specific pair of pseudowords used modulate the consonant bias expressed in toddlers and children (Havy & Nazzi, 2009; Nazzi, 2005; Nazzi & Bertoncini, 2009; Nazzi et al., 2009). Hochmann and colleagues' data should thus be confirmed on a range of other contrasts to further explore the phonetic variability present in each phonological category and allow for generalization of the results. In addition only CVCV words were used, thus making position and phonemic category confounded (consonants in initial position, vowels in word-internal position).

In novel word learning tasks, French-learning infants are thus consistently showing a consonant bias from the age of 14 months onward, and so do Italian-learning 12-month-olds in one study. Could these effects be an artifact of the task (novel word learning) and the status of the words (unknown pseudowords) used in these studies?

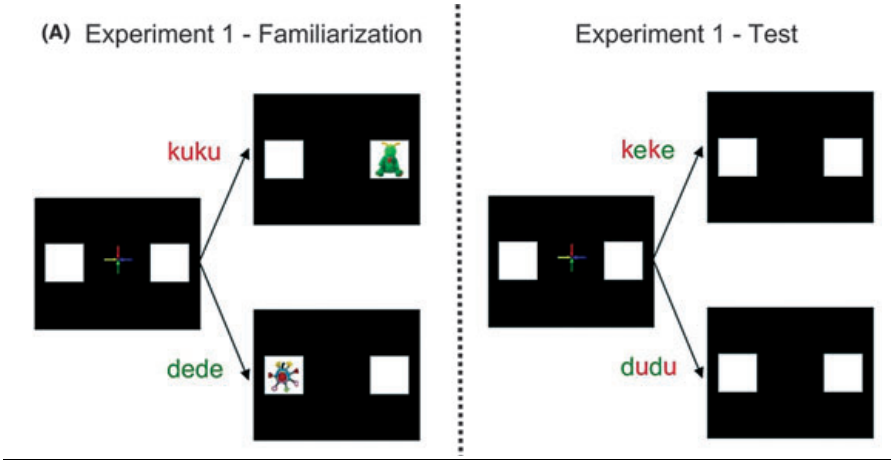


Figure I. 3. *Design of Experiment 1 investigating the consonant bias in lexical processing at 12 months. (from Hochmann et al., 2011).*

Familiar word recognition

To assess the relative role of consonants and vowel in early lexical acquisition using familiar words in French-learning infants, Zesiger and Jöhr (2011) used a label-object matching task in an Intermodal Preferential Looking paradigm (Golinkoff et al., 1987) in which infants were presented with images of two familiar objects and one of the object was named. Results at 14 months showed that infants accepted initial or final vowel mispronunciations (e.g., /*balɛ*/ for /*balõ*/), as well initial consonant mispronunciations (e.g., /*biberõ*/ for /*giberõ*/), but not final consonant mispronunciations of familiar word forms (e.g., /*bazõ*/ for /*balõ*/). These results thus reveal that French-learning 14-month-olds exhibit a consonant bias but only in the final syllable of the words. In a series of word recognition studies using the IPL, Jöhr (2014, chapters 2 to 4) investigated whether the consonant bias would be modulated by the position of the change in the word. In particular, they explored whether the fact that the final syllable is acoustically more salient in French would advantage the detection of final over initial changes in 14-month-olds (Zesiger & Jöhr, 2011), and in 18 and 22 month-olds (Jöhr, 2014). At all ages, infants failed to react to an initial or final vowel mispronunciation (e.g. *poussette* - *peussette* 'stroller' or *ballon-ballin*, 'ball') or to an initial consonant mispronunciation (e.g. *poussette* – *moussette*), but were crucially sensitive to a final consonant change (e.g. *chaussette* – *chaufette* 'sock'). Firstly, these results indicate an effect of position resulting in French-learning toddlers being better able to process the final syllable of familiar words. Secondly, they reveal a consonant bias, leading infants to show sensitivity to a consonant change and not a vowel change in the final syllable of the familiar word. To further explore this effect of position, the authors explored whether a facilitatory context would yield more sensitivity to consonant over vowel changes in the initial syllable of the words. Following White & Morgan (2008), they tried to facilitate the experimental task by presenting words in sentences

("*Trouve la chaussure/",* 'Find the shoe') and by presenting an unfamiliar object as distractor (e.g. an artichoke) instead of a familiar distractor in previous experiments (e.g. a shoe). Results showed that 18-month-olds were still not sensitive to vowel mispronunciations, but were sensitive to initial consonant mispronunciations. In addition, their consonant changes elicited graded sensitivity depending on the degree of the mispronunciation (*chaussure – sossure – fossure*; Jöhr, 2014, chapter 4)

In summary, the results of word recognition studies are confirming the evidence provided by novel word learning studies of a consonant bias in French-learning infants at the even younger age of 14 months. In the following, we will look at the data on word-form recognition studies, which have investigated French-learning infants at even younger ages.

Word form recognition studies

The first investigations of the phonetic specificity in early lexical acquisition looked at consonantal detail. Using the Headturn Preference Procedure (HPP), Hallé and de Boysson-Bardies (1994) examined French-learning 11-month-old infants' representation of phonetic detail in words. They first established that at 11 months, infants prefer to listen to a list of familiar words (*bonjour, lapin…* 'hello, rabbit' etc.) over unfamiliar rare words (*busard, cobaye…* 'harrier, guinea pig' etc.), which 9-month-olds do not. Then, Hallé and de Boysson-Bardies (1996) compared infants' looking times towards phonetically altered familiar words and intact unfamiliar words, and inferred infants' sensitivity to these various alterations in words. Their results show that infants' preference for familiar words did not vary when their initial consonant was altered (either on voicing: *ponjour, rapin…* or manner: *vonjour, napin…).* When the initial consonant was removed they did show sensitivity and no longer preferred the list of familiar words (*onjour*, *apin*…). Interestingly, they showed sensitivity to a medial consonant mispronunciation (manner: *bongour, lafin*…), suggesting a syllable final advantage consistent with Zesiger and Jöhr (2011) and Jöhr (2014). These results revealed for the first time that 11-month-old French-learning infants are reacting to alterations in familiar word forms, but as the effects were mostly found for the suppression of the initial consonant, the authors concluded that 11-month-olds had a global representation of words. To further explore these effects and compare French- and English-learning infants (the two languages having different word prosodic contours), Vihman et al. (2004) analyzed the time course of the results of Hallé and de Boysson-Bardies (1996). These reanalyzes revealed an effect of time on infants' reactions in both the initial and medial consonant change conditions, providing evidence that French-learning 11-month-olds have detailed representations of the consonants in both the

initial and final syllable of words. These studies only investigated consonant changes, what about vowel changes?

Following up on Hallé and de Boysson-Bardies (1996), Poltrock and Nazzi (in revision) compared the weight of consonants vs. vowels in a similar HPP paradigm with 11-month-old monolingual French infants. In Experiment 1, infants were presented with lists of words vs. non-words, and infants preferred real words, indicating a familiarity effect replicating Hallé and de Boysson-Bardies (1994, 1996). In Experiment 2, infants' looking times were compared for two lists of mispronounced familiar words: the consonant or the vowel of the second syllable of the familiar words used in Experiment 1 was minimally mispronounced (e.g. gâteau, /*gato*/, 'cake'; consonant change: /*gapo*/ - vowel change: /*gatø*/). Results demonstrated that infants preferred to listen to the list of vowel mispronunciations over consonant mispronunciations, demonstrating that consonants have a privileged role in 11-month-old French infants' early familiar word-forms. One might argue that reacting to a mispronunciation in a word-form is not tackling lexical processes as much as it reveals low-level acoustic mechanisms. However, the results go against the acoustic and perceptual advantage of vowels over consonants that would be expected at 11 months. Indeed, before they have attuned to their native consonant categories infants are relying more on consonants than on vowels when recognizing a familiar lexical unit. Moreover, 11 month-olds have not yet acquired a sizeable lexicon (Bates, Dale & Thai, 1995), thus, this result is contradicting the 'lexical' consonant bias hypothesis (Keidel, et al., 2007).

To conclude, most lexical studies in French -learning infants report a consonant bias in various lexical tasks, and the earliest attested evidence is yielded by French-learning 11-month-olds in a word-form recognition task (Poltrock & Nazzi, in revision) and at Italian-learning 12-month-olds in a word learning study (Hochmann, et al., 2011). We will now explore whether data in other language groups confirms the consonant bias in early lexical processing.

I .5. a. ii. Evidence in English/German/Dutch/Danish

Word learning studies

The role of vowels in English-learning 14- and 18-month-olds learning novel words was investigated in a semi-interactive task using the IPL (Mani & Plunkett, 2008). Their results showed that they are sensitive to the vowel change in newly learned words when three vocalic features are changed (e.g. *padge/poudge* and *mot/mit)*. The role of consonants in newly learned words was investigated in another IPL study (Ballem & Plunkett, 2005). They found that English-learning 14-month-olds were sensitive to consonant mispronunciations (e.g. *duke-puke* and *fope-zope*). These

results suggested that both consonants and vowels may play a crucial role in lexical learning in English-learning infants at 14 months, although these studies did not directly compare consonants and vowels.

To directly compare the roles of consonants and vowels in English-learning infants, Nazzi et al. (2009) used the conflict word-learning situation previously described for French-learning infants. As their French-learning counterparts, 30-month-old English-learning infants chose more often the target object that preserved the identity of the consonants, thus showing that they were more sensitive to the consonant than to the vowel change. This study showed a consonant bias in English-learning toddlers at 30 months.

Floccia et al. (2014) explored whether this consonant bias could be found earlier in English-learning infants. They first used the same conflict task as in Nazzi et al. (2009) with 16- and 23-month-olds English-learning infants. Contrary to French and English-learning 30-month-olds, both groups of younger English-learning infants chose the target object that preserved the identity of consonants as often as the one that preserved vowels. Again, to provide crosslinguistic data, Floccia et al. (2014) tested English-learning 24-month-olds on the word-learning task that had shown a consonant bias in French-learning 16-month-olds (Havy & Nazzi, 2009). The results revealed above-chance accuracy when learning both consonant and vowel minimal pairs of word-object associations. Altogether this study failed to observe a consonant bias in English-learning infants between 16 and 24 months, given that English-learning toddlers were equally sensitive to consonant and vowel information in word learning tasks. Overall, English-learning infants show a consonant bias in newly learned words only at 30 months (Nazzi et al., 2009), whilst other data report an equal sensitivity to consonants and vowels (Nazzi et al., 2009; Floccia et al., 2014).

Unpublished data on German-learning infants show that at 20 months, they succeed in learning two novel object-labels pairings differing minimally on their initial consonant (/def/-/gef/), thus suggesting that they are sensitive to consonant information when learning words (Schwytay & Höhle, 2014). However, in the original NBC task, they fail to categorize the third object correctly as a /def/ or /gef/, unlike French-learning 20-month-olds (Nazzi, 2005), suggesting a weaker sensitivity to consonant information. Studies comparing the use of consonants to vowels in German-learning infants learning words are missing. Whether German-learning infants are showing a consonant bias in lexical processing thus remains unknown.

Danish-learning infants were tested in the same word-learning task as in Havy & Nazzi (2009), in order to investigate the consonant bias in a language particularly rich in vowels (Højen &

Nazzi, in revision). Indeed, as previously described the phonetic structure of Danish is rich in vocoids (e.g. [*j,w*]) and long vocalic stretches. Results showed that unlike their French counterparts, Danish-learning 20-month-olds failed to learn pairs of words differing in a consonant (/*fan-san*/), even when the change involved two phonetic features *(/sun-pun/*). This suggests a weak sensitivity to consonant information in newly learned words. In addition, Danish-learning 20-month-olds succeeded in the vowel (/*dul-dyl/*) condition and their performance was significantly better in the vowel than in the consonant condition. These are the only results available in Danish -learning infants and they reveal a vowel bias at 20 months in lexical processing.

Familiar Word recognition studies

In an IPL study, English-learning 15-month-olds showed greater sensitivity to a minimal consonant change (e.g. */dog-bog/)* than to a minimal vowel change (/*dog-dig*/, Mani & Plunkett, 2007). This suggests that unlike in novel word-learning situations, English-learning infants might show an early consonant bias in recognizing familiar words. However, 18 and 24 month-olds exhibited a symmetrical processing of consonant and vowel mispronunciations in the same study (Mani & Plunkett, 2007) and so did 12 month-olds in a later study (Mani & Plunkett, 2010). Thus, evidence in English-learning toddlers is more in favor of a symmetric role for vowels and consonants when processing familiar words. Future studies will need to further explore the contradictory result at 15 months, which may be an artifact of the task. Thus, unlike French-learning infants, English-learning infants do not seem to preferentially use consonant over vowel information when recognizing familiar word-object associations. Could word-form recognition tasks be a more appropriate task to reveal a consonant bias in English-learning infants?

Word form recognition

To our knowledge no study directly compared consonant and vowel mispronunciation effects on familiar word-form recognition neither in English-, Dutch-, German- or Danish-learning infants. However some results do provide indications regarding English- and Dutch-learning infants' ability to process each of the speech categories in such tasks.

To examine Dutch-learning 11-month-old infants' sensitivity to onset and coda consonant mispronunciations in familiar monosyllabic word-forms, Swingley (2005) conducted a series of HPP experiments. This study first replicated in Dutch-learning infants the preference for familiar words over non-words originally found in French-learning infants (Hallé & de Boysson-Bardies, 1994).

Follow-up experiments thus showed that this preference disappeared when the initial consonant or coda consonant were mispronounced, although the effect was less strong for coda than for onset consonants (e.g. /*be:n*/, leg, changed to /*de:n*/ or /*be:m*). This study establishes that Dutch-learning 11-month-olds are sensitive to consonant mispronunciations in familiar words.

Vihman et al. (2004) used the same paradigm as Hallé and de Boysson-Bardies (1996) with English-learning infants at 9 and 11 months. They first found that, as for French-learning infants, a word familiarity preference is present in 11-month-olds, but not in 9-month-olds. Moreover, together with a reanalysis of the results of Hallé and de Boysson-Bardies (1996), the time-course of infants' reactions in the first vs. last half of the trials showed that the initial consonant was more important than the medial consonant for English-learning infants. Recall that for French-learning infants, the opposite pattern of results was observed, medial consonants being more important than initial consonants. Given that most English words have a trochaic stress pattern (i.e. strong-weak, Cutler & Carter, 1987), while French exhibits phrase-final lengthening (Delattre, 1965), these opposite patterns of results in the two language groups were interpreted as an effect of acoustic salience of the syllable. Indeed each language group shows higher detail of phonetic detail in the most salient syllable.

Vihman et al. (2004) did not test the same manipulation with vowels. However, in one condition they presented English-learning infants with a list of unfamiliar words over familiar words with reversed stress patterns. They found that infants still preferred the familiar words. The authors argued that reversing the stress pattern in an English word may impact vowels more than consonants due to the correlation between lexical stress and vowel quality (Cutler et al., 1986). Thus, this result indirectly suggests a lack of sensitivity to vowel manipulation in familiar words in 11-month-old English-learning infants (although the time-course analyses reveal some transient sensitivity), unlike what was found for consonant manipulations. However, the weight of consonants and vowels in 11-month-old English-learning infants' lexical representations were not directly compared in this study, thus this interpretation should be taken cautiously.

To conclude on lexical studies looking at the C/V functional asymmetry in various languages, important cross-linguistic differences were observed. In French-learning infants a consonant processing bias is consistently and robustly found from the age of 14 months (Zesiger & Jöhr, 2012; Havy & Nazzi, 2009) and even at 11 months in reaction to mispronunciations in familiar word-forms (Poltrock & Nazzi, in revision). In Italian, a language typologically, lexically and rhythmically close to French, 12-month-olds also revealed a consonant bias in one experiment including novel word learning (Hochmann et al., 2011). English-learning infants are quite consistent in showing a

symmetrical use of consonants and vowels in lexical processing from 12 months onwards, except at 30 months (and possibly after?) when they use consonant information preferentially over vowels. The emergence of the consonant bias in English-learning infants thus seems to be quite late compared to French and Italian-learning infants. Lastly, Danish-learning infants tested at 20 months revealed a strong vowel bias, interestingly in the same novel word-learning task that yielded a consonant bias in French (Højen & Nazzi, in revision).

At present, this pattern of results remains too patchy to fully understand the language-specific trajectories of the C/V functional asymmetry in lexical processing. However, the most informative cross-linguistic comparison comes from the now substantial although incomplete body of literature from French- and English-learning infants. French-learning infants are better at processing lexical information provided by consonants than by vowels, although they do not always fail to use vowels. They do however consistently demonstrate a consonant bias from 11 months onwards. Unlike them, English-learning infants seem equally able to process consonants and vowels in words from the age of 12 months, and except for one result at 15 months, they only show a clear preference for consonant information over vowel information at 30 months of age (Nazzi et al., 2009). Moreover the results of Zesiger and Jöhr (2011) at 14, 18 and 22 months (Jöhr, 2014) using the same IPL task as Mani and Plunkett (2007, 2010) suggest that French-learning infants' preferential use of consonant over vowel information is independent of the task requirements. Thus the results from Nazzi and colleagues are less likely to be linked to the difficulty of the NBC task and its variants, as initially argued by Mani and Plunkett (2007). These results suggest that language-specific consonant and vowel properties play a crucial in the C/V functional asymmetry in lexical processing. However, these results are found in infants who already received substantial exposure to their native language. To test the developmental validity of the universal and innate consonant bias proposed by Nespor et al. (2003), a direct investigation of the development of consonant and vowel processing in younger infants is necessary.

RATIONALE for investigating the Consonant bias at 5 months

The results from French- and Danish-learning infants suggest that at least in the second year of life infants are exploiting the most informative category of speech sounds in their native language to learn words. However, it remains unknown whether this ability emerged from acoustic/phonetic distinctions or their relative importance in the infant lexicon. Evidence of a consonant bias at 11 months (Poltrock & Nazzi, in revision) seems to suggest that infants do not need to have a large lexicon to process consonants preferentially, at least in familiar word-forms. Given that lexical

acquisition starts around 6 months (Bergelson & Swinlgey, 2012; Tincoff & Jusczyk, 1999, 2012) and even at 4.5 months for the infant's own name (Mandel et al., 1995), the studies of Chapter 3 of this thesis will investigate the role of consonants versus vowels in lexical processing at the onset of lexical acquisition in a word-form recognition task. In particular, we will compare French-learning infants' reactions to a mispronunciation in their own name as a function of whether the mispronunciation is a consonant or a vowel change.

I .5. a. iv. Item memorization as a lexically-based process

Before any lexical entry has been built, infants are able to represent word-forms. Indeed neonates are not only able to discriminate phonetic contrasts (Eimas et al., 1971) and extract prosodic information (Mehler et al., 1988) from speech as previously mentioned, but also recognize word-forms after familiarization. For instance, in head-turning experiments newborns recognize individual sound patterns of words after a short retention period of a 100s (*"beagle" or "tinder"*, Valiante, Barr, Zelazo, Papageorgiou & Young, 2006), and even show better results for frequent vs. infrequent words (baby vs. beagle, respectively, Valiante, Barr, Zelazo, Brant, & Young, 2013). Such word-form memorization can be observed even after a 24-hour delay (Swain et al., 1993). In a neuroimaging Near Infrared Spectroscopy (fNIRS) study measuring the newborn brain's hemodynamic activity using a familiarization-recognition design, memory for a bisyllabic word was observed after a 2-min pause (Benavides-Varela, Gomez, Macagno, Bion, Peretz & Mehler, 2011). Thus, newborns are able to represent and memorize consistent strings of phonemes, which may constitute a precursor of lexical-like speech processes.

To evaluate the relative contribution of consonants vs. vowels in word-form memorization at birth, Benavides-Varela, Hochmann, Macagno, Nespor and Mehler (2012) conducted another fNIRS study with a similar design as in Benavides-Varela et al. (2011). Triplets of bisyllabic CVCV-words matching for duration (700ms) and intensity (70dB), were recorded by a female Italian speaker: *sisi, mimi, mama*; *lili, titi, lala.* All newborns were familiarized for 6 minutes with the same word (e.g. *mimi*). In the test phase, half of the participants heard a word with a change in the consonant tier, but not in the vowel tier (*mimi→sisi*) and the other half heard a word that preserved the consonant tier but had a different vowel tier (*mimi→mama*). The oxyhemoglobin concentration increased for the test item sharing the *consonant* tier with the familiarization item (cf. black bar in ***Fig. I. 4***. *mimi→mama*). Such an increase in brain activity between familiarization and test corresponds to a characteristic hemodynamic response to novel words. Moreover, for the test item sharing the *vowel*

tier the oxyhemoglobin concentration *decreased* (cf. white bars in ***Fig. I. 4.*** *mimi→sisi*). This differential pattern of results was found in right frontal, left temporal, and right parietal areas.

The above findings suggest that the newborn brain responded more to a change in the vowel tier of the word than to a change in the consonant tier when memorizing words at birth. These results suggest that unlike older infants and adults, newborns process mainly the vowels and/or the information carried by vowels when they hear words. However, the stimuli were natural recordings, thus did not control for acoustic saliency. In addition the speaker was instructed to record all bisyllabic words with initial syllable stress, which is likely to create pitch contour variations thus favoring vowel over consonant perception (Nespor & Vogel, 1986), and which newborns are sensitive to (Nazzi, Floccia & Bertoncini, 1998). The interpretation would benefit from more acoustically controlled stimuli by avoiding prosodic variations and minimizing acoustic differences between consonants and vowels. Nevertheless, this vowel bias is congruent with previous behavioral evidence that at birth infants react more to a vowel (‘*bi si li mi → bi si li ma’ or ‘bi be bo ba → bi be bo bu’*) than to a consonant change (‘*bi si li mi → bi si li di’ or ‘bi be bo ba → bi be bo da’*) in a syllable sequence, found using a non-nutritive high amplitude sucking procedure (HAS, Bertoncini, Bijeljac-Babic, Jusczyk & Kennedy, 1988).

Importantly, Benavides-Varela et al. (2012) demonstrated that the fNIRS technique is sensitive enough to reveal differential responses to a consonant vs. a vowel change in an item memorization paradigm at birth. Thus, fNIRS is a suitable technique for the investigation of the consonant bias in a lexically-related type of speech processing in the newborn brain.

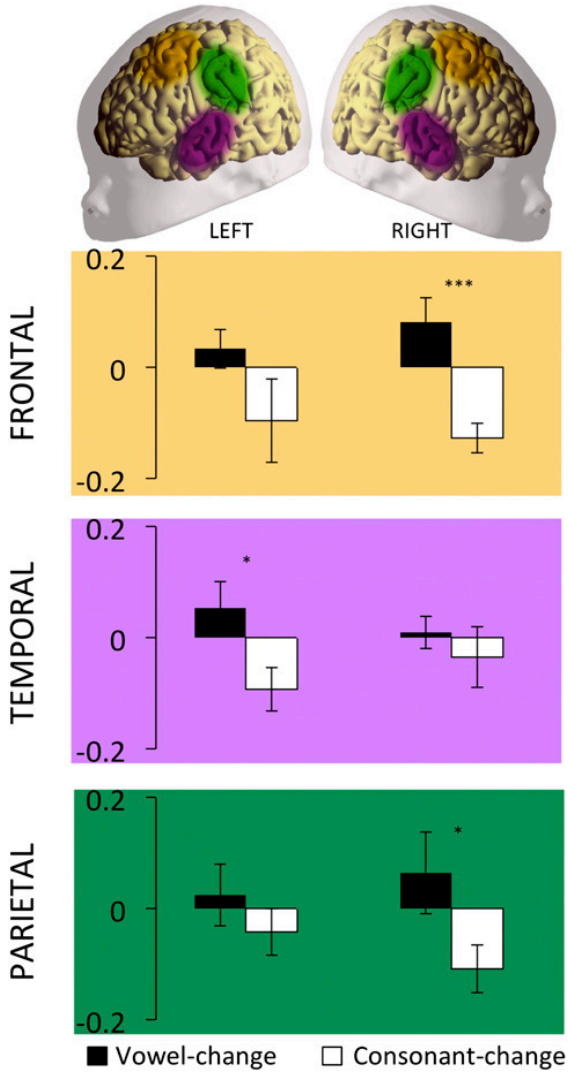


Figure I. 4. *Oxyhemoglobin concentration changes from the last block of familiarization to the first block of the test phase, for a consonant change (white bar, mimi→sisi) and for a vowel change (black bar, mimi→ mama). Colored ellipses on the schematic neonate brain indicate the localization of the channels included in the ROI. Error bars indicate SEM (t test, ***P < 0.0001; *P < 0.05) (from Benavides-Varela et al., 2012).*

RATIONALE for investigating the consonant bias at birth in item memorization

Repetitions of syllables are highly salient and preferred in young infants' input (in *infant directed speech,* Fernald, 1985; Fernald & Kuhl, 1987) and are also frequent in their early vocabularies (Ferguson, 1977). Moreover, in several NIRS studies, (Gervain, Macagno, Cogoi, Pena, & Mehler, 2008; Gervain, Berent, & Werker, 2012) found that sequences carrying syllable repetitions (e.g. *mubaba*) are automatically detected by the newborn brain and learned more easily than non-repetitive control sequences (e.g. *mubage*).

Some of the experiments reported in Chapter 2 aim at investigating whether lexical items carrying repetitions are more likely to be detected and learned in the newborn brain in a fNIRS paradigm, if the repetition is carried by consonants than by vowels, as predicted by the initial bias

hypothesis Nespor et al. (2003), or whether a different pattern of results obtains, e.g. a vowel advantage, as predicted by the other hypotheses.

I. 5. b. The vowel bias in morphosyntax

I. 5. b. i. Rule-learning as a precursor of syntax

Rule learning is a key feature of language development because language is governed by a set of grammatical rules, i.e. structural patterns organizing words into sentences (Chomsky, 1957), that allow the generation of an infinite number of sentences from a finite number of elements (i.e. morphemes). To learn the grammar of their native language from sentences composed of different words, infants must thus extract the underlying patterns the sentences share, and generalize them to different words. This regularity extraction and generalization mechanism has been observed in the first year of life (Gomez & Gerken, 1999; Marcus, Vijayan, Rao, & Vishton, 1999), and even at birth in the case of adjacent repetitions (Gervain et al., 2008; 2012). This mechanism is considered one of the earliest precursors of grammatical acquisition, and could thus serve to investigate the relative role of consonants vs. vowels in a syntax-like process at birth.

Marcus et al. (1999) were the first to show that abstract rule learning abilities are already present at 7 months. In their paradigm, infants were first familiarized with CVCVCV sentences, implementing an ABB (*gatiti, linana*) or an ABA rule (*gatiga, linali*). Then they were tested on their preference for novel test sentences, half of them implementing the familiarized rule, and the other half implementing the other rule for which they had not been trained (*wofewo* or *wofefe*). Their longer looking times to the unfamiliarized grammar during test suggested that the repetition pattern had been extracted and was generalized to novel CVCVCV words. Importantly, the only thing that familiarization and test items had in common was their abstract structure, but not their syllables, allowing this study to be interpreted as the first evidence of elementary syntactic rule learning and generalization, not simple memorization or statistical information extraction.

Gervain et al. (2008) tested the presence of repetition rule learning at birth, using NIRS. Neonates were presented with synthesized monotonous CVCVCV speech sequences implementing a repetitive pattern (ABB in Experiment 1, or ABA in Experiment 2) versus a control sequence ABC, to directly compare each repetitive grammar to a random grammar. In Experiment 1, an increased response in temporal and left frontal areas to ABB over ABC was found (e.g. *mubaba vs. mubage*, respectively): Time-course analyses showed that this initial detection of ABB over ABC occurred in

the first 4 blocks in bilateral temporal areas, and increased over the time course of the experiment in the left frontal area for the ABB grammar. This was interpreted as a possible two-step process, involving a first automatic and rapid response to the ABB grammar with a bilateral temporal location of this initial perceptual response; followed by a higher-level response occurring in a frontal left area (known for its role in higher order learning) suggesting a more abstract process of regularity extraction. On the contrary in Experiment 2, ABA and ABC grammars elicited similar canonical hemodynamic responses in temporal bilateral areas, suggesting that the non-adjacent repetition grammar is not detected over an ABC grammar. Thus ABA sequences do not activate automatic low-level mechanisms, suggesting that at birth perception is not specifically sensitive to distant repetitions, preventing the formation of a specific abstract representation of the ABA grammar. Gervain et al. (2012) replicated the initial detection results for AAB over ABC grammars, interestingly without replicating the time-course effect.

In conclusion of this series of fNIRS studies at birth, Gervain and colleagues postulated two mechanisms of repetition-rule-learning, having two different processing mechanisms: first, a rapid automatic low-level perceptual detection mechanism, present at birth for the adjacent repetition rule, as indicated by the stronger bilateral temporal activation for ABB (or AAB) over ABC in newborns (Gervain et al., 2008, 2012). Secondly, a symbolic representation of the rule is built for ABB and AAB grammars from birth (Gervain et al., 2008, 2012), but needs to develop to represent ABA. Newborns may have difficulties processing ABA repetitions due to a limited memory time-window for repetition detection and/or due to the interfering effect of the item B in ABA sequences.

To investigate the relative role of consonants and vowels at birth in structural learning, repetition-rule learning thus appears like a good candidate given its morphosyntactic importance and the early evidence of a repetition-based regularity extraction mechanism in infants.

I. 2. b. ii. Rule-learning based on vowels at 11-12 months

Only two studies specifically investigated the C/V functional asymmetry in extracting rules in infancy. In the previously described audiovisual paradigm, Hochmann et al. (2011) directly investigated 12-month-old Italian-learning infants' ability to extract and generalize a repetition-based regularity from a list of different CVCV words, and compare their performance whether the repetition was implemented on consonants (i.e. *lula, lalo, dado, dodu, fufa and fofu*) or on vowels (i.e. *dala, dolo, fodo, fudu, lafa and lufu*). Each of six consonant-repetition words was followed by a toy' appearance on one side of the screen, whereas each of six vowel-repetition words was followed by the appearance of a toy on the other side of the screen. At test, infants heard two novel words

with the consonant repetition regularity (i.e. *kike and memi*) or the vowel repetition regularity (i.e. *meke and kimi, **Fig. I. 5***.). Infants looked more at the side predicted by the regularity in the vowel repetition condition, but looked equally on both sides in the consonant repetition condition, thus indicating that they generalized the rule implemented on vowels, and not the one on consonants.

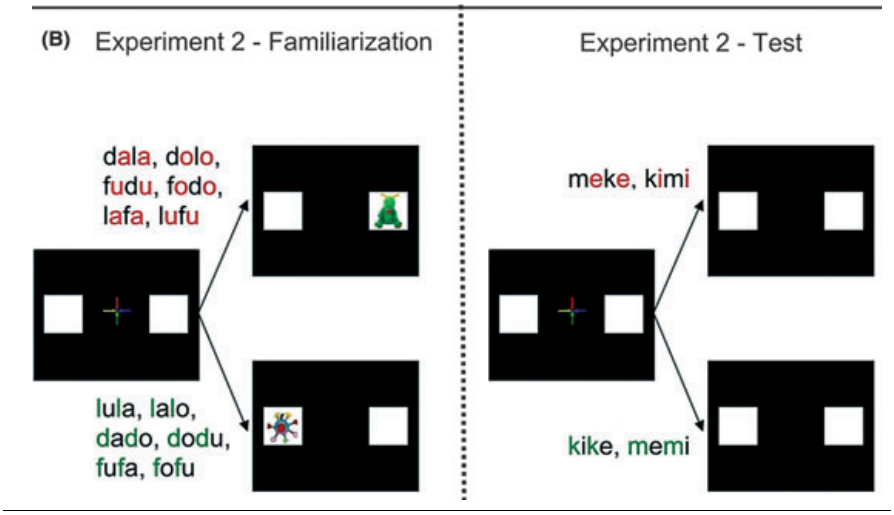


Figure I. 5. *Design of Experiment 2 investigating the vowel bias in rule-learning at 12 months (from Hochmann et al., 2011).*

In a preferential looking time procedure, Pons and Toro (2010) familiarized 11-month-olds with a list of CVCVCV sequences carrying an AAB repetition structure. In Experiment 1, this structure was implemented on the vowel tier (*dabale, tibilo*…) and they were then tested on their preference towards novel test sequences implementing the *same* AAB structure on vowels (*nadato, babalo…)* or a different ABC structure (*dutone, lanude*…). In Experiment 2, familiarization sequences implemented an AAB structure on the consonant tier (*dadeno, lulabo…),* and novel test sequences either implemented the *same* AAB structure on consonants (*titedo, loleni…)* There results show a preference for rule over non-rule words in Experiment 1 but not in Experiment 2 (see ***Fig. I. 6***.), indicating that generalizations of repetition-based regularities are easier over vowels than consonants.

To conclude both studies on the vowel bias in syntactic processing show that by the end of the first year, vowels appear to be a privileged category for extracting and generalizing abstract structures, in line with the prediction of Nespor et al. (2003).

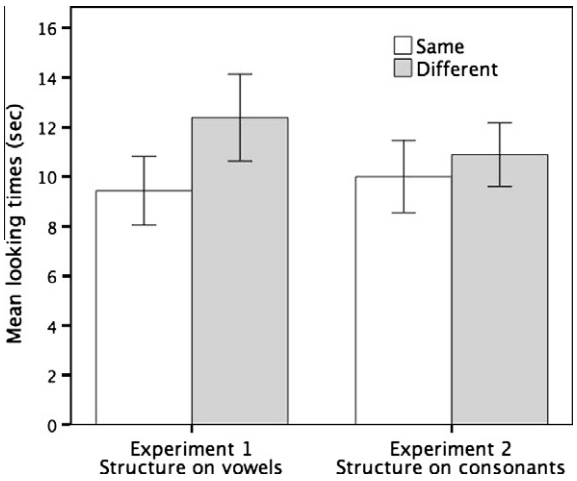


Figure I. 6. *Mean looking times (in seconds) and SEs for Same and Different test trials when the AAB structure was implemented over the vowels (Experiment 1) or over the consonants (Experiment 2, Pons and Toro, 2010)*

RATIONALE for investigating the vowel bias in rule-learning at birth

It is well established that newborns detect and learn repetitive structures based on syllables from implemented on various CVCVCV sequences using fNIRS (Gervain et al., 2008, 2012). Whether the preferential role of vowels in early syntax-related processing is already present at birth as predicted by Nespor et al. (2003) remains unknown and will constitute the second complementary aim of the experiments in Chapter 1 on the origins of the C/V functional asymmetry at birth.

EXPERIMENTAL CHAPTER 2 – Investigation of the Consonant/Vowel asymmetry in speech processing at birth, using Functional Near-Infrared Spectroscopy

The functional asymmetry between consonants and vowels potentially offers a helpful learning mechanism for infants to break into language. For this bootstrapping effect, it is important that the asymmetry be present from an early age, possibly from birth or even prenatally. This chapter therefore investigates the very origins of this bias, by evaluating how repetition-based regularity detection, originally found for *syllable*-repetition structures (ABC vs. ABB, Gervain et al., 2008), generalizes to *segment*-repetition structures, and whether consonants and vowels are already functionally differentiated at birth as predicted by the initial division of labor hypothesis (Nespor et al., 2003). Three experiments were conducted with newborn participants using functional Near Infrared Spectroscopy (fNIRS). Experiment 1 is a replication of the original syllable-repetition paradigm (Gervain et al. 2008), except that it uses and validates a new material (a new repertoire of consonants and vowels) that was necessary to create a more balanced stimulus set allowing a better comparison between consonants and vowels. Experiments 2 and 3 are two complementary experiments that directly address the question of the origins of the C/V asymmetry. These two complementary experiments share the same grammars and stimuli, i.e. consonant repetitions (ABBc), vowel repetitions (ABBv) and unstructured sequences (ABC), but differ crucially in their experimental designs. Indeed the context of presentation of the stimuli is manipulated in order to explore the relative roles of consonants and vowels. In particular, Experiment 2 consists in presenting the stimuli 'by type', thus inducing rule-learning, predicting a stronger reliance on vowels, while in Experiment 3 infants are exposed to stimuli "by token", thus eliciting item/lexical memorization, predicting a stronger reliance on consonants.

We used fNIRS to investigate newborns' brain responses in all three experiments. Thus, before discussing the experiments themselves, the following section presents the general principles and properties of the fNIRS technique.

II. 1. Introduction to fNIRS

II. 1. a. Brief history and general properties of fNIRS

NIRS (or optical imaging) is a non-invasive neuroimaging technique that exploits the principles of near-infrared spectroscopy (***Fig. II. 1***), which is based on the properties that human tissues are relatively transparent to the Near Infrared (NIR) spectral range of light, i.e. wavelengths beyond 650nm (***Fig. II. 2.,*** Jöbsis, 1977).

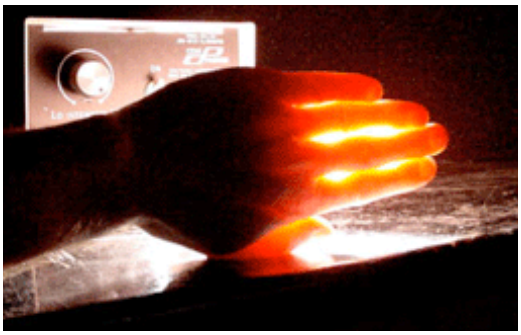


Figure II. 1. *Photons can be absorbed by the pigmented compounds (chromophores), scattered in tissues, or transmitted, i.e. they can continue their trajectory unaffected by the medium. Thus, light penetrates up to several centimeters in biological tissues. Here a bright white source transmitted through a hand and causing a reddish glow due to the absorption of photons by the pigmented hemoglobin protein in red blood cells (picture from www.nirx.net).*

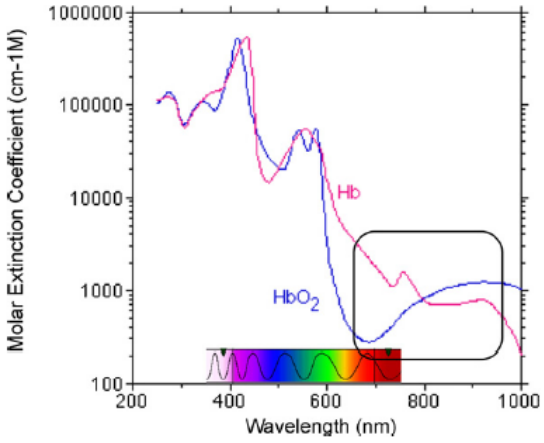


Figure II. 2. *The molar extinction coefficients of oxyHb (blue curve, 'HbO₂') and deoxyHb (pink curve, 'Hb') as a function of light wavelengths. Biological tissues (e.g. skin, bone) are transparent to light in the spectral wavelengths covered by the black square, which are relevant for fNIRS, except for the isosbestic point where extinction coefficients of oxy and deoxyHb are equal (from Gervain et al., 2011).*

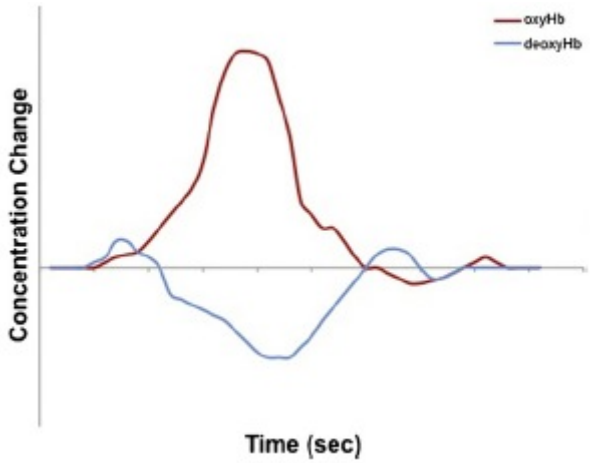


Figure II. 3. *A typical hemodynamic response function (HRF) accompanying an increase in brain activity evoked by a stimulation (from Gervain et al., 2011).*

The relatively high attenuation of light in parts of the spectrum other than the NIR range in tissue is due to absorption of photons by water as well as the main chromophore, hemoglobin (Hb), the protein that transports oxygen in blood cells and is either oxygenated (oxyHb) or deoxygenated (deoxyHb). In the NIR range, absorption by water and Hb is lower, thus tissues are more transparent.

By measuring the absorbance of NIR light, optical imaging allows for the detection of changes in the optical properties of brain tissue related to changes in blood oxygenation, from which neural activity can be inferred. Indeed, NIRS measures the hemodynamic correlates of brain activity and thus relies on neurovascular coupling, i.e. the principle that typically links focal neuronal activity to a local increase in cerebral blood flow (that supplies oxygen and glucose to neurons). The canonical neurovascular response to stimulation consists of an increase in oxyHb and a decrease in deoxyHb concentrations (***Fig II.3***).

For 20 years now, fNIRS has provided an indirect measure of cortical activity reflected in oxyHb and deoxyHb concentration changes, in response to a stimulus (Chance et al., 1993; Hoshi and Tamura, 1993; Kato et al., 1993; Villringer et al., 1993). The properties of the fNIRS technique have made it particularly suitable for infant brain imaging investigations, first in clinical settings, then in basic research, first using simple sensory stimulations

(acoustic tones: Sakatani, et al., 1999; odors: Bartocci et al., 2000; stroboscopic flashing light and other classic visual stimuli: Kusaka et al., 2004) and then more complex stimuli (multisensory, audio-visual: Bortfeld, Wruck & Boas, 2007; Bortfeld, Fava & Boas, 2009; Fava, Hull & Bortfeld, 2014; Taga & Asakawa, 2007; social communication, biological motion processing and face processing: Lloyd-Fox et al., 2009; Otsuka et al., 2007; Southgate et al., 2014). Research on the development of language in the brain has particularly benefited from fNIRS and the methodological innovations developed in the last two decades (for reviews: Gervain et al., 2011; Minagawa-Kawai et al., 2008; Quaresima et al., 2012; Rossi et al., 2012; Dieler et al., 2012; Homae, 2014).

Most NIRS systems use continuous-wave light (CW), i.e. constant illumination of tissue with monochromatic NIR light or pulsated narrow-bandwidth LED light, and measure changes in NIR light intensity at different wavelengths, permitting estimation of cerebral hemoglobin concentration *changes*, thus *relative* oxyHb and deoxyHb concentrations. More costly and technologically complex NIRS illumination techniques allow for the absolute optical properties of tissues and thus absolute concentrations to be obtained (for reviews see: Frequency-Domain and Time-Domain systems: Lloyd-Fox et al., 2010; Wolf et al., 2007; Torricelli et al., 2008; for a comparison of CW, Frequency-Domain and Time-Domain NIRS systems: Scholkmann et al., 2014). The explanations below are restricted to CW/pulsated relative systems, which are the simplest and most affordable NIRS devices, and provide sufficient measures for research-related applications of fNIRS.

II. 1. b. Principles of fNIRS

The Beer–Lambert law (1) is the equation that allows for the calculation of the concentration of an idealized, homogeneous, non-scattering medium that absorbs light. It establishes the relationship between the proportion of light absorbed, the properties of the medium and the intensity of the incident light. Light attenuation (*A,* the loss of light intensity) is calculated from: the known properties of incident light (before tissue penetration: intensity I_0, wavelength λ), the same properties of NIR light measured after penetration through the tissues, and the estimated optical properties of the tissues (*c* density or concentration of the medium, ε_λ molar extinction coefficient of the medium, *l* the distance travelled by light in the medium).

$$A = -\log\left(\frac{I}{I_0}\right) = c \times \varepsilon_\lambda \times l \qquad (1)$$

An adapted version of the Beer-Lambert law (2) yields a more accurate measure of *A* in highly scattering, non idealized media, such as biological tissues (Delpy et al., 1988) by taking into account the specific scattering properties of the various media that the light is penetrating (skin, skull and brain, *G* the scatter) and thus its non-linear trajectory (using the *DPF*, an estimated coefficient factor accounting for the non-linear trajectory of light in biological, differential pathlength, and dependent on source-detector distance)

$$A = -\log\left(\frac{I}{I_0}\right) = (c \times \varepsilon_\lambda \times l \times DPF) + G \qquad (2)$$

Given that two wavelengths of NIR light are used, λ_{oxy} *and* λ_{deoxy} (Fig II.2), the resulting two-equation system yields the following equation (3). ΔA is measured, and the constant scatter coefficient G, assumed to remain constant in time in a single participant, can thus be eliminated. Thus the unknown changes in concentration can be calculated: Δc_{oxy} relative concentrations of oxyHb and Δc_{deoxy} deoxyHb respectively:

$$\Delta A = \left[\left(\Delta c_{oxy} \times \varepsilon_{\lambda oxy}\right) + \left(\Delta c_{deoxy} \times \varepsilon_{\lambda deoxy}\right) \right] \times l \times DPF \qquad (3)$$

The use of multiple-channel NIRS systems in the early 90's led to two different NIRS techniques: optical topography and optical tomography. The latter allows a three-dimensional reconstruction of the NIRS signals of the cortex. It requires long data acquisition and image reconstruction time, is costly and has a low spatial resolution; it is therefore mostly used in clinical applications only and much less in basic research (Hebden et al., 2002, Hebden et al., 2004).

The most common multiple-channels NIRS system is the optical topography (OT) system. It pairs sources and detectors (optodes) to form measurement channels at the surface of the scalp and provides a two-dimensional sampling of the cortical surface. The source–detector distance (S-D distance) determines the depth of penetration, and the spatial resolution. The optical pathlength from source to detector, i.e. length and spatial distribution of NIR light through tissue layers, is longer than the actual S-D distance, and its shape is modulated by the scattering properties and optical heterogeneity of the various tissue layers, the intensity and wavelengths of the light, the age of the subjects and the head region measured, which determine the thickness of the layers. The portion of tissue penetrated by light from the source to the detector has a banana-shaped trajectory (***Fig. II. 4***.) and defines the cortical areas reached by a given channel. The larger the S-D distance, the deeper the brain structures explored. But placing sources and detectors too far apart reduces the signal-to-noise ratio and the spatial resolution, and the number of optodes that can be placed on the head. In newborns and young infants, the optimal S-D distance providing a satisfactory trade-off between depth of penetration and spatial resolution, which yielded replicable results, is between 2.5 and 5 cm (Gervain et al., 2011).

NIR light is generally carried by flexible optic fibers to and from tissues (***Fig. II. 5***.). The different types of light sources can either be laser diodes and/or light emitting diode (LED) sources. Each detector records the amount of light coming from the coupled sources, and different channel configurations can be arranged depending on the research question and age of subjects. With newborns optical fibers are arranged into probes or caps with different channel configurations, typically using 2.5-4 cm source–detector separations.

Data analysis

Typical pre-processing of NIRS data converts raw optical measures to oxyHb and deoxyHb concentration changes using the modified version of the Beer-Lambert law, and improves the data quality through a series of steps (Huppert et al., 2009). First the data are filtered to remove low-frequency oscillations originating from slower changes in systemic cardiovascular properties, e.g. blood pressure and high frequency noise such as instrumental or other noise as well as signals not related to stimulation (heart beat, breathing, sucking on a pacifier etc.). Then motion artifact rejection detects unreasonably large changes in concentration within trials or blocks. Such pieces of data are rejected or replaced using interpolation techniques.

After pre-processing, several statistical analysis methods are available. The analysis method used in this thesis is time series averaging, typically used in NIRS infant studies (e.g. Gervain et al. 2008; Otsuka et al., 2007; Peña et al., 2003; Shibata et al., 2007). However, others use model-based fNIRS analysis inspired by fMRI research (e.g. Shimada & Hiraki, 2006; Telkemeyer et al., 2009;

Wartenburger et al., 2007), and correlate NIRS data to a predictor function of the typical adult HRF and involve general linear model (GLM) and statistical parametric mapping (SPM) approaches.

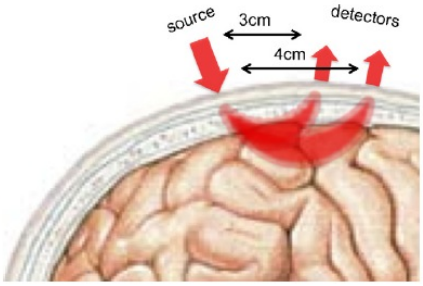


Figure II. 4. *The optical pathlength from source to detector at two different Source-Detector distances in optical topography (from Gervain et al., 2011).*

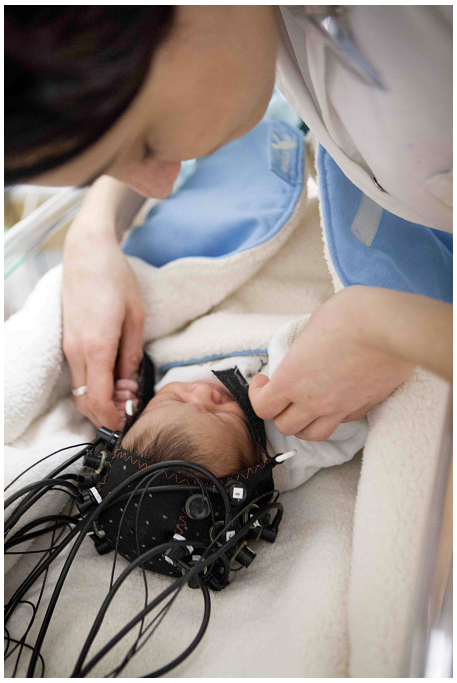


Figure II. 5. *Picture of a newborn participant in his crib, before an experimental session in the current studies (NIRSCOUT NIRx). Optical fibers are conducting NIR light from the NIRS system to led sources inserted in the cap placed upon the head of the participant. Optode configuration of the 24 channels used in the current fNIRS experiments.*

II. 1. c. Main advantages and limitations of NIRS for the study of speech processing in infancy

NIRS presents a number of practical advantages over other techniques, especially for studying the functional brain organization and the dynamics of speech processing in development. However, some limitations also have to be taken into account when designing a NIRS experiment to answer a specific research question, and when analyzing and interpreting the data (for a more general review of NIRS advantages and limitations, see Ferrari, Mottola, & Quaresima, 2004; Aslin, 2012).

I. 1. b. i. Advantages

Unlike MRI, PET, and other nuclear imaging techniques, NIRS is completely safe and non-invasive, thus more suited to research purposes especially in young infants. Moreover younger participants have thinner tissue layers and lighter hair and skin, thus allowing for light to travel longer distances with less attenuation before reaching the cortical layers than in older participants. Thus younger populations typically yield a better NIRS signal-to-noise ratio.

Unlike fMRI, NIRS is completely silent, which is crucial for easily presenting young infants with auditory stimuli such as speech. Moreover NIRS is much less sensitive to movements than MRI, EEG and MEG. Therefore the infant does not need to be physically constrained to a rigid position, reducing the risk of infants awakening and/or becoming fussy.

As fMRI, NIRS measures hemodynamic correlates of brain activity and there are strong correlations between the NIRS response and the BOLD signal measured in fMRI, which is based only on concentration changes in deoxyHb (Alderliesten et al. 2014). NIRS measures oxyHb besides deoxyHb, thus yields a more complex evaluation of the cortical hemodynamic response than fMRI. A decrease in deoxyHb is the major source of the BOLD contrast (Kleinschmidt et al., 1996; Obrig & Villringer, 2003), therefore areas in which fNIRS shows a decrease in deoxyHb should correspond to areas showing a BOLD response. Therefore some NIRS studies primarily use the deoxyHb signal as their main measure (Obrig et al., 2000; Telkemeyer et al., 2009; Rossi et al., 2011; 2012). However, the oxyHb signal more often yields significant results, especially in newborns and young infants, most probably due to technical properties of the commercially available NIRS systems, such as the material of the optical fibers etc. (Sato et al., 2012; Gervain et al., 2008). Additionally some studies with infants also show a decrease in oxyHb upon stimulation (Obrig et al., 2010) instead of the canonical increase in oxyHb (***Fig. II. 3***.). Several explanations have been proposed, such as young infants' immature and more complicated neurovascular regulation (different metabolic oxygen rates, non-myelinated axons…), or the reallocation of cognitive resources having particular blood

flow redistribution consequences ('stealing/plumbing effect') during stimulus processing (Quaresima et al., 2012). The origins of negative activation and the neurovascular coupling specificities in infants are not well understood, thus it appears more adequate to report both the oxyHb and the deoxyHb signals (Rossi et al., 2012). NIRS undoubtedly offers a more complete measure of the cortical hemodynamic response than fMRI, which takes on a special importance for infant research.

In terms of temporal resolution the data acquisition sampling rate is usually around 10 Hz and can go up to 100 Hz, allowing to measure the time course of cortical hemodynamic changes in response to certain stimuli (Gervain et al., 2008) with better temporal resolution than fMRI. Moreover, unlike EEG, fNIRS provides precise spatial localization, as it is not subject to the inverse problem of source localization. Its spatial resolution, however, is in the centimeter range, i.e. worse than that of MRI.

Lastly, contrary to MRI measurements, the NIRS signal does not interfere with electrophysiological measures. Thus, combined EEG and fNIRS measurements have the advantage of providing excellent temporal resolution and better spatial localization (Rossi et al., 2011; Wallois, Mahmoudzadeh, Patil, & Grebe, 2012).

I.1.b.ii. Limitations

Although better than in fMRI, the temporal resolution of fNIRS makes it unsuitable for the exploration of very rapid neural reactions, which are often crucial for research questions related to speech processing. Electro- and magneto-physiological measures with temporal resolutions ranging in milliseconds should then be preferred.

The main disadvantage of fNIRS lies in its relatively shallow penetration into brain tissue, allowing only surface structures to be seen. Thus, NIRS is not an appropriate measure for targeting cognitive functions involving structures lying deep in the brain (memory and emotions), and in general only regions close to the optodes, such as classical language areas, should be investigated with NIRS. Moreover, it is difficult to separate the NIRS signal originating from the cortex from those originating in other non-cortical tissues the light encounters (scalp, temporal muscle, skull, frontal sinus, cerebrospinal fluid and dura), which attenuate the NIR light before it reaches the cortex.

Additionally, optodes are most often positioned on the scalp relative to surface landmarks and not relative to individual participants' brain anatomy (contrary to fMRI). As a consequence, individual channels are likely probing slightly different anatomical areas in different subjects.

Another limitation is the lack of standardization even among the most widely used relative NIRS systems, with different laboratories using different wavelengths (the most effective wavelengths differ across NIRS systems), source-detector distances and probe placements, criteria

for data rejection, signal processing, and statistical analysis, which makes it still difficult to compare results across studies.

Since its first use to measure brain function, NIRS has found numerous applications as a research tool, among which cognitive development in infants and children is dominant (Boas et al., 2014; Benavides-Varela, et al., 2011; Cristià et al., 2013; Lloyd-Fox et al., 2010; Quaresima et al., 2012). It is also very effective to study normal and pathological brain physiology in general (Ferrari & Quaresima, 2012; Scholkmann et al., 2014; Nguyen et al., 2011; Watanabe et al., 2002), psychiatric conditions such as depression and schizophrenia (Ehlis et al., 2014), and strokes and brain injuries (Obrig, 2014).

II. 1. d. NIRS in our experiments

In this thesis, we have used fNIRS to explore speech perception and learning in the newborn brain. In particular three NIRS experiments aimed at exploring the C/V functional asymmetry at birth. For this purpose, we used the fNIRS paradigm that revealed the newborn brain's ability to detect repetition-based regularities carried by *syllables* (e.g. *mubaba vs. mubage*) (Gervain et al., 2008) in Experiments 1, 2 and 3. Our goal was to explore the relative role of *consonants and vowels* in speech processing. We therefore adapted the original Gervain et al. (2008) study for our purposes and in particular built a new adapted linguistic material. This material minimizes the acoustic and positional differences between consonants and vowels. To validate it, Experiment 1 replicates the ABB vs. ABC paradigm (Gervain et al., 2008) with this new material, and thus serves as a baseline experiment before Experiment 2 and 3 actually compare the roles that consonants and vowels play in the detection of repetition-based regularities at birth. Experiment 1 has already been written up and submitted as a paper on its own, with an emphasis on the effect of stimulus variability on neural repetition effects in the NIRS signal at birth, and it will be presented in the form it was submitted in the following section.

II. 2. EXPERIMENT 1: Baseline Experiment

Bouchon, C., Nazzi, T. & Gervain, J. (in revision). Hemispheric asymmetries in repetition enhancement and suppression effects in the newborn brain.

Section numbers have been added to the original paper.

Abstract

The repeated presentation of stimuli typically attenuates neural responses (repetition suppression) or, less commonly, increases them (repetition enhancement) when stimuli are highly complex, degraded or presented under noisy conditions. In adult functional neuroimaging research, these repetition effects are considered as neural correlates of habituation. The development and respective functional significance of these effects in infancy remain largely unknown.

This study investigates repetition effects in newborns using functional near-infrared spectroscopy, and specifically the role of stimulus complexity in evoking a repetition enhancement vs. a repetition suppression response, following up on Gervain et al. (2008). In that study, abstract rule-learning was found at birth in cortical areas specific to speech processing, as evidenced by a left-lateralized repetition enhancement of the hemodynamic response to highly variable speech sequences conforming to a repetition-based ABB artificial grammar, but not to a random ABC grammar. Here, the same paradigm was used to investigate how simpler stimuli (12 different sequences per condition as opposed to 140), and simpler presentation conditions (blocked rather than interleaved) would influence repetition effects at birth.

Results revealed that the two grammars elicited different dynamics in the two hemispheres. In left fronto-temporal areas, we reproduce the early perceptual discrimination of the two grammars, with ABB giving rise to a greater response at the beginning of the experiment than ABC. In addition, the ABC grammar evoked a repetition enhancement effect over time, whereas the ABB grammar no longer evoked a repetition effect. Right fronto-temporal areas showed neither initial discrimination, nor change over time to either pattern. Taken together with Gervain et al. (2008), this is the first direct evidence that stimulus variability influences the direction of the neural repetition effect in newborns. Further, this temporal modulation is restricted to the left hemisphere, confirming its specialization for learning linguistic regularities from birth.

II. 2. a. Neural repetition effects

Introduction Neurons often change their response frequency or intensity to repeated presentations of similar or identical stimuli. This was found at different levels of brain organization and using different measurement techniques. Single-neuron recordings (Miller & Desimone, 1994) directly show the reduced firing rate of neurons for stimulus repetition (Desimone & Duncan, 1995). Repetition suppression was also measured and interpreted as neuronal adaptation in electrophysiological activity measured over large populations of neurons over the scalp (EEG, Haenschel, Baldeweg, Croft, Whittington, & Gruzelier, 2000; Näätänen, 1984; May et al., 1999). Repetition effects have also been found in the hemodynamic correlates of brain activity (using

functional magnetic resonance imaging, fMRI, or near-infrared spectroscopy, NIRS), which is the focus of the current study.

In adults, fMRI studies have documented systematic decreases in the blood oxygen level-dependent (BOLD) response to repetitions of identical stimuli as compared to varying stimuli (Henson, 2003; and Grill-Spector et al., 2006, for recent reviews). fMRI experiments have thus widely used adaptation paradigms to probe brain regions specifically primed by repetitions of stimuli such as visual objects and their geometric properties (Wiggs & Martin, 1998; Grill-Spector & Malach, 2001), numbers (Naccache & Dehaene, 2001), as well as auditory and speech stimuli (Näätänen & Rinne, 2002; Belin & Zatorre, 2003; Cohen et al., 2004; Dehaene-Lambertz, Dehaene et al., 2006).

Although repetition suppression in fMRI is considered as the hemodynamic signature of neuronal adaptation, the relation between reduction in firing rates at the cellular level, electrophysiologically observed repetition suppression and its hemodynamic correlates is very complex and still debated (see Grill-Spector et al., 2006; Baldeweg, 2006). In clear illustration of this complex relationship, repeated stimuli are sometimes found to elicit increased hemodynamic responses, giving rise to a repetition enhancement rather than a repetition suppression effect (Shapiro et al., 2013). Repetition enhancement typically occurs when the repeated stimulus is initially unfamiliar (Henson et al., 2000), unattended (Vuilleumier et al., 2005) or under certain conditions affecting the quality of the stimulus presentation: short exposure time (Grill-Spector, et al., 2000), low visual stimulus quality or visibility (Dolan et al., 1997. Turk-Browne et al., 2007), low discriminability (Kourtzi et al., 2005), or stimulus degradation (James et al., 2006; Turk-Browne et al., 2007, 2006). Independently of the direction of change, repetition effects are often used to investigate cognitive functions and learning in the adult brain in various modalities, including language processing (see Segaert et al., 2013).

Much less is known about repetition effects in infancy. Recently, the cognitive/functional interpretation of repetition effects as indicators of habituation has been extended to the developing brain (Turk-Browne et al., 2008), suggesting that repetition effects could be neural signatures of habituation to repetitive stimuli and learning. This is based on the fact that behavioral studies with infants have shown decline in responding to a repeated stimulus, and these habituation effects have been successfully used to reveal many early abilities in various cognitive domains such as vision, hearing, memory, number cognition, language, and social cognition (see Aslin, 2007, for a review). This decreased behavioral responsiveness to a repeated stimulus has been proposed to reflect the simplest form of learning, already present prenatally (Sandman et al., 1997). So far, very few neuroimaging studies have measured the hemodynamic correlates of repetition effects in

developmental populations. The present study will contribute to this literature by investigating early repetition effects during speech processing/learning in the newborn brain.

In an fMRI study, 3-month-old infants showed an increase of activity, i.e. repetition enhancement, in their left inferior frontal region (especially Broca's area) when sentences were repeated every 14 seconds (Dehaene-Lambertz, Hertz-Pannier, Dubois, Mériaux, Roche, Sigman & Dehaene, 2006), whereas adult data had shown a repetition suppression effect in broad temporal and inferior frontal regions (Dehaene-Lambertz, Dehaene, et al., 2006). This developmental change (from enhancement to suppression) was attributed to infants' less mature linguistic and memory capacities than those of adults. But importantly, both the adult and the infant results were interpreted as evidence for memory function.

As NIRS is increasingly becoming the method of choice for imaging newborns and young infants (Lloyd-Fox, et al., 2010; Obrig, Rossi, Telkemeyer, & Wartenburger, 2010; Minagawa-Kawai et al., 2011; Gervain et al., 2011), it is important to better understand the dynamics of the metabolic response in these developmental populations, which is what our current study undertakes. Repetition effects in the NIRS signal have not yet been investigated systematically, but a few NIRS studies did observe changes in neural activity over time in response to repeated speech stimuli in newborns (Gervain et al., 2008; Wagner, Fox, Tager-Flusberg & Nelson, 2011; Benavides-Varela, Gomez, Macagno, Bion, Peretz & Mehler, 2011). Benavides-Varela et al. (2011) revealed effects of dishabituation in the NIRS signal in newborns tested on a novel word, *pelu*, after they had been familiarized to repetitions of the word *mita.* Another NIRS study looked at the time course of responses to repeated speech stimuli in newborns and was the first to report repetition effects at birth using this technique (Gervain et al., 2008; see Wagner et al., 2011, for a replication with 7- and 9-month-olds). The original paradigm by Gervain et al. (2008) aimed at testing rule learning abilities at birth, by presenting neonates with trisyllabic sequences implementing a repetition-based ABB pattern or a random ABC control pattern (e.g. "mubaba" vs. "mubage", respectively). The experiment consisted of 28 blocks, 14 per pattern. Each block contained 10 sequences, and no sequence was repeated over the experiment. Thus a total of 140 *different* sequences were presented for both patterns. On average, NIRS responses were stronger for ABB than for ABC sequences in both the left and right temporal areas and in the left frontal area. Additionally, a time course analysis showed that in the left frontal area, this increased response for ABB over ABC sequences was already present in the first 4 blocks, and further increased over the course of the study. This was interpreted as a two-step process, involving a first automatic and rapid detection of the repetition-based structure signaled by the initially superior activation in response to ABB over

ABC, followed by a gradual learning process of the abstract repetition-based structure, signaled by the increasing difference in activation between ABB and ABC in the course of the 14 blocks.

Although the above pattern of results in the left frontal area, known for its role in higher order learning (Dehaene-Lambertz et al., 2006), did suggest an abstract regularity learning mechanism, it is not clear why an enhanced, rather than a suppressed neural response was observed. One possible explanation is that the entire set of 140 repetition-based sequences taken together constituted a highly variable material to which habituation had not occurred during the time course of the study. Indeed, although all ABB sequences conformed to the same ABB regularity, they were all different, and used all the possible combinations of a large number of different syllables, resulting in a highly complex and variable material the learning of which might have required considerable processing effort from the newborn participants. Following Turk-Browne et al.'s (2008) analogy between repetition effects and habituation, the left lateralized repetition enhancement effect in the NIRS signal observed in Gervain et al. (2008) might indicate a novelty effect, i.e. not fully completed learning, for the ABB grammar. According to this interpretation, then, a simpler stimulus material might lead to complete learning and thus repetition suppression over time.

The current study seeks to directly address this hypothesis, which, if confirmed, would contribute both to our understanding of repetition effects in the infant hemodynamic response and to how regularities in speech are learned at birth. To address these issues, the current study tested newborns' detection and learning of the ABB vs. ABC patterns in trisyllabic sequences, as in Gervain et al. (2008), but with two substantial differences in the construction of the stimulus material, which might favor more complete learning and thus an absence of a repetition enhancement effect. First, in order to increase levels of cortical activation and thus improve statistical power in the time course analysis, we did not interleave blocks of ABB and ABC sequences as in Gervain et al. (2008), but rather presented each condition separately (Friston et al., 1999; Buckner et al., 1996; Aguirre & D'Esposito, 1999; Chee et al., 2003). Second, and more importantly, the current material contained only a limited number of different sequences to increase the probability of habituation to the material (see details below).

I. 2. b. Materials and methods

NIRS is a neuroimaging technique that measures the hemodynamic correlates of neural activity, using near infra-red light shed at the scalp. From the differential photon absorption properties of oxygenated and deoxygenated hemoglobin, changes in oxy- and deoxy-hemoglobin concentration can be calculated (Jöbsis, 1977; Meek, 2002), thus investigating cortical activity with a penetration depth of approximately 10-15mm into the neonate brain (Okada et al. 1995; Gervain et

al., 2011). fNIRS, has been widely used in developmental cognitive neuroscience over the last two decades due to its advantages over fMRI (non-invasiveness, relative tolerance to head movements, absence of acoustic noise, relative affordability etc.), as well as its precise spatial localization, which EEG lacks (Obrig & Villringer, 2003).

Participants

Twenty-four full-term, healthy neonates (ranging in age from 1 to 3 days, mean age: 1.8 days; Apgar score 10 minutes after birth: 10; 13 girls; mean head circumference = 34.5 cm) born to French-speaking mothers were included in the analyses. An additional 23 infants were tested, but their data were excluded due to the infant becoming awake or fussy and failing to complete the procedure (11), equipment failure (2), or insufficient analyzable data (10). Data were considered insufficient when more than 50% of the channels had less than 50% non-rejected blocks. All infants were tested while asleep. All parents of infants gave informed consent prior to beginning the experiment.

Stimulus material

As in Gervain et al.'s (2008) Experiment 1, newborns were exposed to a repetition-based ABB artificial grammar (e.g. "mulele", "junana") and random ABC control grammar (e.g. "mulevi", "junary"). Both grammars generated trisyllabic sequences and the two grammars were matched on all nonstructural properties: syllabic repertoire (***Table II. 1.***); frequency of the A, B, and C syllables; phonological characteristics; flat prosody; and transitional probabilities between syllables. Sequences were synthetized using the fr4 female voice of the MBROLA diphone database, in a monotonous pitch of 200 Hz, with the same length for all phonemes (150ms).

However, unlike in Gervain et al. (2008), the current experiment used simpler material. Specifically, we used 12, rather than 20 different syllables, made up of 6 different consonants (Cs) and 6 different vowels (Vs), rather than of 12 Cs and 5Vs. Furthermore, these syllables were combined to yield 12, rather than 140 different trisyllabic sequences per grammar. As a result, in the current experiment the same sequence was presented multiple times during the session (6 times exactly), introducing redundancy in the sequences presented, which was completely absent from Gervain et al. (2008). The mean transitional probability (TP) between adjacent syllable pairs was 0.33 for a total number 48 combinations here, rather than a mean TP of 0.10 for 254 different combinations in Gervain et al., (2008), thus yielding a relatively more predictable material. Details are summarized in ***Table II. 1***.

	CV inventory	Number of different sequences generated by CV combinations	Adjacent Syllables TPs Mean(SD)	Total number of sequences presented in experiment	Sequence frequency in experiment
Experiment 1 (Gervain et al., 2008)	20 CV syllables 12 Cs, 5 Vs	280 CVCVCVs: • 140 ABB • 140 ABC	0.10 (0.07)	280 (14 blocks of 10 sequences per condition)	Exactly 1 time
Current experiment	12 CV syllables 6 Cs, 6 Vs (*li le ry ra mu mo ny na vi ʒu ʒo ve*)	24 CVCVCVs: • 12 ABB • 12 ABC	0.33 (0.12)	144 (12 blocks of 6 sequences per condition)	Exactly 6 times, in 6 different blocks

Table II. 1. Experiment 1. Details on the stimuli used in Experiment 1 Gervain et al. (2008) Experiment 1 and in the current Experiment 1.

Procedure Neonates were tested in a local maternity hospital, in a silent experimental room, while asleep or at rest in a bassinet. Each infant was presented with 24 blocks of stimuli. Each block consisted of six CVCVCV sequences and lasted 9.4 sec. Within a block, the six sequences were separated by brief pauses of variable length (0.5 – 1.5 s). Blocks were also spaced at time intervals of varying duration (20 –25 s) to avoid inducing phase-locked brain responses. The total testing time for each infant was 14.6 min. The block design used is presented in ***Figure II. 6.*** All participants were presented with both the ABB and ABC conditions. Condition order was randomized and counter-balanced across participants, such that half of the participants were presented with the ABB condition first, the other half with the ABC condition first.

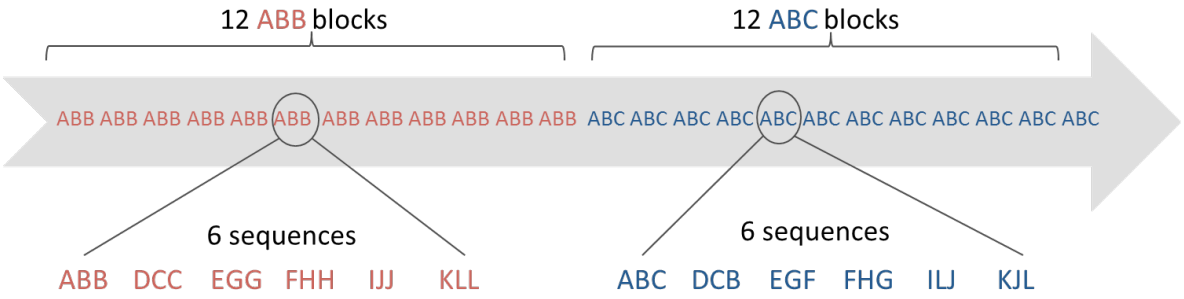


Figure II. 6. *Experiment 1. The block design used in the current study. Condition order was randomized and counter-balanced across infants. Letters from A to L represent the 12 syllables extracted from the CV inventory.*

Optical imaging was performed with an optical topography imager (*NIRScout* 816, *NIRx Medizintechnik GmbH*, Berlin, Germany) using pulsated LED sequential illumination (5 mW) with two wavelengths of 760 nm and 850 nm to record the NIRS signal at a 10.417 Hz sampling rate. Four LED sources were placed on each side of the head in analogous positions, and were illuminated sequentially. They were coupled with 5 detectors on each side of the head. The configuration of the 24 channels (12 per hemisphere) created with the 4 sources and 5 detectors per hemisphere is

shown in ***Figure II. 7. A.*** To keep the optodes in place with a standard 3 cm separation, they were embedded in a cotton cap (*Easycap*) of 36 cm diameter (***Fig. II. 7. B.)***. A MacBook laptop running PsyScope controlled the experiment, played the language stimuli and sent markers to the NIRS machine. The stimuli were played through two speakers approximately 1 m from the infants' head.

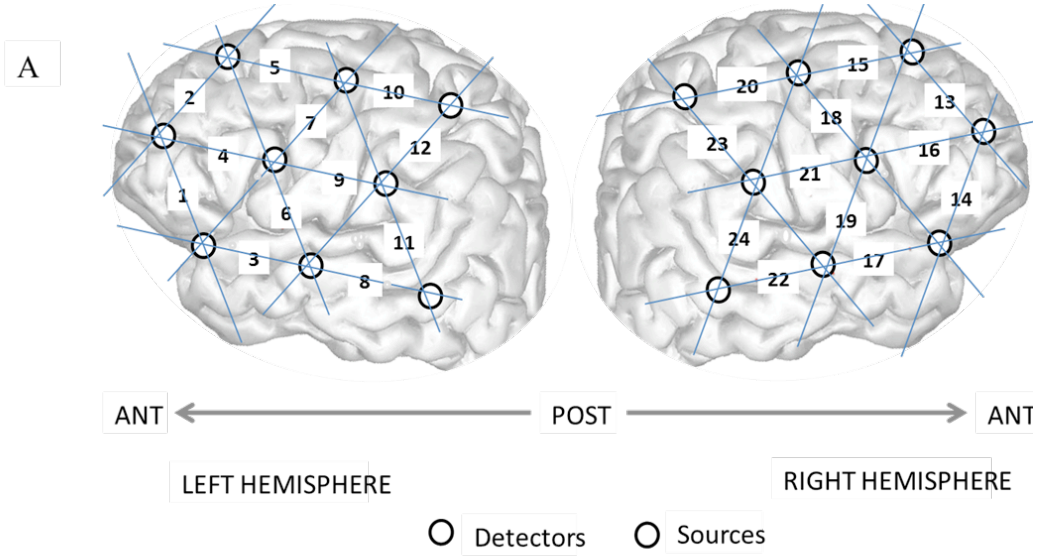


Figure II. 7*. A. Configuration of our probe sets overlaid on a schematic infant brain.* ***B.*** *Picture of a newborn participant with optodes placed upon the head*

Data analysis

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

block-channel pairs in which a change in concentration greater than 0.1 mmol × mm over a period of 0.2 s, i.e. two samples, occurred, and rejecting the block for that channel. Channels with data for less than 5 out of 12 blocks per condition were discarded. A baseline was established by using a linear fit over the 5 s time-window preceding the onset of the block and the 5 sec window beginning 15 s after the end of the block. The 15 s resting period after stimulus offset was used to allow the hemodynamic response function (HRF) to return to baseline (Peña et al., 2003; Gervain et al., 2008). Analyses were conducted in MATLAB (version R2012b) with custom analysis scripts.

I. 2. c. Results

The grand average results are presented in ***Figure II. 8.*** The figure shows the oxyHb and deoxyHb concentration changes averaged across all blocks of each condition and across all infants. The channel-by-channel *t*-test results for the mean oxyHb and deoxyHb changes are shown in ***Figure II. 9.*** for the following comparisons: the average of the two conditions vs. baseline, ABB vs. baseline, ABC vs. baseline, and ABB vs. ABC.

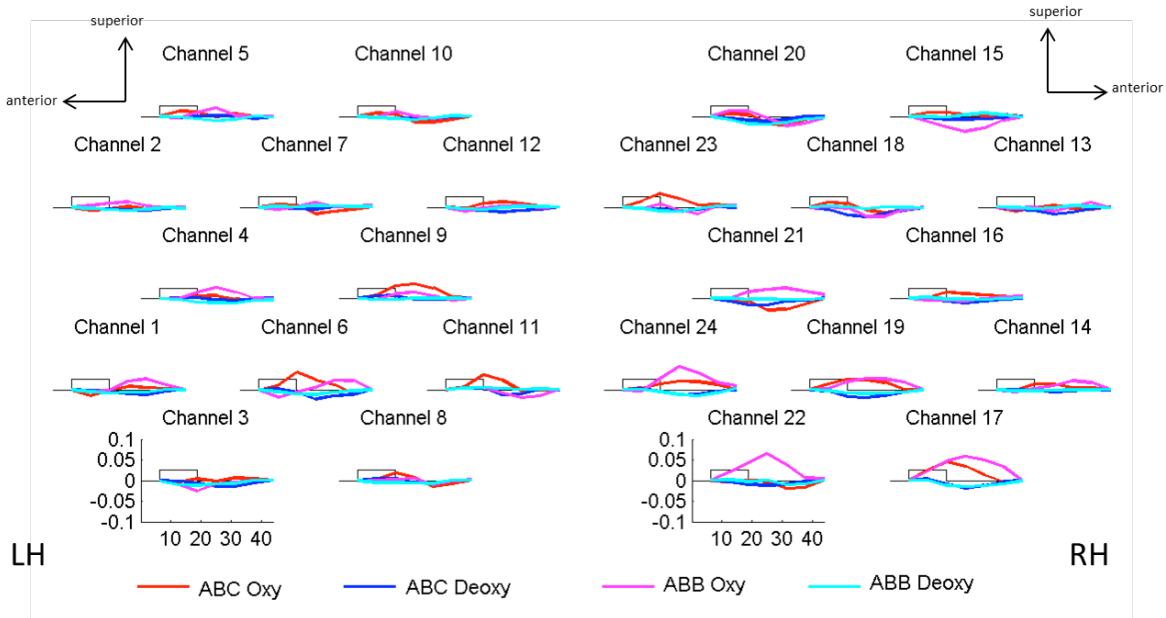
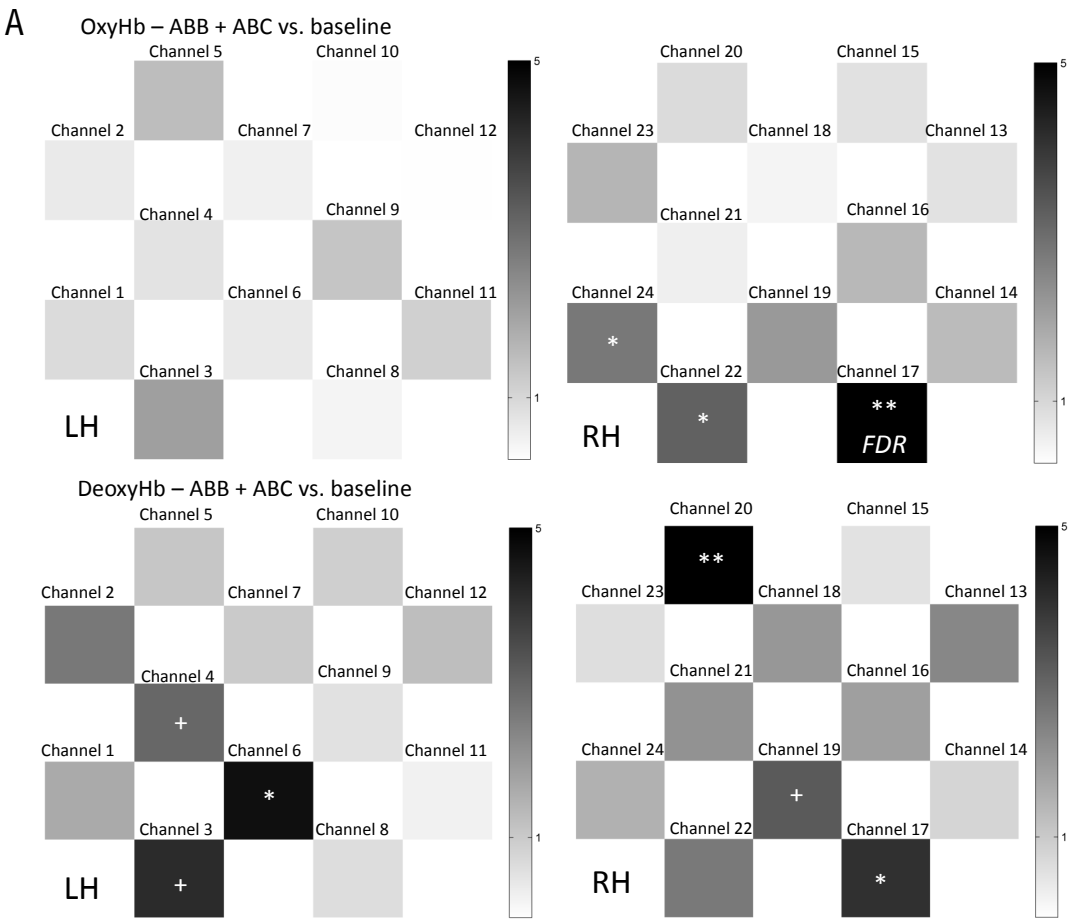
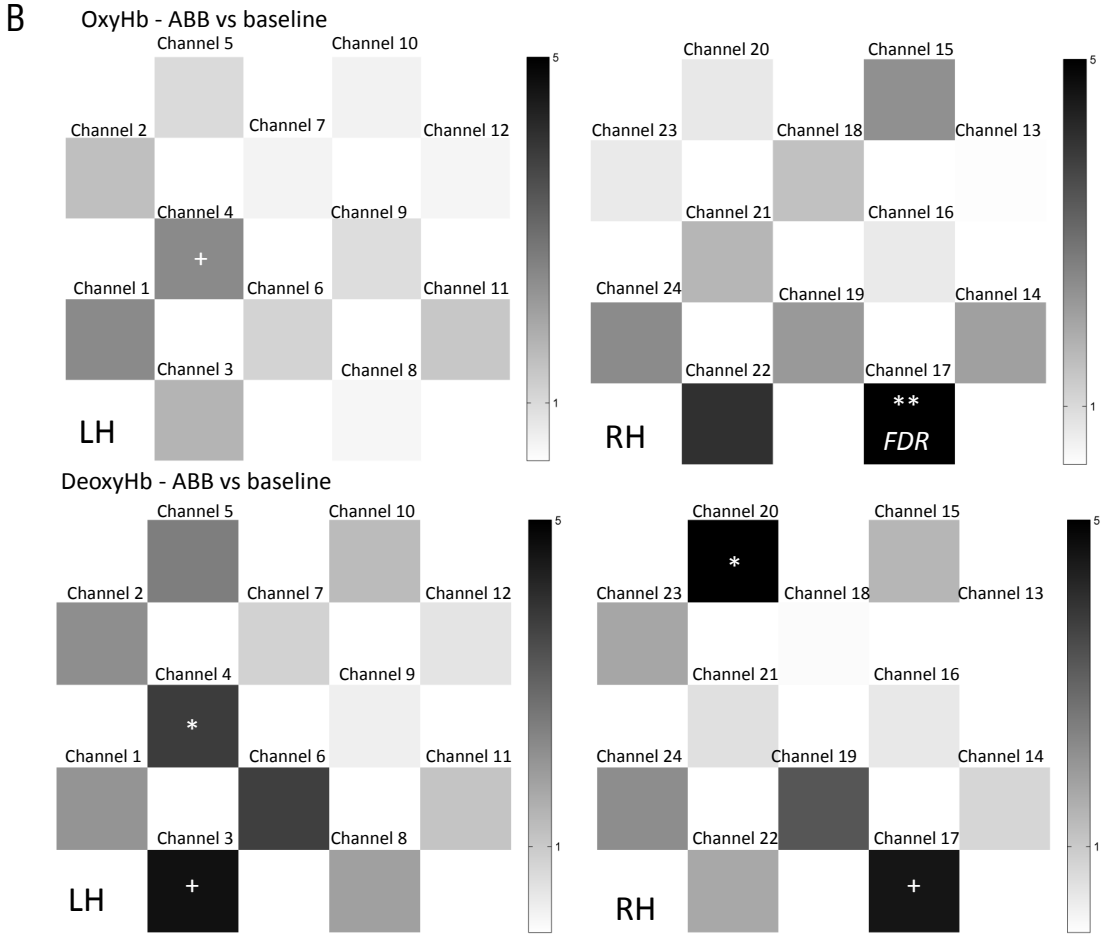


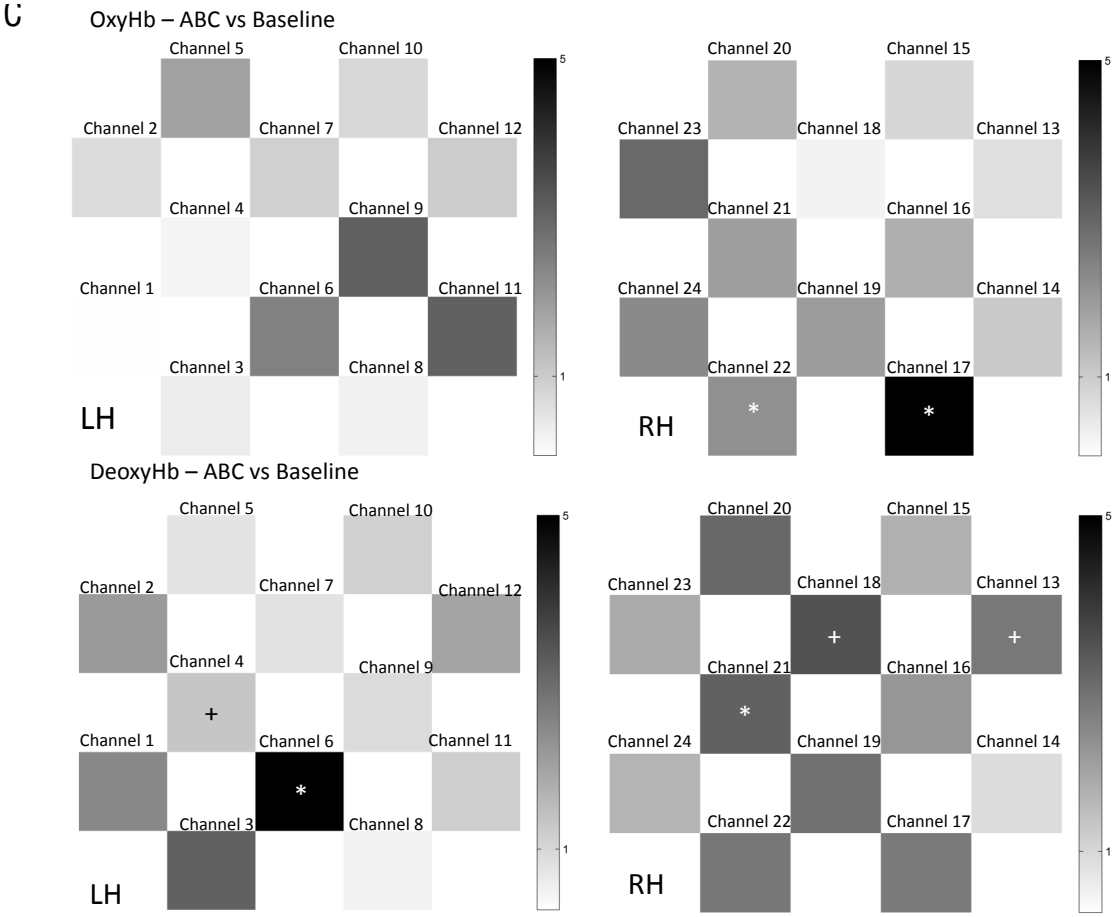
Figure II. 8. *Experiment 1. Grand average results. Numbers and location of channels correspond to the placement shown in Figure II. 7. A. The x-axis represents time in seconds; the y-axis shows concentration in mmol x mm. The rectangle along the x-axis indicates time of stimulation. The continuous red and blue lines in the graphs represent oxyHb and deoxyHb concentrations, respectively, in response to the ABC grammar. The dashed magenta and cyan lines represent oxyHb and deoxyHb concentrations, respectively, in response to the ABB grammar.*


A
OxyHb – ABB + ABC vs. baseline
Channel 5
Channel 10
Channel 2
Channel 7
Channel 12
Channel 4
Channel 9
Channel 1
Channel 6
Channel 11
Channel 3
Channel 8
LH
5
1
Channel 20
Channel 15
Channel 23
Channel 18
Channel 13
Channel 21
Channel 16
Channel 24
Channel 19
Channel 14
Channel 22
Channel 17
RH
**
FDR
DeoxyHb – ABB + ABC vs. baseline
+
*
+
**
*

B


OxyHb - ABB vs baseline
Channel 5
Channel 10
Channel 2
Channel 7
Channel 12
Channel 4
Channel 9
+
Channel 1
Channel 6
Channel 11
Channel 3
Channel 8
LH
5
1
Channel 20
Channel 15
Channel 23
Channel 18
Channel 13
Channel 21
Channel 16
Channel 24
Channel 19
Channel 14
Channel 22
Channel 17
**
FDR
RH
5
1
DeoxyHb - ABB vs baseline
Channel 5
Channel 10
Channel 2
Channel 7
Channel 12
Channel 4
Channel 9
*
Channel 1
Channel 6
Channel 11
Channel 3
Channel 8
+
LH
5
1
Channel 20
Channel 15
*
Channel 23
Channel 18
Channel 13
Channel 21
Channel 16
Channel 24
Channel 19
Channel 14
Channel 22
Channel 17
+
RH
5
1

C


OxyHb – ABC vs Baseline
Channel 5
Channel 10
Channel 2
Channel 7
Channel 12
Channel 4
Channel 9
Channel 1
Channel 6
Channel 11
Channel 3
Channel 8
LH
5
1
Channel 20
Channel 15
Channel 23
Channel 18
Channel 13
Channel 21
Channel 16
Channel 24
Channel 19
Channel 14
Channel 22
Channel 17
RH
*
*
5
1
DeoxyHb – ABC vs Baseline
Channel 5
Channel 10
Channel 2
Channel 7
Channel 12
Channel 4
Channel 9
+
Channel 1
Channel 6
Channel 11
*
Channel 3
Channel 8
LH
5
1
Channel 20
Channel 15
Channel 23
Channel 18
Channel 13
+
+
Channel 21
Channel 16
*
Channel 24
Channel 19
Channel 14
Channel 22
Channel 17
RH
5
1

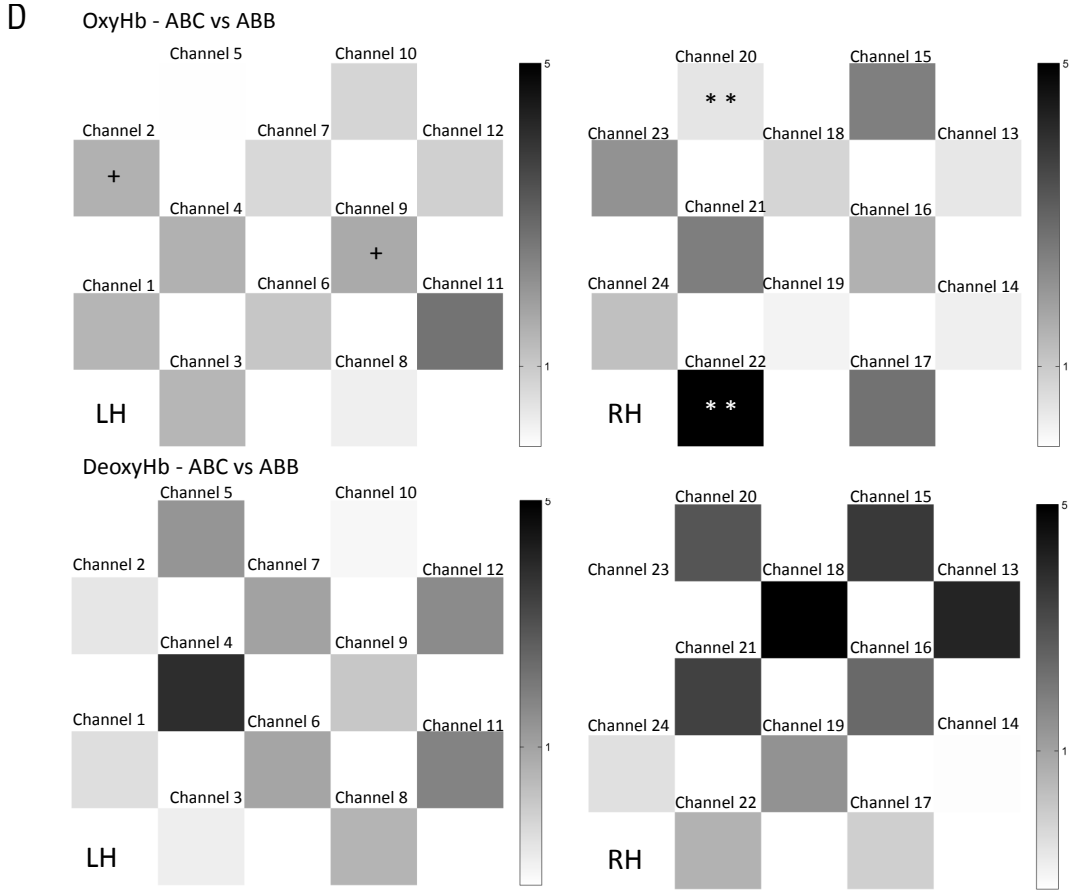


Figure II. 9*. Experiment 1. Statistical maps (t-maps) comparing oxyHb and deoxyHb responses to: (A) ABB+ABC vs. baseline, (B) ABB vs. baseline, (C) ABC vs. baseline, (D) ABB vs. ABC. Channels are plotted following the same placement as in Figure II. 7. A. The t-values for each channel are color-coded as indicated on the color bar. Significance levels are indicated for each channel by p-values: '+' marginally significant without correction for multiple comparisons;'**' p < .01 significant without correction; '*' p < .05 significant without correction; 'FDR' p < .05 significant after correction for multiple comparisons using the False Discovery Rate (Benjamini & Yekutieli, 2001).*

As Condition Order did not have significant main effects or interactions in preliminary analyses, we collapsed over this factor in any further analyses. A first analysis of variance (ANOVA) with the within-subject factors of Grammar (ABB/ABC), Hemisphere (Left/Right) and Region of Interest (frontal/temporal) was run on average Hb concentration to evaluate whether the two grammars are processed differently overall, as in Gervain et al. (2008). These ROIs were chosen given our probe configuration (***Fig. II. 7***.) in order to evaluate the responses in temporal areas (LH: channels 3, 6, 8, 11; RH: channels 17, 19, 22, 24) involved in auditory processing and in the frontal areas (LH: channels 1, 2, 4, 5; RH: channels 14, 13, 16, 15) involved in structural and more abstract processing in infants (Peña et al., 2003; Friederici, 2002; Gervain et al., 2008). Separate ANOVAs were conducted for oxyHb and deoxyHb.

The ANOVA for oxyHb yielded a significant interaction between Hemisphere × ROI [$F(1, 23) = 5.99$, $p = .02$], as the activity in the RH was significantly greater in the temporal than in the frontal ROI , and temporal activity was significantly greater in the RH than in the LH (RH: frontal < temporal: $p = .004$; temporal ROIs: left < right: $p = .01$, these and all further post hoc p-values indicate results of Scheffe's post hoc test for the relevant pairwise comparisons, unless otherwise specified). The interaction between Hemisphere × ROI × Condition was also significant [$F(1, 23) = 4.71$, $p = .03$] due to a significant right-lateralized advantage for the ABB condition in temporal ROIs (RH: ABB in temporal ROI > ABB in frontal ROI, $p = .001$; temporal ROIs: ABB right > ABB left, $p = .003$), and to several marginal tendencies (the ABB grammar yielded marginally greater activity than the ABC grammar in the left frontal ROIs ($p = .08$) and in the right temporal ROIs ($p = 07$); and the left frontal ROI was marginally more active than the right frontal ROI in response to ABB ($p = .09$). Moreover, the ANOVA for deoxyHb yielded a marginal Hemisphere × ROI interaction [$F(1, 23) = 2.97$, $p = .09$] due to significantly greater activation in the right temporal ROI over the left temporal ROI (p = .03). No other effects or interactions approached significance.

Together with the *t*-tests results (***Fig. II. 9.***), this first analysis revealed a somewhat different pattern than in Gervain et al. (2008). Contrary to the generally stronger activation found for ABB over ABC, establishing a clear general difference in the processing of the ABB and ABC grammars, we found no such effect. However, the greater response in the temporal as compared to the frontal channels was similar to the pattern observed in Gervain et al. (2008), suggesting that the stimuli were successfully processed by the auditory cortex in the current study, too. Additionally though, stronger responses were found in the right than in the left hemisphere, the opposite of what Gervain et al. (2008) found.

Importantly for the purposes of the current study, we ran a second ANOVA to examine the temporal dynamics of the responses to the ABB and ABC grammars during the course of the 12 consecutive blocks. This analysis also allowed us to test whether the apparent absence of repetition detection in the first, overall analysis as well as the apparent right hemisphere advantage could be attributed to a masking effect of a potential general decrease in activation occurring over time in response to the current material. In a similar analysis, Gervain et al. (2008) found that the perceptual detection of the ABB grammar was immediately observable in the first 4 blocks, and it further increased over the remaining blocks, whereas the response to the ABC grammar did not change throughout the experiment. Based on previous results regarding auditory perception in newborns (Peña et al., 2003; Gervain et al., 2008), and the results of *t*-tests and the first analysis, a single ROI per hemisphere was used for this analysis: channels [1, 3, 8] corresponded to the left fronto-temporal area and channels [14, 17, 22] to the right fronto-temporal area (***Fig. II. 7.*** *A*).

We conducted an ANOVA on average Hb concentrations within the target ROI with the factors Grammar (ABB /ABC), Time (3 initial blocks/ 3 final blocks), and Hemisphere (LH/RH). The ANOVA for oxyHb yielded a marginal main effect for Grammar [F(1, 23) = 3.80, p = .053], as ABB gave rise to a larger response than ABC. The Time × Grammar interaction was marginally significant [F(1, 23) = 3.63, p = .059], as initial blocks showed an advantage for ABB over ABC activation, whereas activations to both grammars were equivalent in the final blocks. Importantly, the Time × Grammar × Hemisphere interaction was significant [F(1, 23) = 5.85, p = .017; ***Fig. II. 10.***]. This was due to the fact that in the LH, the ABC grammar gave rise to an increasing response between initial and final blocks (p = .0005). Moreover, in the LH, there was a significantly stronger initial activation in response to the ABB over the ABC grammar (p = .002), but no such difference in the final blocks, where the activation for the ABC grammar had reached the same level as for ABB. In the RH, there was no difference between ABB and ABC in both the initial and the final blocks.

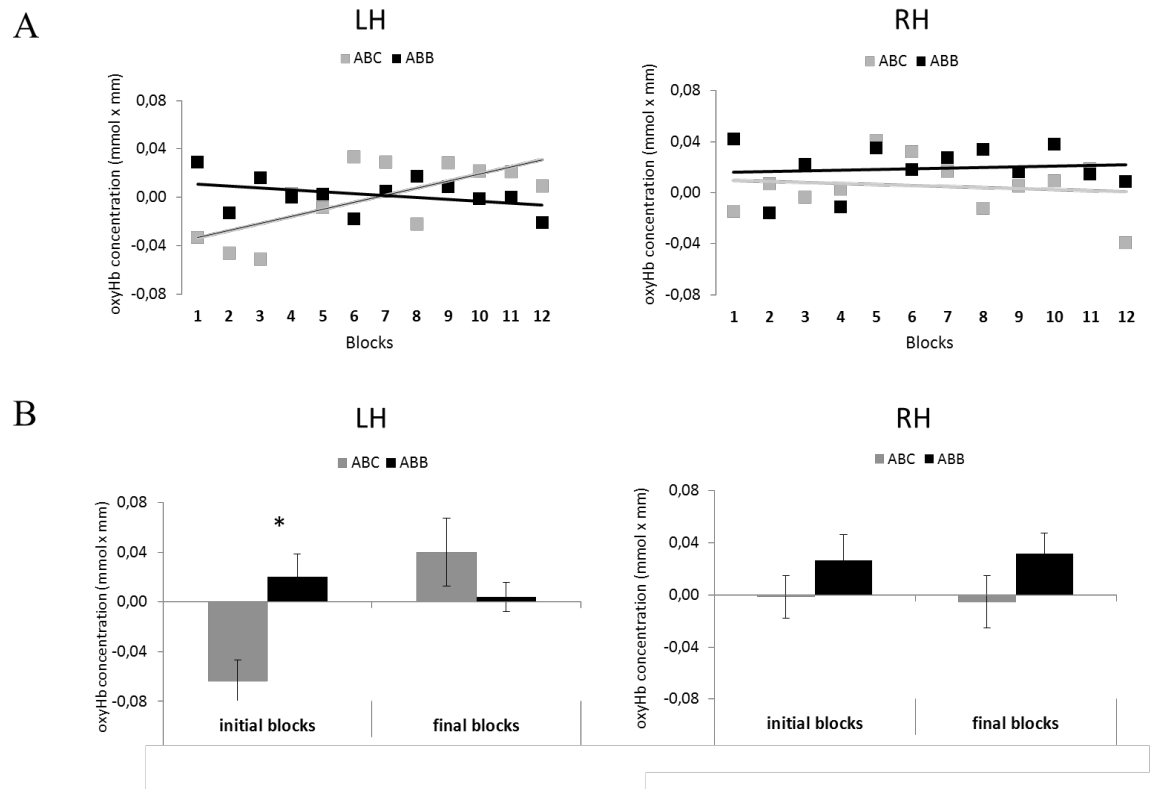


Figure II. 10. *Experiment 1. The time course of the responses for the two grammars in the left fronto-temporal ROI (channels 1, 3, 8) vs. the right fronto-temporal ROI (channels 14, 17, 22). The y-axis shows the average oxyHb concentration in mmol x mm. ABC is plotted in light-grey; ABB in black.* ***A.*** *Linear regression lines of the oxyHb concentrations fitted on the data points provided by the 12 consecutive blocks for the two grammars, plotted on the x-axis.* ***B****. The bars indicate the average oxyHb concentration for the first ("initial blocks") and the last three blocks ("final blocks").*

The same ANOVA for deoxyHb also yielded a significant Time × Grammar × Hemisphere interaction [$F(1, 23) = 6.51$, $p = .012$]. This was due to ABB giving rise to a stronger initial deoxyHb response than ABC in the LH ($p = .040$), and ABC giving rise to a significant increase in the LH from

initial to final blocks (p = .027). No other effects or interactions approached significance. In the RH, there was no difference between ABB and ABC in both the initial and the final blocks.

Overall, the oxy- and deoxyHb responses show a left-lateralized advantage for ABB over ABC in the first blocks, a significant increase in neural activity over time in response to ABC, and a stable or slightly decreasing response to ABB between the initial and final blocks.

I. 2. d. Discussion

The goal of the present experiment was to assess the stimulus presentation conditions under which repetition enhancement (Gervain et al., 2008) and repetition suppression effects might be observed in newborn hemodynamic activity in response to language stimuli. Specifically, we hypothesized that the repetition enhancement, rather than suppression, effect found in Gervain et al. (2008) was due to the high variability in the stimulus material used, requiring too long and effortful learning for neural habituation to occur during the time course of that study. Therefore, we tested whether the enhancement effect might be attenuated under conditions of relatively low stimulus variability, which might allow the newborn brain enough time and repetitions to show habituation. Accordingly, the current experiment tested the detection and learning of a repetition-based ABB grammar vs. an ABC random grammar as in Gervain et al. (2008), but using a relatively simpler and much more redundant list of stimuli. Indeed, while Gervain et al. (2008) used 20 different syllables combined into 140 different sequences per condition, here we used a list of 12 different syllables combined into 12 different sequences. Moreover, instead of interleaving the two conditions, they were presented consecutively.

Our study has three major findings. First, we observed a fronto-temporal left-lateralized advantage for the ABB over the ABC grammar during the three initial blocks of the experiment. This initially stronger response to ABB over ABC, occurring especially in the left fronto-temporal channels, converges with the results of Gervain et al. (2008) as well as with behavioral findings (Endress, Nespor & Mehler, 2009) and thus confirms an automatic detection mechanism specific to the structured ABB pattern over the random pattern. This finding thus confirms the neural sensitivity of auditory and speech processing areas to adjacent repetitions at birth.

Secondly, we observed a strong hemispheric asymmetry, with the LH showing differential response dynamics, while activations in the RH, although relatively strong overall, remained constant in time for each condition. This result is in line with previous accounts of speech processing being left lateralized from birth (Peña et al., 2003).

Thirdly, and most importantly, we uncovered significant modulations in the time course of the responses to the two grammars, which, when compared with the results of Gervain et al. (2008),

reveals the importance of variability/redundancy in the stimuli for inducing repetition effects. Here, after a small initial response to ABC, newborns' left fronto-temporal cortex increased its activity over time to the ABC grammar, while the response to ABB, which was significant in the initial blocks, remained constant throughout the 12 blocks. A comparison with Gervain et al. (2008), where left fronto-temporal responses to ABB continued to increase during the course of the experiment while they remained constant for ABC, suggests that whether sequences were highly variable or very redundant has an impact on the direction of changes in neural response over time to the two grammars.

However linking the current findings to habituation and cognitive processes is not straightforward. The facilitation model proposed by Grill-Spector et al. (2006) suggested that the greater the number of repetitions (here the redundancy in the material), the faster and the earlier the memory trace formation is detected, diminishing the hemodynamic response until no activation exists. Interestingly, our simpler and more redundant material shaped the direction of the resulting learning curves in opposite directions for the two grammars. This may be explained considering the different underlying learning processes that the two grammars may have evoked. While both grammars contained 12 different sequences repeated 6 times, allowing the memorization of specific sequences, only the ABB sequences carried a structured pattern that may have elicited abstract rule learning. Therefore it is not surprising that repetition effects are affected in a complex manner, and our material is not simply yielding repetition suppression instead of enhancement as Grill-Spector's model would have predicted, because decreasing variability and increasing predictability in the sequences compared to Gervain et al. (2008) is likely to affect ABB and ABC stimuli differently since different learning processes are likely to take place in the two conditions.

Indeed, to explain the absence of repetition enhancement (i.e. the increased, but constant activation) for the ABB grammar in our study, we hypothesize that the repetitive ABB pattern being implemented in a smaller number of sequences than in Gervain et al. (2008) leaves less opportunity for abstract learning processes to occur. This is consistent with many behavioral results indicating that variability in the input modulates the nature of learning mechanisms. In particular, in speech perception and language acquisition, introducing variation in the input was found to help rule extraction (Gomez, 2002), while increasing redundancy, i.e. the number of repetitions of the same item in the input, led to rote memorization (Endress & Bonatti, 2005). This suggests that when learners are extracting the information that stays constant across variable forms, more variability yields more abstract learning. Conversely, one possible explanation of the enhancement effect observed here for ABC is that the repetitive presentation of only 12 ABC sequences induced memorization processes (Benavides-Varela et al., 2011), which were not induced in Gervain et al.

(2008) in which the same sequence never occurred twice and altogether the stimuli were much less predictable. This was not observed for the ABB condition mostly likely because, as discussed above, it elicited more abstract rule extraction and generalization. Our findings thus highlight the importance of variability, complexity and redundancy in the stimuli in inducing repetition effects, and the need for further investigation of the complex links between stimulus variability/predictability, repetition suppression/enhancement evoked, and underlying learning processes.

It might be argued that the other methodological difference alone, i.e. the blocked rather than interleaved presentation of the ABB and ABC blocks, could account for the differing results between the two studies. It is possible that the current experiment, contrary to Gervain et al. (2008), provided an opportunity for uninterrupted learning, which facilitated the abstract learning of the ABB grammar and led to the absence of repetition enhancement effect. However, based on previous adult fMRI results where repetition suppression persisted even with intervening stimuli (Henson et al., 2000), it is unlikely that interleaved presentation alone could reverse the direction of the repetition effects. Moreover, the early automatic advantage for ABB sequences over ABC was replicated in the current experiment, suggesting that the basic detection of repetition did occur whether blocks were interleaved or not.

The present results also contribute to a better understanding of how speech is processed and how regularities are learned at birth. First, the adjacent repetition pattern seems to be automatically detected, as was found consistently at birth (Gervain et al., 2008, 2012), and this advantage occurred in the initial blocks, suggesting that this detection mechanism is context-independent, automatic, rapid, and perceptually based, allowing for the immediate discrimination of a local adjacent repetition carried by syllables. The changing dynamics over time further indicates that in addition to this automatic detection mechanism, a more abstract learning process might also take place. Moreover, differential response dynamics to the two grammars occurred only in the left hemisphere, especially in left infero-frontal and temporal channels, known to be responsible for speech and language processing. Although of interest, our optode configuration did not allow us to test responses in more anterior frontal areas, known to be involved in abstract, higher order learning (Minagawa-Kawai et al., 2011).

In conclusion, these results provide new evidence of repetition effects at an early age in the auditory modality and more specifically for speech processing. Specifically, they highlight the importance of an experimental variable, stimulus complexity, which impacts the amount of processing and learning required, and thus the direction of the repetition effect.

I. 2. e. From syllable to segment-based repetition detection

In the paper "*Hemispheric asymmetries in repetition enhancement and suppression effects in the newborn brain*", the findings of Experiment 1 were mainly interpreted in terms of their contribution to the understanding of neural repetition effects in the NIRS signal at birth. Importantly, by confirming the detection of a *syllabic* repetition-based regularity in the newborn brain using a novel stimulus material, Experiment 1 has several methodological implications for the investigation of the origins of the C/V functional asymmetry conducted in Experiments 2 and 3. First, Experiment 1 validates the NIRx NIRS system and set-up. Second, the new material with its reduced variability as compared to Gervain et al. (2008) built purposely for this goal has also been validated. Third, the blocked presentation of the conditions adopted here led to similar conclusions as with interleaved conditions (Gervain et al, 2008, 2012). Thus Experiment 1 paves the way for the comparison of *consonant* vs. *vowel* repetition detection in Experiments 2 and 3 with the same system, material and blocked-condition design.

II. 3. EXPERIMENTS 2 and 3

II. 3. a. Segmental repetition detection and the C/V functional asymmetry

The initial proposal by Nespor et al. (2003) stated that the origin of the C/V functional asymmetry observed in the world's languages and in adult speech processing (Toro, Shukla, Nespor, & Endress, 2008; Toro, Nespor, Mehler, & Bonatti, 2008; Bonatti, Peña, Nespor, & Mehler, 2007) could be due to innate biases. Many recent studies have investigated the relative contributions of consonants and vowels to lexical and grammatical processing in children and infants. However the consonant bias for lexical processing and the vocalic bias for prosodic/syntactic processing have never been directly and simultaneously investigated at birth. This lack of investigation may be explained by the theoretical and methodological difficulty to reliably evaluate abilities related to lexical and prosody/morphosyntax in newborns in a single paradigm. However, recent studies have revealed precursors of these complex linguistic functions that are operational from birth. Experiment 2 and 3 follow up on these findings in order to test the presence of the C/V asymmetry at birth. Indeed, both *a syntax-like mechanism* consisting of the extraction and learning of structural regularities (Gervain et al., 2008, 2012) and *lexical item memorization*, whereby newborns recognize familiarized word forms (Benavides-Varela et al., 2011, 2012) have been observed at birth. Moreover

these studies used fNIRS imaging, a technique particularly adapted to young populations (Gervain et al., 2011).

The aim of Experiment 2 and 3 is thus to explore the innateness of the consonant bias in item memorization and of the vocalic bias in rule learning, using the single paradigm of repetition-based regularity detection (Gervain et al., 2008, 2012). For the first time, these two complementary experiments are evaluating the bootstrapping power of these biases for lexical and syntax acquisition, as proposed by Nespor et al. (2003). Gervain and colleagues (2008, 2012) have set up a paradigm that robustly reveals the extraction of the common structure shared by a list of word-like units (e.g. ABB: *mubaba, talulu, penana…*) and differentiate them from unstructured units (ABC: *mubage, talupi, penaku*…). The specific neural response found in newborns' left fronto-temporal cortex for syllabic repetition-based regularities is interpreted as evidence of a structural extraction and generalization ability, thus demonstrates a precursor of syntactic processing at birth. Therefore, this paradigm can be adapted to explore whether this ability applies to repetition regularities carried by *segments* instead of syllables (e.g., *mubabe* for an ABB consonant repetition; *mubaga* for an ABB vowel repetition). This would be important for language acquisition given that segmental regularities are relevant to signal morphological and syntactic functions in languages (Berent, Marcus, Shimron, & Gafos, 2002). In a behavioral task similar to the original repetition-based rule-learning experiment by Marcus et al. (1999), Pons and Toro (2010) found that 11-month-olds were better at extracting an ABB task implemented on vowels than on consonants. Here we will test whether this vowel bias is present in repetition-based regularity detection in the newborn brain using Gervain et al. (2008)'s paradigm.

However an advantage for vowels over consonants in newborns' speech processing could also be explained by the acoustic and developmental advantage for vowels over consonants (Pisoni, 1973; Benavides-Varela et al., 2011, 2012; Bertoncini, et al., 1988). Thus, to distinguish between a functionally driven advantage for vowels from a general acoustically-based sensitivity for vowels over consonants, it is necessary to conduct a second experiment assessing the other part of the C/V functional asymmetry hypothesis, i.e. the consonantal bias for lexical processing. Accordingly, in Experiment 3, a lexically-based instead of syntax-like type of processing will be induced. The absence of a vowel bias in lexical processing in Experiment 3 would establish that the vowel bias observed in syntax-like processing is functional. Importantly, Experiment 3is based on the same paradigm and uses the same consonant vs. vowel repetition stimuli as Experiment 2 to allow for a comparison. Gervain and colleagues' paradigm may be readily adapted to fit the requirements of lexical item memorization and thus elicit lexical-based processes by presenting the same word repeatedly within

a block (instead of different tokens sharing a common rule). It is well established that word memorization is present from birth (Swain, et al., 1993; Valiente et al., 2006, 2013) and it has been demonstrated using NIRS that newborns can identify and recognize specific words or word-like items after repeatedly hearing them in short familiarization periods (Benavides-Varela et al., 2011, 2012).

In sum, the paradigm of Gervain et al (2008) is used here in two experiments comparing the processing of three different artificial grammars, one unstructured control grammar (ABC) and two carrying a segmental repetition, yielding two ABB grammars: ABBconsonant and ABBvowel (later on, ABBc and ABBv respectively). Crucially, we also used two different experimental designs to induce structural vs. lexical processing, the former favoring learning over vowels, the latter learning over consonants according to the functional hypothesis of Nespor et al. (2003). To achieve this differential processing, in Experiment 2, the blocks were variable, i.e., made up of 6 different sequences sharing the same pattern. Thus the grammars were presented by type, as in Gervain et al. (2008, 2012) and Experiment 1. In Experiment 3, the blocks were uniform, i.e., made up of 6 repetitions of the same sequence (e.g. *mubabe mubabe, mubabe …*), so grammars were presented by 'token'. If the processing of consonants and vowels is already driven by a functional division of labor, then, given that Experiment 2 and 3 are meant to selectively induce structural processing and item-based memorization respectively, ABBv should provide better input for learning in Experiment 2, while ABBc should be preferentially processed in Experiment 3.

The pattern of results observed in Experiment 1 is specific to our NIRS system, our material and our design, and thus allows specifying the predictions for the outcome of the current experiments at least for Experiment 2. Like Experiment 1, Experiment 2 and 3 consist of 12 consecutive blocks for each condition and crucially the ABBc and ABBv lists of sequences are generated from the ABB list validated in Experiment 1. Crucially the detection of the repetition regularity in Experiment 1 was marked by a left-lateralized fronto-temporal neural advantage for ABB over ABC during the three initial blocks, followed by a constant level of activation throughout the 12 ABB blocks, while ABC evoked a repetition enhancement effect, i.e. an increasing activity between initial and final blocks. If the neural signatures evoked by segmental repetitions are similar to those of syllabic repetitions, we may expect the neural response to ABB grammars in Experiment 2 to look like the response to ABB in Experiment 1. If a vowel bias for rule learning is present at birth as suggested by Nespor et al. (2003), we may expect the neural response to the ABBv grammar in particular to look like the response to ABB in Experiment 1. Making precise predictions for Experiment 3 is more delicate given that an entirely different learning mechanism is expected to be induced by the experimental design (homogeneous/uniform instead of heterogeneous/variable

blocks of stimuli). However, we can assume that a consonant advantage should lead to a differential response to ABBc as opposed to ABBv and ABC.

II. 3. b Experiment 2: Syntax-like context

Materials and Methods

Participants

Eighteen (10 girls and 8 boys) healthy neonates (all Apgar scores 10 minutes after birth: 10) born to French-speaking mothers were included in the analyses. Their age ranged from 1 to 4 days (mean age: 2.2 days). They had a mean head circumference of 33.9 cm. An additional 32 infants were tested but their data were excluded due to the infant becoming awake or fussy and failing to complete the procedure (14), equipment failure (2), parental intervention (2) or insufficient analyzable data (14).

Data were considered as insufficient when more than 50% of the channels had less than 50% non-rejected blocks. All infants were tested while asleep. All parents of infants gave informed written consent prior to beginning the experiment. The CERES (Comité d'Evaluation des Projets de Recherche en Santé) has approved this study under the approval nr. 2011-13-2.

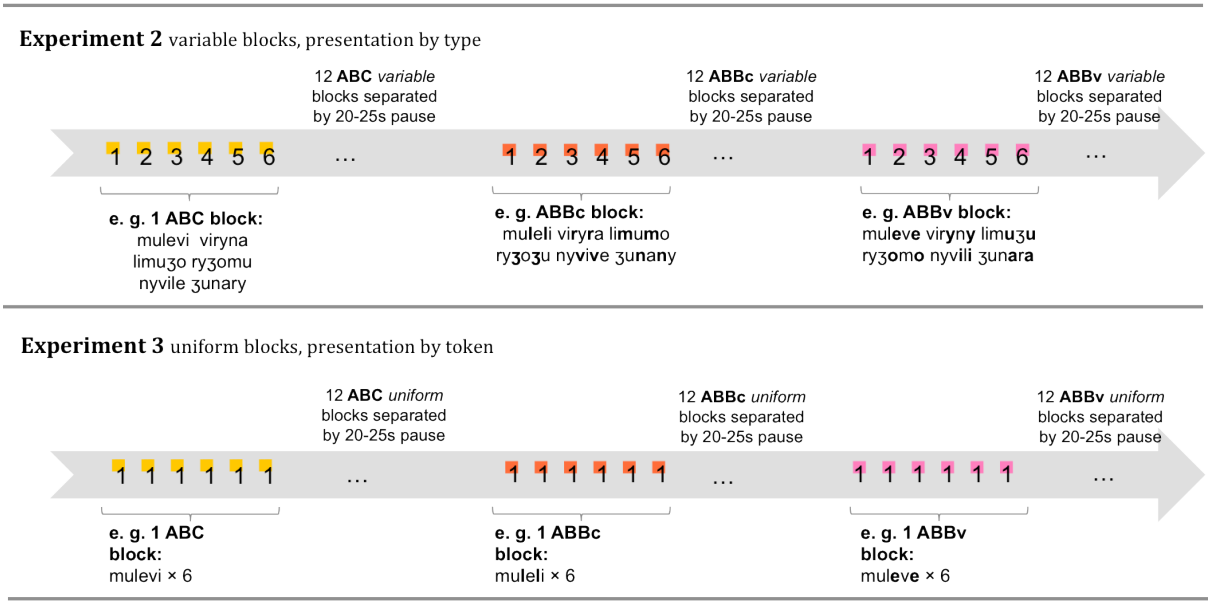


Figure II. 11. *The block designs used in Experiment 2 (upper panel) and Experiment 3 (lower panel).*

Material

Participants were exposed to the same random ABC control grammar as in Experiment 1 (described in section *I.3.b.)* (e.g. *mulevi, junary*) and to two repetition-based ABB artificial grammars derived from the ABB grammar in Exp 1 : ABBc with repetitions of consonants, but not of vowels (e.g. *muleli,*

junany), and ABBv with repetitions of vowels, but not of consonants (e.g. *muleve, junara).* Using the same 12-syllable repertoire, the two ABBc and ABBv grammars generated trisyllabic sequences and were matched on all nonstructural properties: syllabic repertoire; frequency of the A, B, and C syllables in each position; phonological characteristics; flat prosody; and transitional probabilities between syllables. Sequences were synthetized using the fr4 female voice of the MBROLA diphone database (Dutoit et al., 1997), in a monotonous pitch of 200 Hz, with the same length for all phonemes (150ms). The optical topography imager and configuration of probe set was the same as in Experiment 1. In Experiment 2, the sequences are presented by type, i.e. the 6 sequences composing a block are all different and only share the same structure: ABBc, ABBv or ABC (***Fig. II. 11.*** upper panel).

Procedure

Neonates were tested in a local maternity hospital, in a silent experimental room, while asleep in a bassinet. Each infant was presented with 36 blocks of stimuli. Exactly as in Experiment 1, each block consisted of six CVCVCV sequences and lasted 9.4 sec. Within a block, the six sequences were separated by brief pauses of variable length (0.5 – 1.5 s); blocks were also spaced at time intervals of varying duration (20 –25 s) to avoid inducing phase-locked brain responses. The total testing time for each infant was 21.9 min. The block designs used in Experiment 2 and in Experiment 3 are presented in **Figure *II. 11.*** All participants in both experiments were presented with all conditions, ABBc, ABBv and ABC. Condition order was randomized and counter-balanced across participants.

Data analysis

Same as in Experiment 1.

Results

The grand average results for Experiment 2 are presented in ***Figure II. 12***., with the oxyHb and deoxyHb concentration changes averaged across all blocks of each condition and across all infants.

The channel-by-channel t-test results for the mean oxyHb and deoxyHb changes are shown in ***Figure II. 13,*** for the following comparisons: the average of the three conditions vs. baseline (A), each condition vs. baseline (B-D), conditions to each other (E-G). For ease of exposition, ***Figure II. 14.*** (upper panel) plots all channels for which significant t-test comparisons were found for oxyHb or deoxyHb and the direction of their effect. After correction for multiple comparisons using the False Discovery Rate (Benjamini & Yekutieli, 2001), none of the t-test remains significant, except a greater deoxyHb response for ABC over the baseline in channel 12.

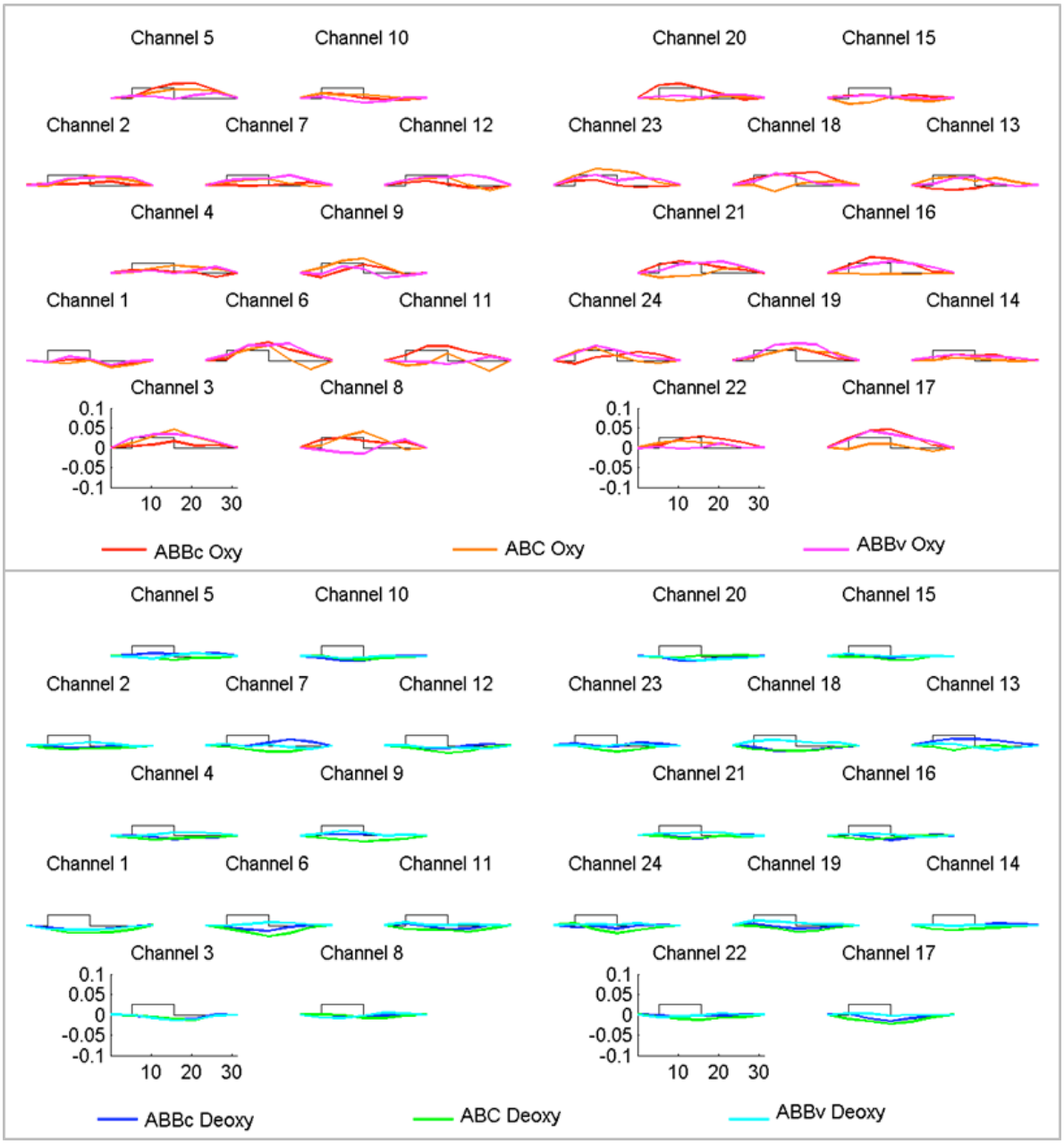
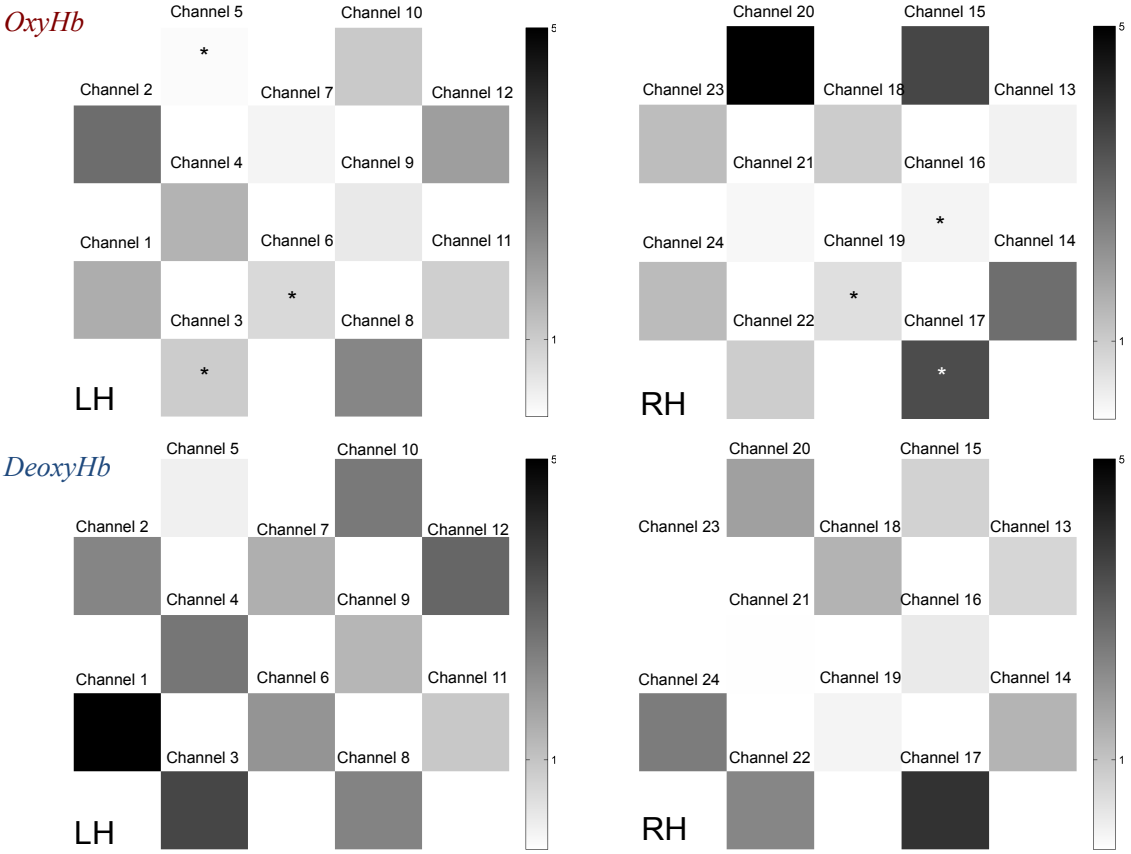
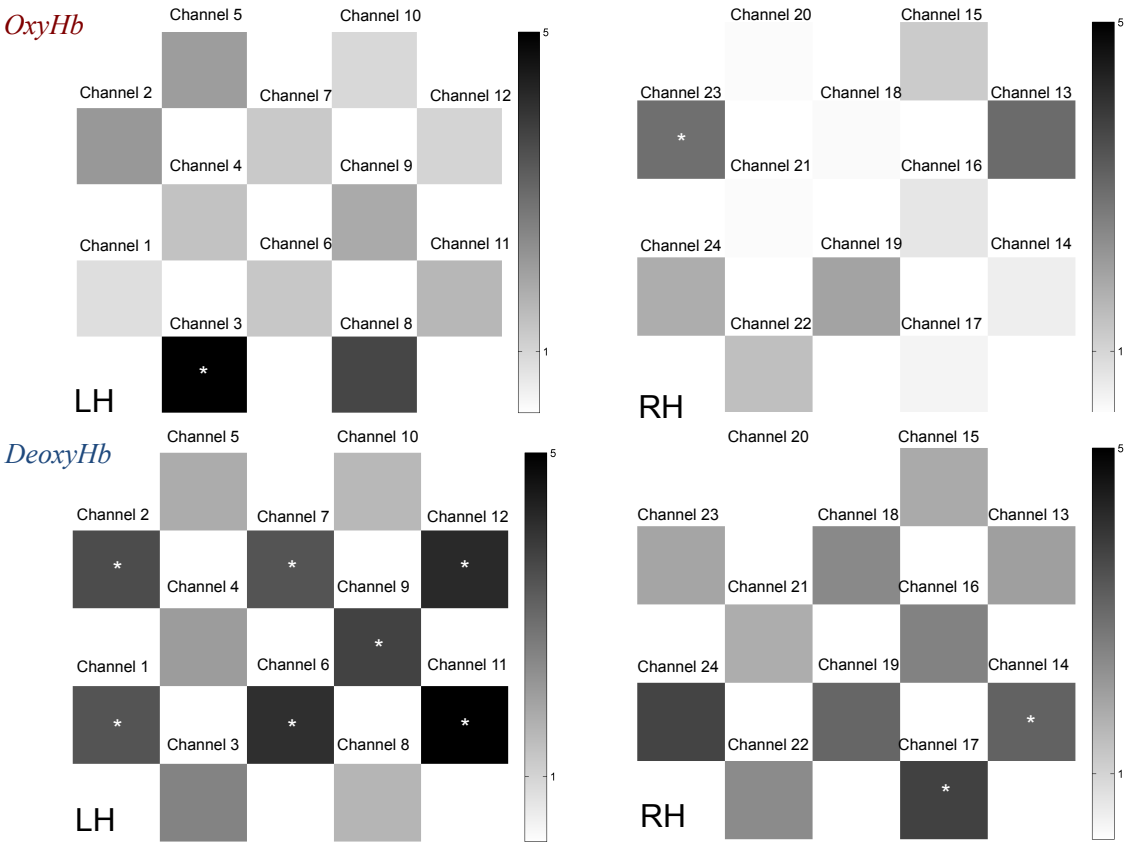


Figure II. 12. *Experiment 2. Grand average results of Experiment 2. Numbers and location of channels correspond to the placement shown in Figure II. 7. A. The x-axis represents time in seconds; the y-axis shows concentration change in mmol x mm. The rectangle along the x-axis indicates time of stimulation. The upper panel represents oxyHb concentrations, and the lower panel deoxyHb concentrations in response to the 3 grammars. The orange and green lines represent oxyHb and deoxyHb concentrations, respectively, in response to the ABC grammar. The red and navy blue lines in the graphs represent oxyHb and deoxyHb concentrations, respectively, in response to the ABBc grammar. The magenta and cyan lines represent oxyHb and deoxyHb concentrations, respectively, in response to the ABBv grammar.*

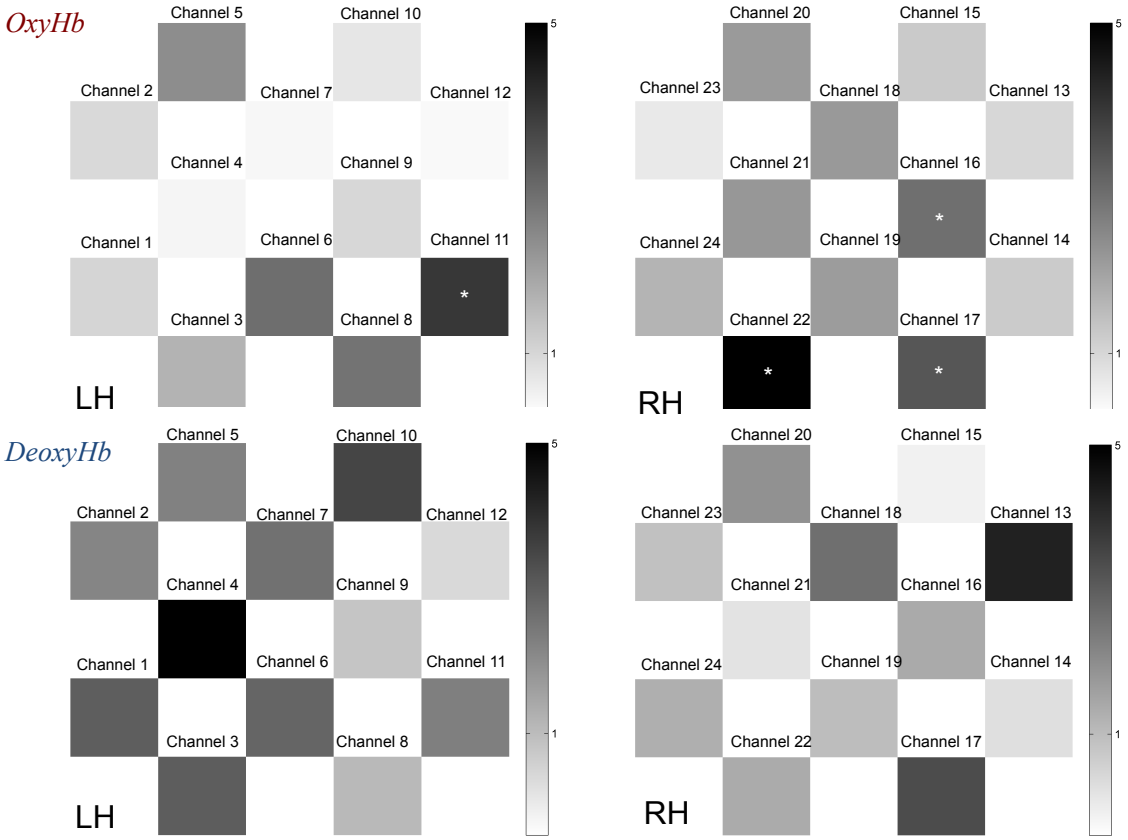
A. Mean activity to baseline



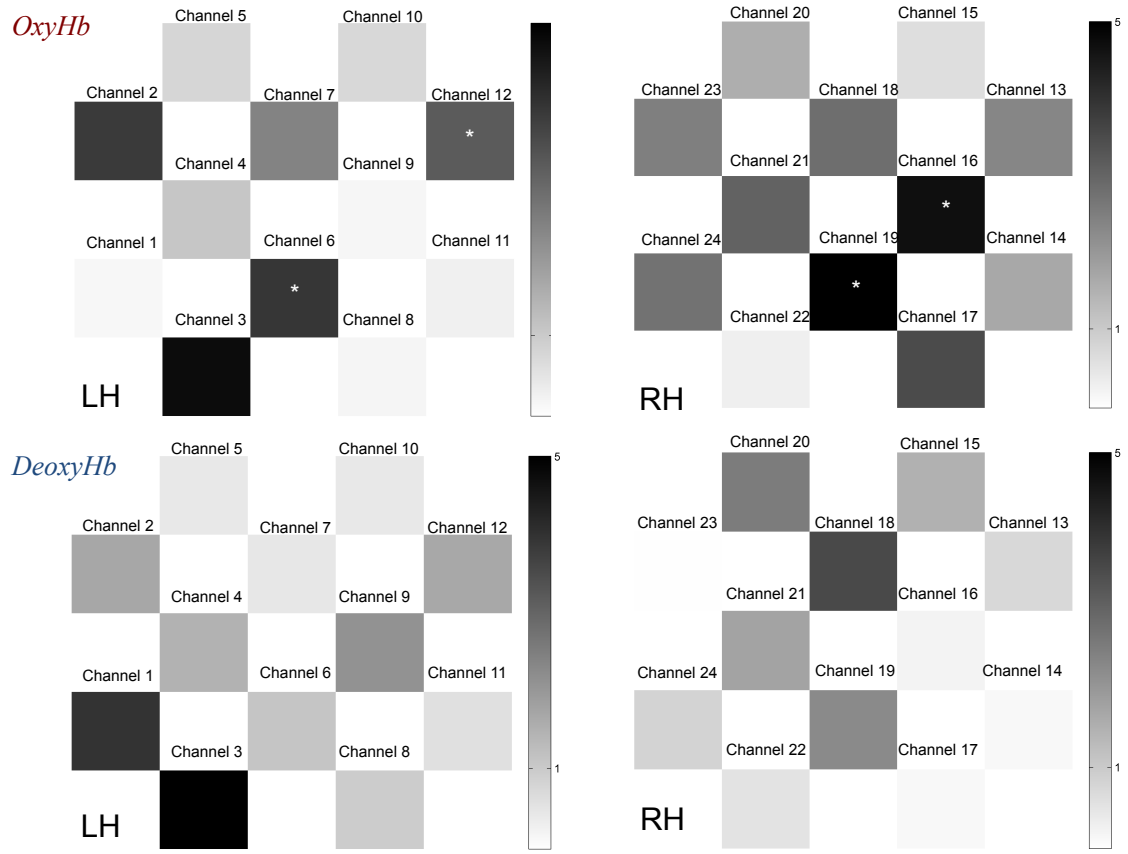
B. ABC vs. baseline



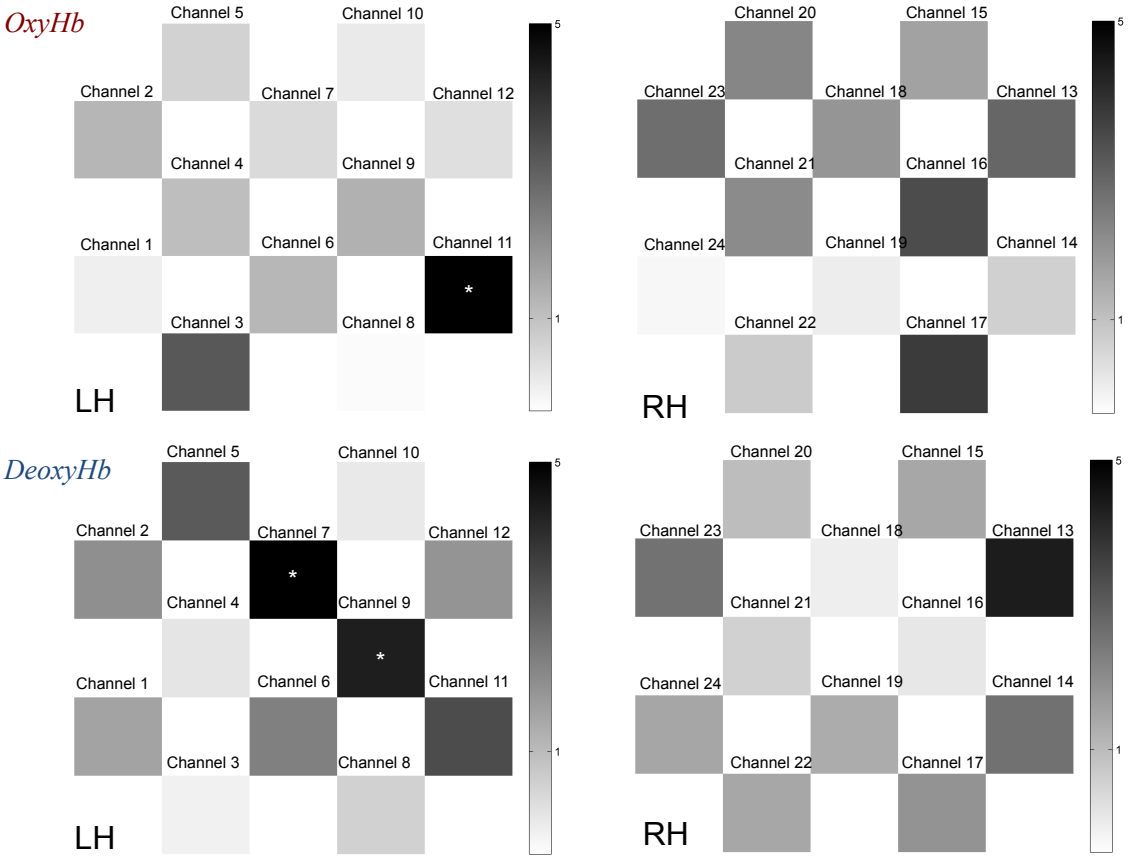
C. ABBc vs. baseline



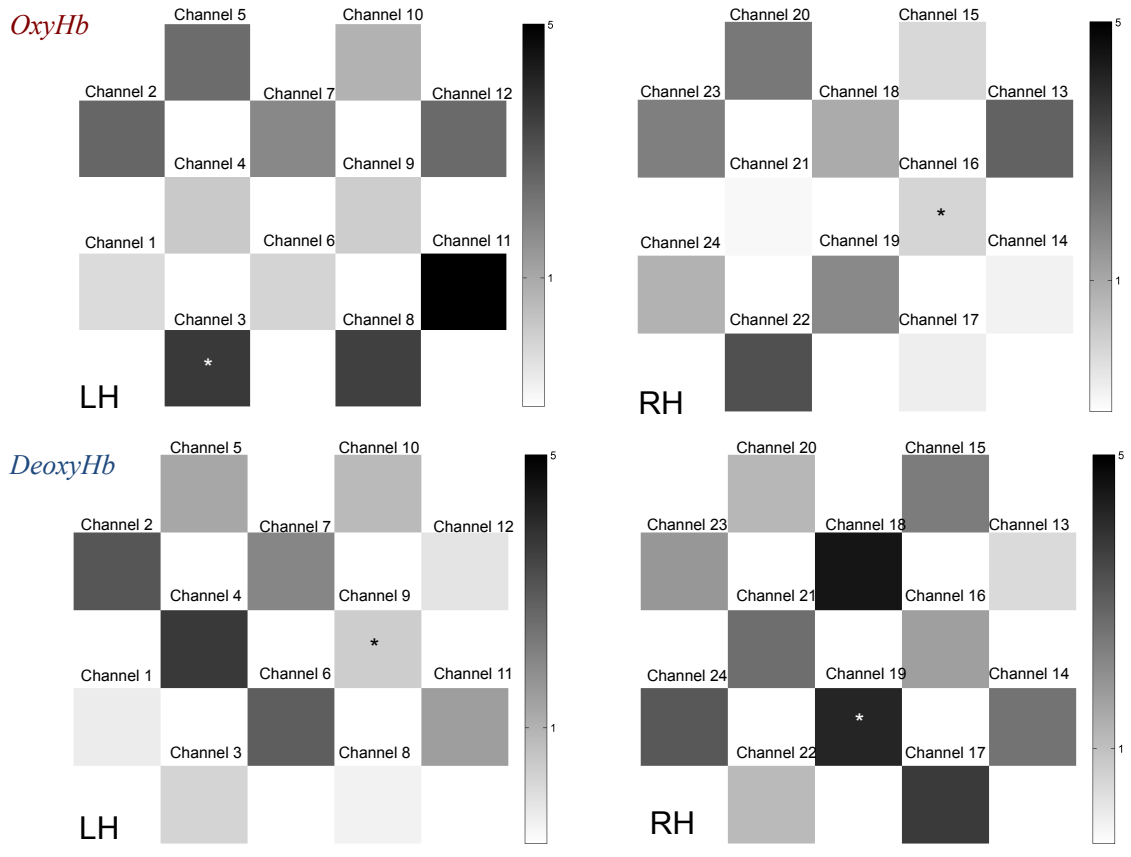
D. ABBv vs. baseline



E. ABBc vs. ABC



F. ABBv vs. ABC



G. ABBc vs. ABBv.

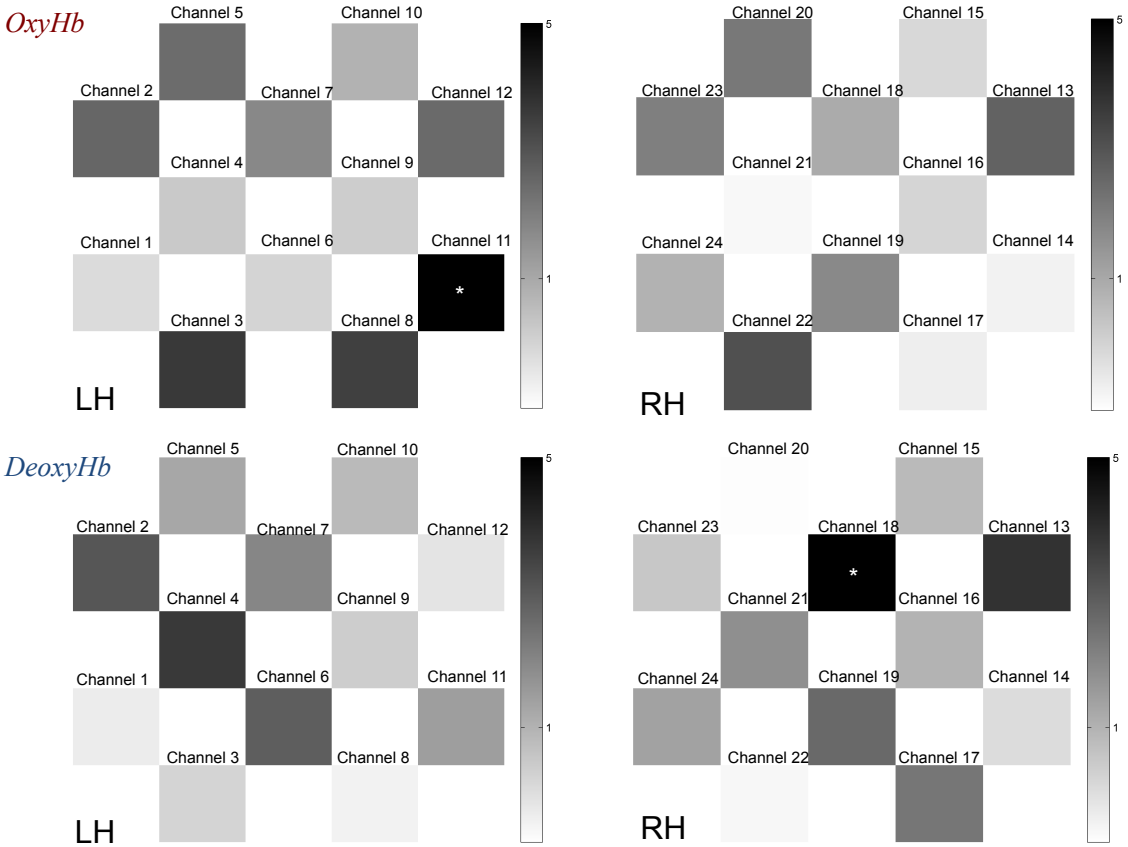


Figure II. 13*. Experiment 2. Statistical maps (t-maps) comparing oxyHb and deoxyHb responses to: (A) ABBc + ABBv +ABC vs. baseline, (B) ABC vs. baseline, (C) ABBc vs. baseline, (D) ABBv vs. baseline, (E) ABBc vs. ABC, (F) ABBv vs. ABC; (G) ABBc vs. ABBv. Channels are plotted following the same placement as in Figure II. 7. A. The t-values for each channel are color-coded as indicated on the color bar. Significance levels are indicated for each channel by p-values: '*' p < .05 significant without correction.*

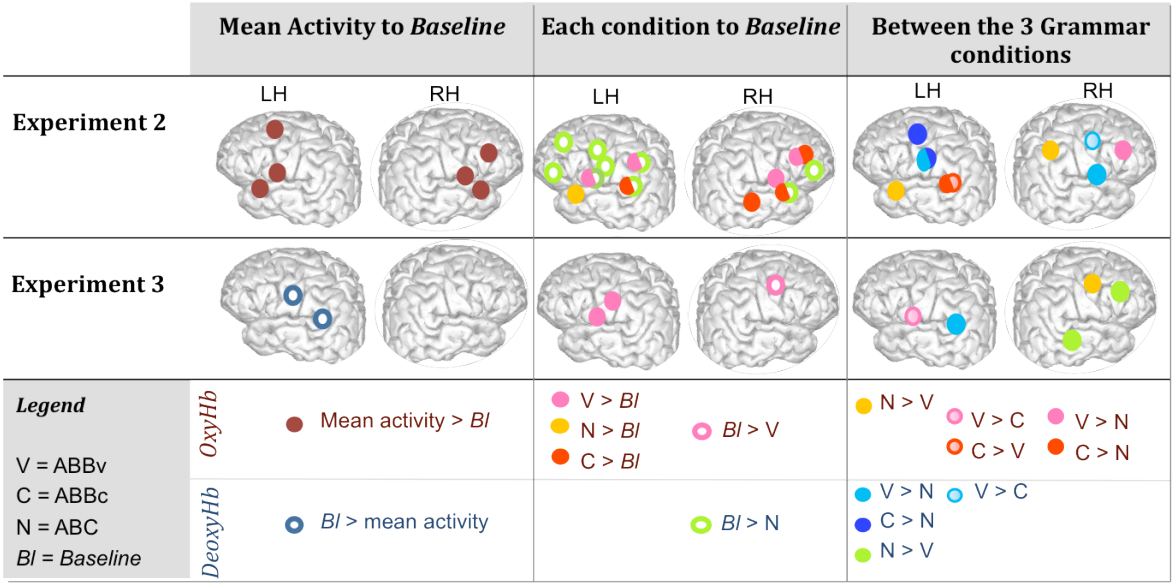


Figure II. 14. *Experiments 2 and 3. Schematic representation of a newborn brain showing the pattern of neural activity elicited in Experiment 2 (upper panel) and Experiment 3 (lower panel) as evidenced by significant channel-by-channel t-tests comparisons. Each dot represents the outcome of a significant t-test on a given channel. T-tests comparing: averaged activity for all 3 conditions to the baseline are shown in the left column; activity elicited by each condition (V, C or N) to the baseline in the middle column; and activity elicited by each grammars to each other in the right column. The dots are color-coded as indicated in the legend below the panels, for oxyHb concentration changes in the first row and deoxyHb concentration changes in the second row.*

The channel-by-channel paired t-tests of Experiment 2 (uncorrected, ***Fig. II. 13***. and ***Fig. II. 14.,*** upper panel) reveal some significant comparisons, but fail to elicit a clear pattern of activity in a specific area to the advantage of any of the three grammars. Note however that most of these significant comparisons correspond to an increase in activity for the ABB grammars, while the ABC condition yields decreases in deoxyHb concentration (see ***Fig. II. 13.*** *B.* DeoxyHb; and ***Fig. II. 14***).

The ANOVA for oxyHb yielded a marginal effect of ROI [$F(1,17) = 2.98$, $p = .086$] as the activity in temporal ROIs tended to be stronger than in frontal ROIs. The ANOVA for deoxyHb yielded a significant effect of Grammar [$F(2,16) = 7.43$, $p = .0008$], as both the activity for ABBv and ABBc was stronger than for ABC (Scheffe's post-hoc test: ABC < ABBv, $p = .0001$; ABC < ABBc, $p = .049$). All other effects and interactions failed to reach significance.

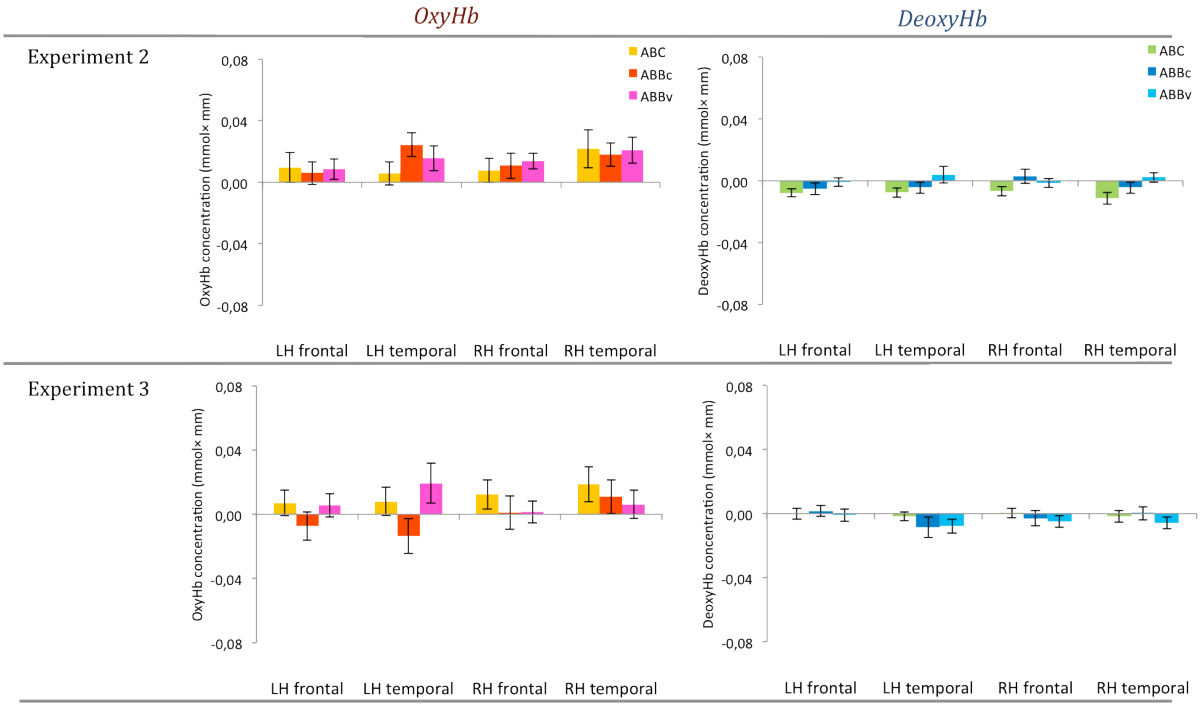


Figure II. 15. *Experiments 2 and 3. The averages of the responses for the three grammars in the frontal vs. temporal ROI in the left (LH) and right hemispheres (RH) in Experiment 2 (upper panel) and Experiment 3 (lower panel). The y-axes show the average Hb concentration in mmol x mm: oxyHb on the left graphs, and deoxyHb on the right graphs. Error bars represent standard errors. (Channels in LH Frontal ROI = [1,2,4,5]; in LH temporal ROI=[3,6,8,11]; in RH temporal ROI = [17,19,22,24]; and in RH Frontal ROI = [14,13,16,15]).*

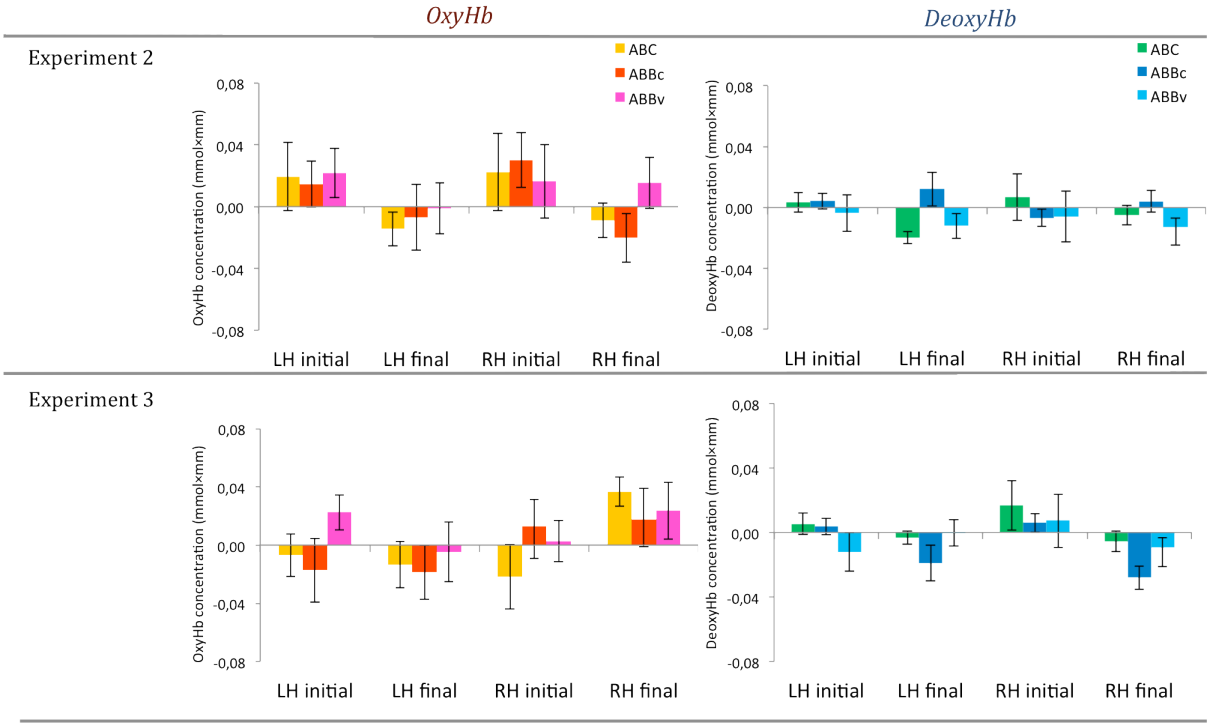


Figure II. 16. *Experiments 2 and 3. The averages of the responses for the three grammars in the left fronto-temporal ROI (ch. [1, 3, 8]) vs. the right fronto-temporal ROI (ch. [14, 17, 22]), for the three first blocks ("initial") and the three last blocks ("final "). The y-axes show the average Hb concentration in mmol x mm: oxyHb on the left graphs, and deoxyHb on the right graphs. Error bars represent standard errors.*

As in Experiment 1, a second ANOVA was run to examine the temporal dynamics of the responses to the ABBc, ABBv and ABC grammars during the course of the 12 consecutive blocks. As in Experiment 1, a single fronto-temporal ROI per hemisphere was used: channels [1, 3, 8] in the LH, and channels [14, 17, 22] in the RH. This ANOVA was conducted on average Hb concentrations in a time window between the 0-35s after stimulus onset, in the bilateral target ROI with the within-subject factors Grammar (ABBc/ABBv/ABC), Time (3 initial blocks/ 3 final blocks), and Hemisphere (LH/RH) (***Fig. II. 16,*** upper panel). This analysis explores whether, as in Experiment 1, the time course of neural responses differs for the three grammars in the fronto-temporal ROI, with a difference response for ABBv as predicted by Nespor et al. (2003).

The ANOVA for oxyHb yielded a significant main effect for Time [$F(1, 17) = 5.07$, $p = .026$], as initial blocks gave rise to a larger response than final blocks. Similarly, the ANOVA on deoxyHb yielded a significant main effect for Time [$F(1, 17) = 6.42$, $p = .012$], as initial blocks gave rise to a larger response than final blocks. All other effects and interactions failed to reach significance.

Discussion

Experiment 2 had the same design and similar material as Experiment 1, but crucially opposed the ABC unstructured control grammar to two ABB grammars (instead of one): ABBc and ABBv. First, a stronger activity is evoked in temporal over frontal areas bilaterally, showing that the stimuli are processed by the auditory areas. Second, oxyHb and deoxyHb concentrations decrease significantly from initial to final blocks in bilateral fronto-temporal ROIs, suggesting an attenuation of the neural response over the time-course of the experiment. Finally, and most importantly, both ABB grammars seem to elicit larger neural responses than the ABC unstructured condition. Therefore, like syllabic repetitions in Experiment 1, segmental repetitions are differentially processed from ABC sequences, and both consonant and vowel repetitions-based regularities are detected in the newborn brain. Experiment 2 thus fails to find a vowel bias for syntax-like processing at birth, as predicted by the innate C/V functional asymmetry. It also goes against the general advantage for vowel over consonant information in early speech processing. Experiment 3 will present the same grammars and stimuli, rearranged in a new design to induce an item-based, more lexical-like mode of processing in the newborn brain, and thus explore the origins of the consonant bias at birth.

II. 3. c. Experiment 3: Lexical context

Materials and Methods

Participants

Eighteen (9 girls and 9 boys) healthy neonates (all Apgar scores 10 minutes after birth: 10) born to French-speaking mothers were included in the analyses. Their age ranged from 1 to 4 days (mean age: 2.1 days). They had a mean head circumference of 33.8 cm. An additional 10 infants were tested in Experiment 3, but their data were excluded due to the infant becoming awake or fussy and failing to complete the procedure (4), equipment failure (4), or insufficient analyzable data (2).

Material

Same as in Experiment 2, except that in Experiment 3, the sequences were presented by token, i.e. the 6 sequences composing a block were identical (***Fig. II. 11***., lower panel).

Procedure

Same as in Experiment 2. The block design used in Experiment 3 is presented in **Figure *II. 11*.** (Lower panel)**.**

Data analysis

Same as in Experiment 1 and 2.

Results

The grand average results for Experiment 3 are presented in ***Figure II. 17***. The figure shows the oxyHb and deoxyHb concentration changes averaged across all blocks of each condition and across all infants.

The channel-by-channel t-test results for the mean oxyHb and deoxyHb changes are shown in ***Figure II. 18.*** for the following comparisons: the average of the three conditions vs. baseline (A), each condition vs. baseline (B-D), conditions to each other (E-G). For ease of exposition, ***Figure II. 14.*** (Lower panel) plots all channels for which significant t-test comparisons were found for oxyHb or deoxyHb. After correction for multiple comparisons using the False Discovery Rate (Benjamini & Yekutieli, 2001), none of the t-test remains significant.

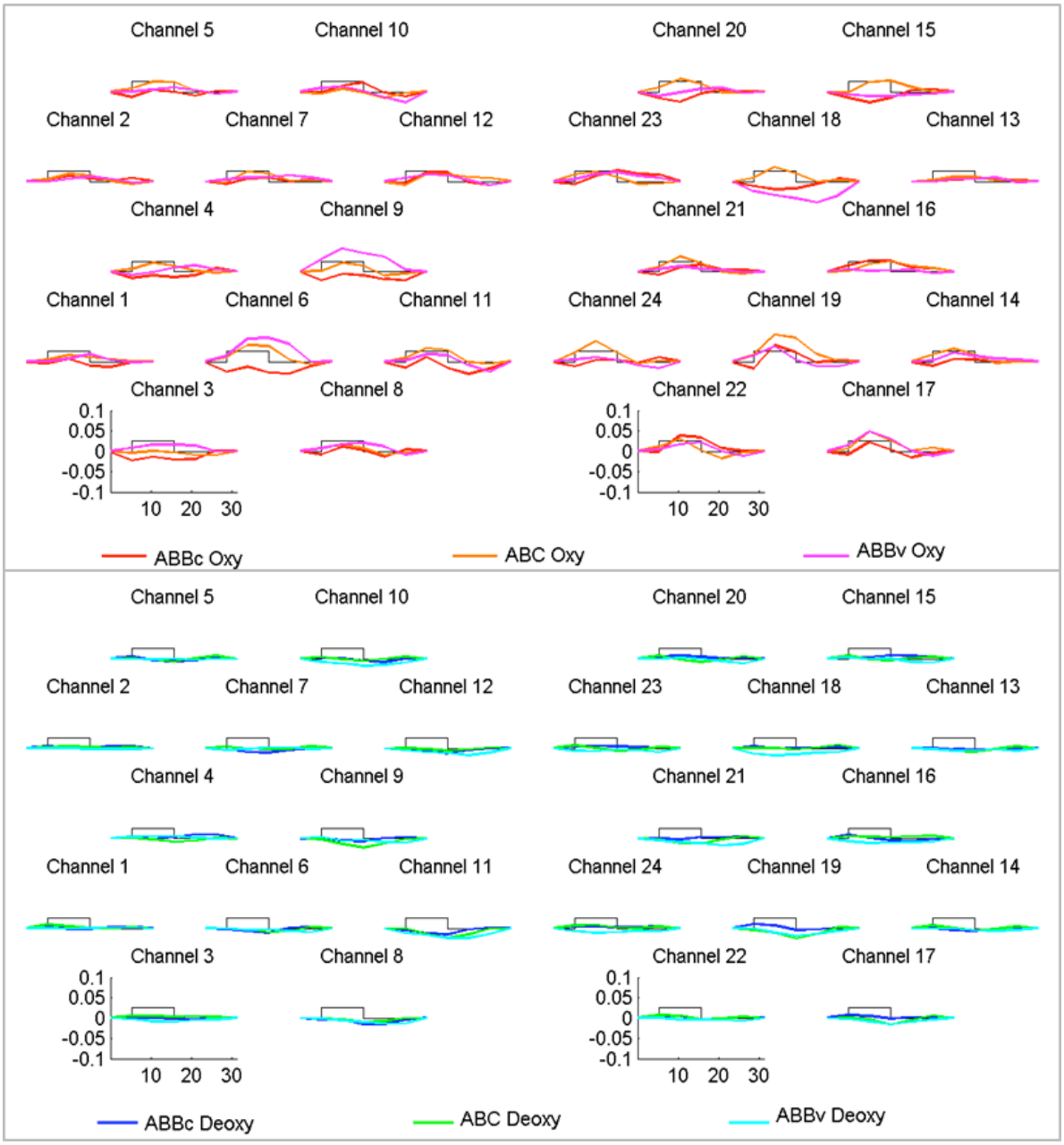
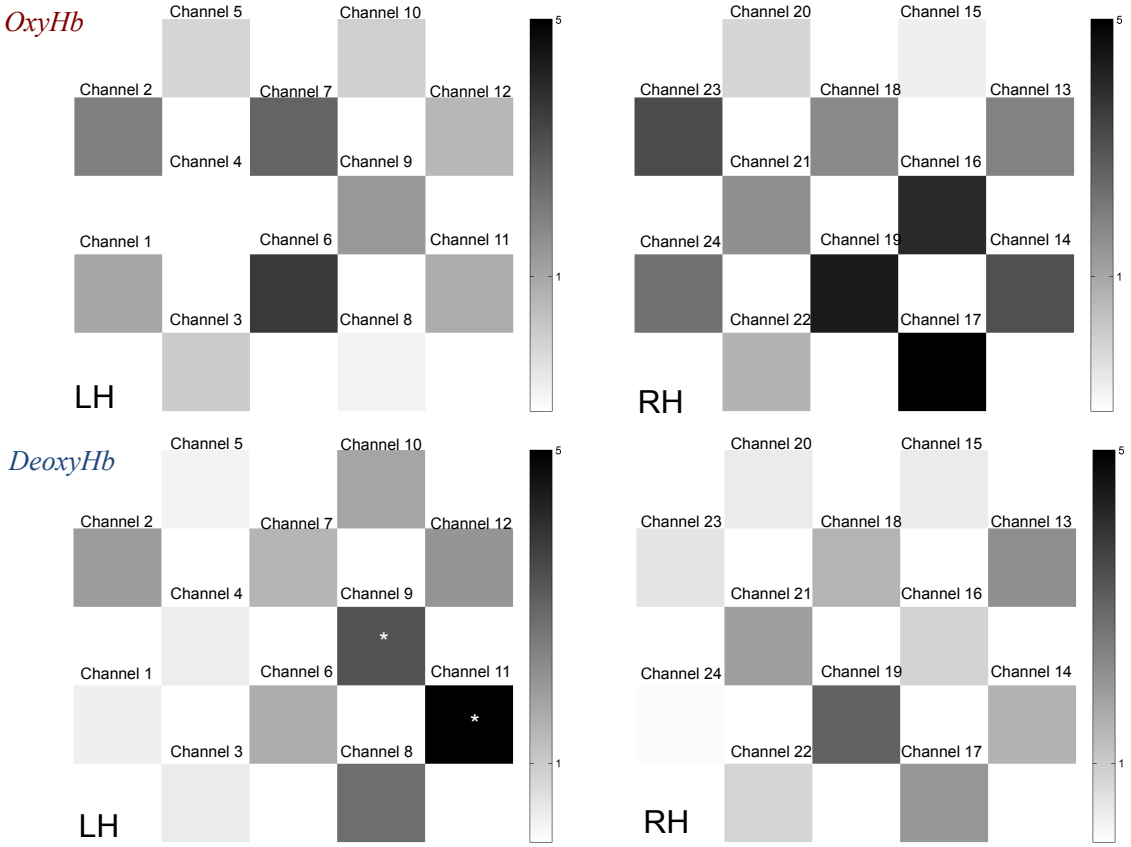
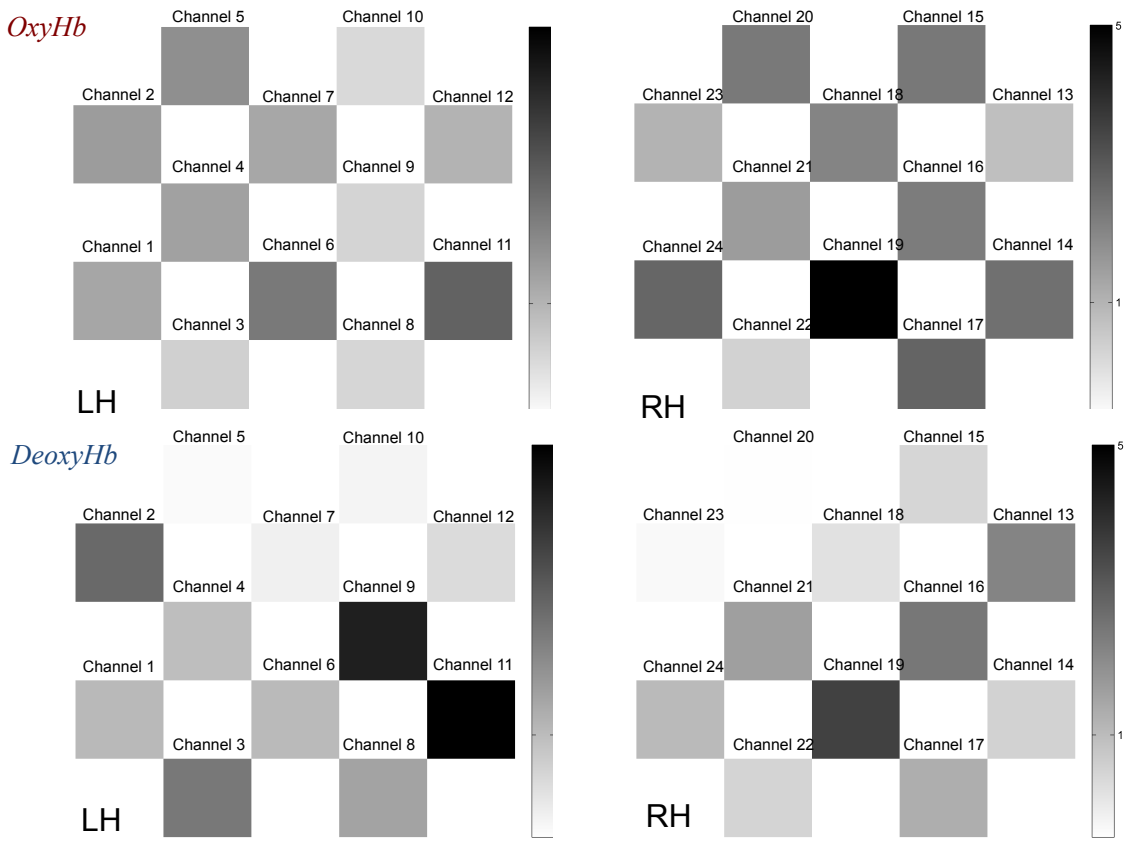


Figure II. 17. *Experiment 3. Grand average results. Same plotting conventions as in Figure II. 12.*

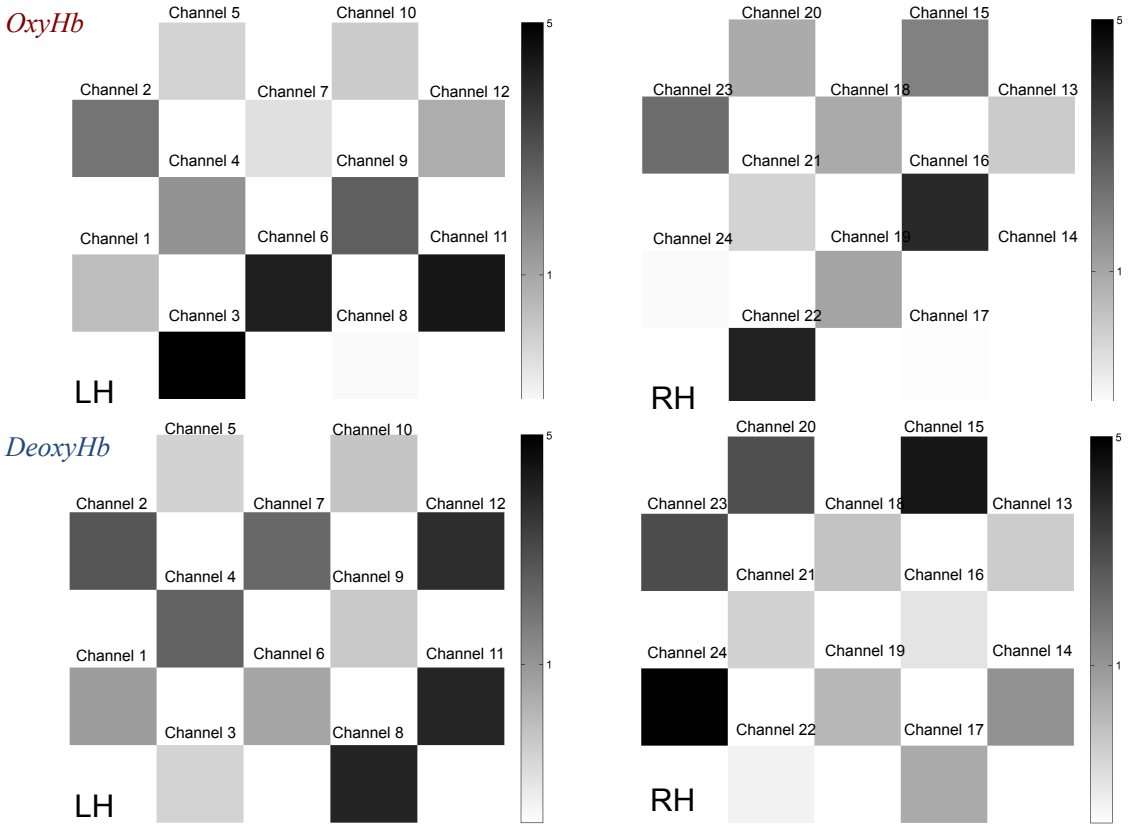
A. Mean activity to baseline



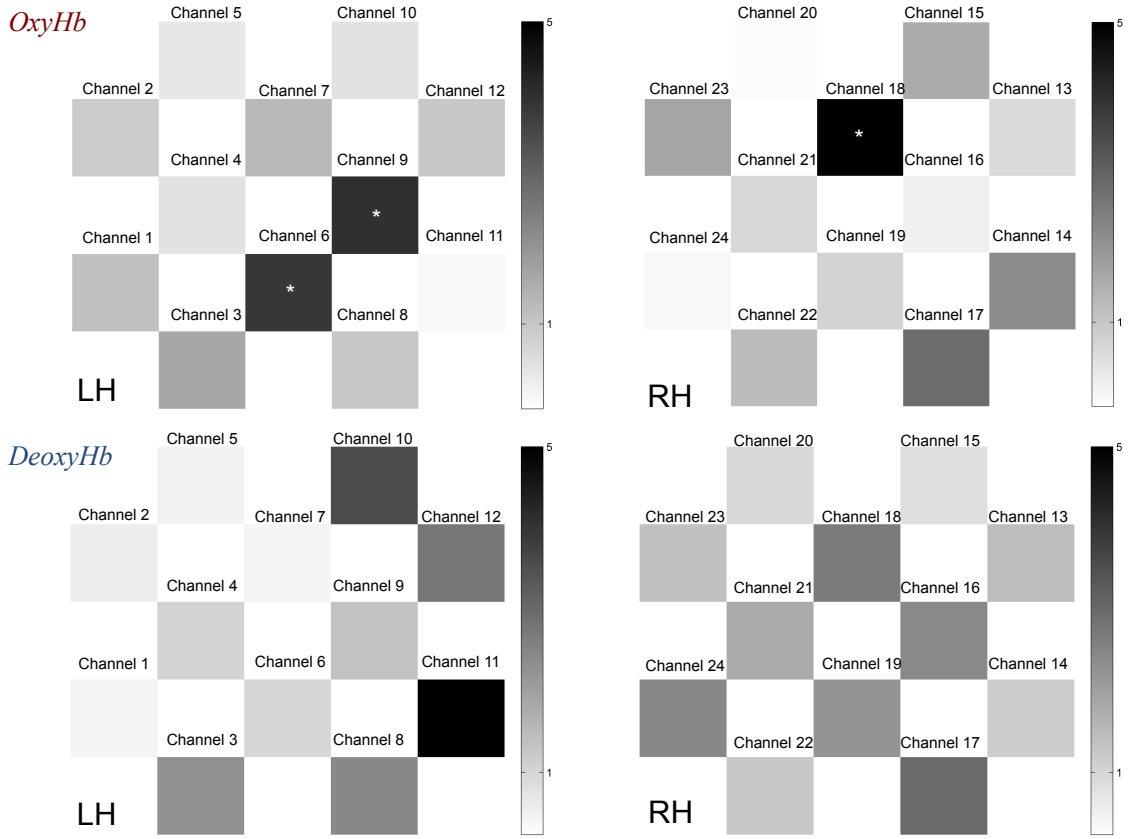
B. ABC vs. baseline



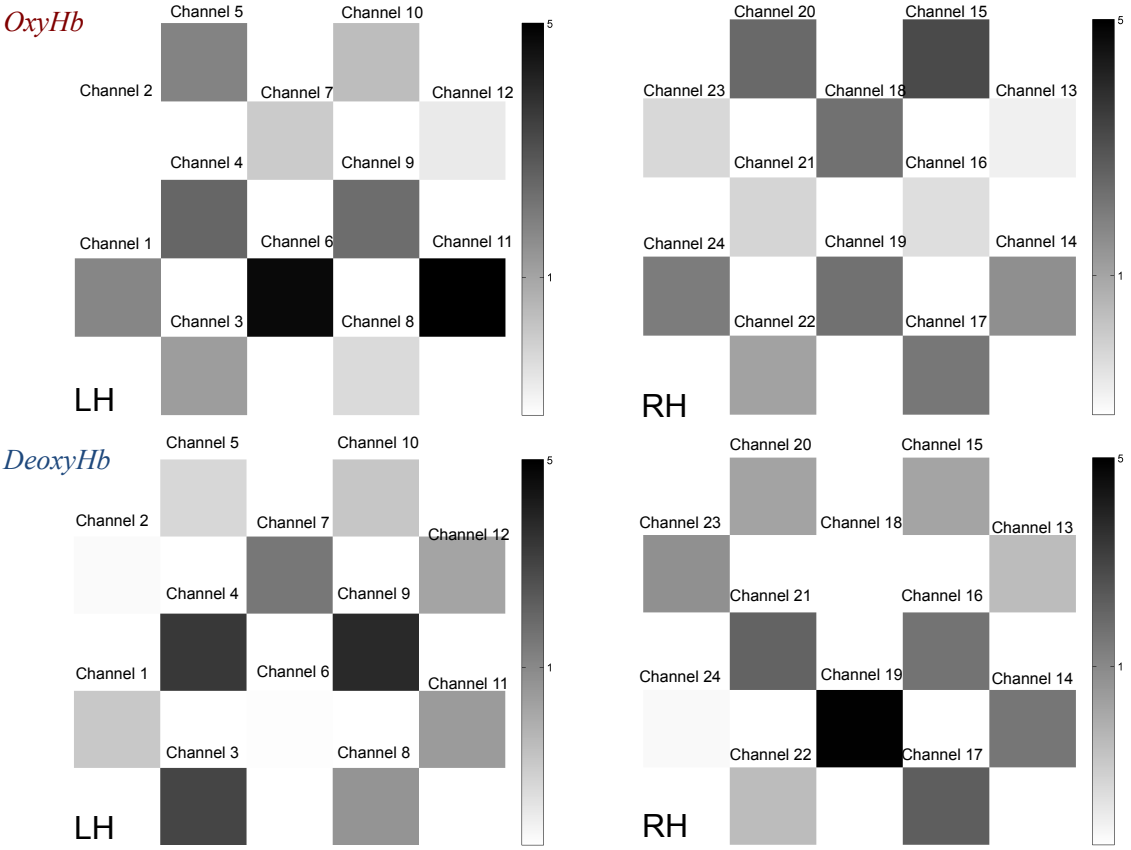
C. ABBc vs. baseline



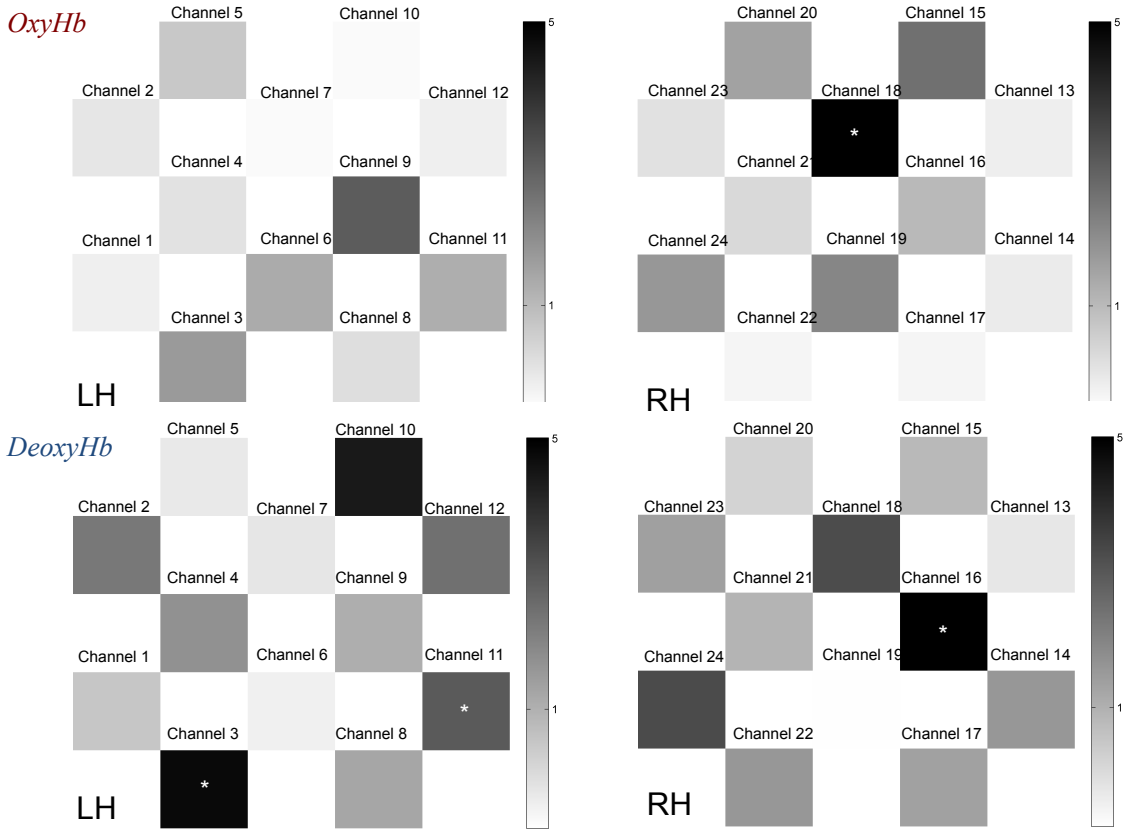
D. ABBv vs. baseline



E. ABBc vs. ABC



F. ABBv vs. ABC



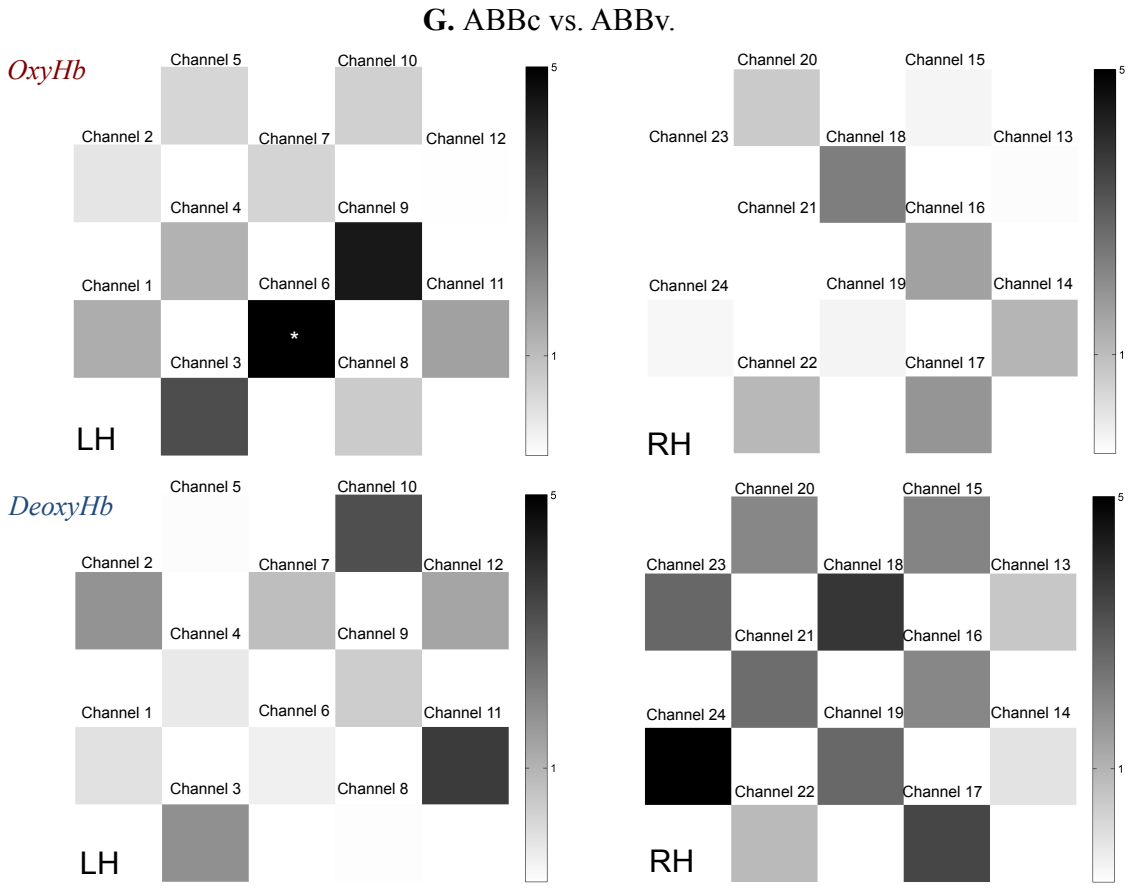


Figure II. 18. *Experiment 3 Statistical maps (t-maps) comparing oxyHb and deoxyHb responses to: (A) ABBc + ABBv +ABC vs. baseline, (.B) ABC vs. baseline, (C) ABBc vs. baseline, (D) ABBv vs. baseline, (E) ABBc vs. ABC, (F) ABBv vs. ABC; (G) ABBc vs. ABBv. Same plotting conventions as in Figure II. 13.*

The results of channel-by-channel paired t-tests of Experiment 3 yield very few significant comparisons (***Fig. II. 18.*** *and* ***Fig II. 14.*** lower panel), reflecting the very low average neural activity evoked in this experiment.

As in Experiment 1 and 2, a first ANOVA was run with the within-subject factors of Grammar (ABBc/ABBv/ABC), Hemisphere (Left/Right) and ROI (frontal/temporal) on average Hb concentration, in a time window between 0-35s after stimulus onset. This analysis evaluated whether the ABB grammars were processed differently as in Gervain et al. (2008), Experiment 1 and Experiment 2, and in addition whether the ABBc grammar was processed differently as predicted by Nespor et al. (2003). Separate ANOVAs were conducted for oxyHb and deoxyHb. They yielded no significant effect or interaction (cf. ***Fig. II. 15.*** *Lower panel)*.

As in Experiment 1 and 2, a second ANOVA was run to examine the temporal dynamics of the responses to the ABBc, ABBv and ABC grammars during the course of the 12 consecutive blocks. This ANOVA was conducted on average Hb concentrations in a time window between 0-35s after stimulus

onset, in the target fronto-temporal ROI (channels [1,3,8] in the LH, and channels [14,17,22] in the RH) with the within-subject factors Grammar (ABBc/ABBv/ABC), Time (3 initial blocks/ 3 final blocks), and Hemisphere (LH/RH). This analysis explores whether, as in Experiment 1, the time course of neural responses differentiated the three grammars in the fronto-temporal ROI. Separate ANOVAs were conducted for oxyHb and deoxyHb. They yielded no significant effect or interaction (cf. ***Fig. II. 16***. *Lower panel*).

Discussion

Experiment 3 used the same material as Experiment 2, but presented the stimuli by token, a design supposed to induce lexical item memorization. Very few channels gave rise to significant t-test results and the ANOVAs did not yield any significant effect or interaction. Therefore the overall neural activity evoked in Experiment 3 is weak, which may be attributed to the repetition of the same token six times in each block. This repetitive presentation might have led to a strong habituation/attenuation of the response, so no differences between consonants and vowels could be observed. This effect of context of presentation is congruent with our discussion of the differences in results between Experiment 1 and Gervain et al. (2008) in terms of stimulus variability.

II. 4. General discussion

The goal of the present experimental chapter was to investigate the role of consonants versus vowels in speech processing at birth. Our experiments followed up on Gervain et al.'s original paradigm (2008), which first demonstrated newborns' ability to detect and learn repetition-based regularities over syllables (Gervain et al., 2008, 2012). Experiment 1 first tested a new material better suited to address the question of the C/V functional asymmetry. Using a less variable material, Experiment 1 successfully confirmed the specific detection of a repetition rule over syllables in the newborn brain, although eliciting a different time course pattern in response to ABC and ABB patterns than found in Gervain et al. (2008). Experiment 2 and 3 then directly addressed the question of the C/V asymmetry at birth. In these experiments, infants were exposed to the same ABC unstructured sequences as in Experiment 1, and to two repetition-based regularities: in ABBc sequences the consonantal tier carried a consonant repetition while the vocalic tier did not (e.g. *lumumo)*, and vice versa for ABBv sequences (e.g. *limuju*). The lists of ABBc and ABBv sequences were derived from the material in Experiment 1 (e.g. ABC *limujo* → ABBc *limumo +* ABBv *limuju)*. In Experiment 2, as in previous experiments by Gervain and colleagues and as in Experiment 1, the

sequences were presented in variable blocks in order to favor the detection of the *type* represented within blocks. In Experiment 3 the sequences were presented in uniform blocks of 6 repetitions of the same *token* in order to induce memorization of lexical items (Endress & Bonatti, 2005). According to the division of labor hypothesis proposed by Nespor et al. (2003) vowels should be better targets than consonants for rule-learning processes induced in Experiment 2, while sequences carrying a consonant repetition should provide better input for lexical learning in Experiment 3.

The first finding of the three experiments is related to the effect of input variability on infants' brain responses. Taken together, the three experiments show that a gradual decrease in stimulus variability leads to an increasingly attenuated NIRS response. Indeed, in Experiment 1 and 2 using a more redundant material than Gervain et al. (2008), the evoked neural activity was reduced compare to Gervain et al. (2008), where stimulus variability was maximal (each sequence was presented only once for a total of 280 different sequences). Activity was even further reduced in Experiment 3, where the stimuli were presented in uniform instead of variable blocks. Comparing Experiment 1 and 2, the attenuation of the cortical response can be attributed to the fact that Experiment 2 had one additional grammar condition. Consequently, it lasted a third longer than Experiment 1 (21 vs. 14 minutes respectively), and had considerably more redundant stimuli (e.g. the pair of sequences *nyvile / nyvivi* in Experiment 1 would yield the triplet *nyvile / nyvive / nyvile* in Experiment 2 & 3). The decrease in evoked brain activity from Experiment 2 to Experiment 3 can be attributed to the additional effect of token presentation of 6 identical items in uniform blocks, which were expected to induce rote memorization and resulted in a considerable increase of the predictability of the input within each block. Taken together, our findings highlight the crucial role of variability in modulating neural responses to the stimuli.

The second finding is that the newborn brain was found to detect segment-level repetitions. Indeed, in Experiment 2, both ABBc and ABBv patterns elicited an increase in activity as opposed to the ABC condition, and although this distinction did not give rise to lasting effects during the time-course of the experiment, ABBc and ABBv appear to be relevant targets for repetition-based regularities. These findings thus extend the existing evidence for repetition-based detection found at the syllabic level to the segmental level (Gervain et al., 2008, 2012).

The third and most important finding is related to the origins of the C/V functional asymmetry at birth. Indeed, the absence of a differential pattern of activation in Experiment 2 between consonants and vowels at birth goes against the proposal of an initial vowel bias in syntax-related processing (Nespor et al., 2003) and differs from other studies suggesting a possible general acoustically-driven vowel bias in early speech processing. Nevertheless, the detection of segmental repetitions (Experiment 2 and 3) appears weaker than that of syllabic repetitions (Experiment 1). Given this weaker effect, a methodological caveat is in order: it is possible that our sample size may

be too small to elicit the predicted differential response between consonants and vowels. Testing additional infants would be a way to address this concern. Lastly in Experiment 3, no effect of the type of grammar was found. Again testing more infants would be important to make sure this lack of effect is not due to a statistical power issue. If confirmed, it would go against the proposal that the consonant bias in lexical-like processing is present from birth, as proposed by Nespor et al. (2003).

In summary, in this chapter on the origin of the consonant and vowel asymmetry at birth, the three fNIRS experiments conducted fail to reveal that consonant- and vowel-based repetition regularities are differently detected in the newborn brain. Experiment 2 and 3 do not confirm the strong assumptions that consonants serve lexical-related processes while vowels are better targets for repetition detection from birth. This paradigm thus fails to show an innate functional distinction between consonants and vowels at birth. But how early can such biases be observed in development? So far, the earliest evidence of a lexically-related consonant bias and a syntactically-related vowel bias was found at 11 months of age (Poltrock & Nazzi, in revision). In the next experimental chapter, we will investigate whether the consonant bias in lexical processing can be observed in the first semester of life.

EXPERIMENTAL CHAPTER 3 – Investigation of the Consonant/Vowel asymmetry in early lexical processing, measuring reactions to mispronunciations in the infant's own name

In most languages consonants carry more lexical information than vowels do, and Nespor, Peña, & Mehler (2003) proposed that infants would therefore be innately and universally biased towards the use of consonants over vowels when processing lexical information in speech, thus facilitating their lexical acquisition. So far, cross-linguistic investigations converge to the conclusion that infants in their second year of life use this property differently depending on their native language specificities (Danish: (Højen & Nazzi, in revision; English: Mani & Plunkett, 2007; 2010; Floccia et al., 2014; Nazzi et al., 2009; French; Havy & Nazzi, 2009; Poltrock & Nazzi, in revision; Zesiger & Jöhr, 2011; Jöhr, 2014; Italian: Hochmann, et al., 2011). However, the origins and the first year developmental trajectories that led to these different patterns of C/V (a)symmetry in toddlers remain unknown. Concerning the origins at birth, in the first experimental chapter we have failed to find that precursors of syntax and lexical processing use one or the other category of speech sounds preferentially, thus suggesting that the consonant and the vowel biases may not be innate. The emergence of the consonant bias in the first year of life remains unexplored. Considering that a consonant bias could bootstrap the acquisition of lexical knowledge (Bergelson & Swingley, 2012; Tincoff & Jusczyk, 1999, 2012) from very early on, this second experimental chapter investigates whether French-learning infants are already able to use the greater lexical informativeness of consonants when processing their first familiar word-forms, in particular their own name.

III.1. Main Study – EXPERIMENTS 4 to 7: Comparing own name recognition with initial Consonant versus Vowel changes (C1/V1) in French-learning monolinguals

This next section corresponds to the text of the following manuscript in press:

Bouchon, C., Floccia, C., Fux, T., Adda-Decker, M., & Nazzi, T. (in press). Call me Alix, not Elix: Vowels are more important than consonants in own name recognition at 5 months. *Developmental Science.*

Section numbers have been added to the original paper.

ABSTRACT

Consonants and vowels differ acoustically and articulatorily, but also functionally: Consonants are more relevant for lexical processing, and vowels for prosodic/syntactic processing. These functional biases could be powerful bootstrapping mechanisms for learning language, but their

developmental origin remains unclear. The relative importance of consonants and vowels at the onset of lexical acquisition was assessed in French-learning 5-month-olds by testing sensitivity to minimal phonetic changes in their own name. Infants' reactions to mispronunciations revealed sensitivity to vowel but not consonant changes. Vowels were also more salient (on duration and intensity) but less distinct (on spectrally-based measures) than consonants. Lastly, vowel (but not consonant) mispronunciation detection was modulated by acoustic factors, in particular spectrally-based distance. These results establish that consonant changes do not affect lexical recognition at 5 months, while vowel changes do; the consonant bias observed later in development does not emerge until after 5 months through additional language exposure.

III. 1. a. Introduction

Consonants and vowels are the two basic sound categories central to the structure of speech in all languages (Ladefoged, 1993). They differ in many respects: for example, vowels tend to be longer and louder than consonants (Repp, 1984), and are thus perceived more clearly in utero (Granier-Deferre, Ribeiro, Jacquet, & Bassereau, 2011) and at birth (Bertoncini, Bijeljac-Babic, Jusczyk & Kennedy, 1988; Benavides-Varela, Hochmann, Macagno, Nespor & Mehler, 2012). Vowels are also less numerous than consonants in most languages (Maddieson, 1984). Furthermore, consonants are processed more categorically than vowels (Fry, Abramson, Eimas & Liberman, 1962), and by partly different brain areas as shown by neuropsychological (Caramazza et al., 2000) and electrophysiological/brain imaging studies (Carreiras & Price, 2008). During development, native vowel categories are learned earlier (6 months: Kuhl, Williams, Lacerda, Stevens, & Lindblom, 1992) than consonant categories (10-12 months: Werker & Tees, 1984).

These differences between consonants and vowels have led to propose a "division of labor" that could help infants learn their native language (Nespor, Peña & Mehler, 2003). Specifically, two complementary functional biases have been proposed, such that *"…consonants, rather than vowels, are most relevant to build the lexicon, and vowels, rather than consonants, are most relevant for grammatical* [and prosodic] *information" (page 224).* This hypothesis has been extremely influential in the field (see below), but very few studies have explored the origins of these hypothesized learning biases. Here we explore the respective role of consonants and vowels in the earliest steps of lexical acquisition, focusing on the emergence of the consonant bias for lexical processing.

Two kinds of hypotheses regarding the origin of the consonant bias have been offered. The "initial bias" hypothesis states that infants start processing consonants and vowels as distinctive linguistic categories from birth, ascribing a limited role to input characteristics and thus predicting no developmental or cross-linguistic differences (Nespor et al., 2003; Bonatti, Peña, Nespor & Mehler,

2007; Pons & Toro, 2010). Alternatively, "learned bias" hypotheses propose that this bias emerges during development, as a result of infants' acquisition of the acoustic-phonetic (Floccia, Nazzi, Delle Luche, Poltrock & Goslin, 2014) or lexical (Keidel, Jenison, Kluender & Seidenberg, 2007) properties of their native language.

Many adult studies, using various tasks, have shown that consonants are privileged over vowels in lexical processing in French, Spanish, Italian, English, and Dutch. This was found in tasks measuring lexical access in both auditory (Cutler, Sebastián-Gallés, Soler-Vilageliu & van Ooijen, 2000; Delle Luche, Poltrock, Goslin, New, Nazzi & Floccia, 2014) and written modalities (Acha & Perea, 2010; New, Araújo & Nazzi 2008), detection of word–forms from continuous speech (Toro, Nespor, Mehler & Bonatti, 2008), and novel word learning (Havy, Serres & Nazzi, 2014).

Several studies also tested the consonant bias in toddlers and children. These studies, initially conducted in French, revealed a consonant bias in novel word learning between 16 months and 5 years of age (Nazzi, 2005; Havy, Bertoncini & Nazzi, 2011; Havy & Nazzi, 2009; Havy, Serres & Nazzi, 2014), and a consonant bias in familiar word recognition at 14 months (Zesiger & Jöhr, 2011). These findings establish that the consonant bias in (French) lexical processing is consistently present from early in development, and converging evidence was found at 12 months in Italian-learning infants (Hochmann, Benavides-Varela, Nespor & Mehler, 2011). However, studies on English-learning infants offer a different picture. A consonant bias was found in word learning tasks at 30 months but not 16 and 23 months (Floccia et al., 2014; Nazzi, Floccia, Moquet & Butler, 2009) and in word recognition tasks at 15 months but not 12, 18 or 24 months (Mani & Plunkett, 2007, 2010). These latter findings appear to go against the initial bias hypothesis, which predicts an early and language-independent consonant bias. However, it remains possible that the consonant bias would only be momentarily masked in English-learning toddlers.

Therefore, the current study aimed to specify its origin and early trajectory by testing whether the consonant bias is present from the very beginning of lexical acquisition, by 5 months of age. This age was chosen given data establishing recognition of some familiar word forms (Mandel, Jusczyk & Pisoni, 1995; Bortfeld, Morgan, Golinkoff, & Rathbun, 2005; Mersad & Nazzi, 2012) and comprehension of some words (Tincoff & Jusczyk, 1999, 2012; Bergelson & Swingley, 2012) at 5/6 months. Moreover, while some of these earlier studies had examined the effects of mispronunciations on early recognition of familiar words (consonant mispronunciation in English: Bortfeld et al., 2005; vowel mispronunciation in French: Mersad & Nazzi, 2012), none had directly compared the effects of consonant and vowel mispronunciations. Here, we tested whether recognition of their own names by French-learning 5-month-olds is affected by a consonant change or a vowel change. Lastly, we also conducted acoustic analyses on the stimuli used in order to

determine whether infants' performance is related to acoustic properties of the stimuli presented, in order to evaluate the acoustic/phonetic learning hypothesis (Floccia et al., 2014).

III. 1. b. Methods

Following Mandel et al. (1995), we used the Headturn Preference Procedure (HPP) to test the sensitivity of French-learning 5-month-olds to a consonant change (Consonant change condition) versus a vowel change (Vowel change condition) in their own name. Infants in the test groups were presented with repetitions of their own correctly-pronounced name (CPs) on half of the trials, and repetitions of their own mispronounced name (MPs) on the other half of the trials. We only used one-feature phonetic changes (Consonant change condition: place, voicing or manner; Vowel change condition: place, roundedness or height). Since all phonetic changes were native French contrasts (Dell, 1985), most, if not all, contrasts were expected to be distinguishable by our French-learning participants, given evidence that most native contrasts can be discriminated by infants from birth (for reviews, see Werker, 1994; Kuhl et al., 2008). Importantly, the use of infants' own names allowed us to test many different consonant and vowel changes (25 in the Consonant change condition and 28 in the Vowel change condition, due to some infants having the same names), providing generalizable results. A preference for CPs over MPs (indicated by the difference in looking times towards CPs and MPs, referred to as *LT.diff* later on) would indicate sensitivity to the mispronunciation (e.g., in the Consonant change condition, a baby named *Victor* should prefer listening to *Victor* over *Zictor*). As a precaution, to rule out effects due to pure acoustic preferences, yoked control infants were tested on the same stimuli as test infants with the main constraints that they had a different name, did not know anyone with the critical name, and had a name starting with the same phoneme category (consonant or vowel) as the critical name (e.g. a baby named *Martin* was presented with *Victor* vs *Zictor*). In this way, CPs and MPs were equally unfamiliar to control infants, and no preference should be observed for CPs over MPs (i.e., a baby named *Martin* should not prefer *Victor* over *Zictor*).

According to the initial bias hypothesis (Nespor et al., 2003), 5-month-olds should already be more sensitive to a consonant change than to a vowel change in their own name. Therefore, test infants should exhibit a larger preference for CPs over MPs in the Consonant change condition compared to the Vowel change condition, and no preference should be found in the yoked control groups of both experiments. Alternatively, if this prediction were not confirmed, it would suggest that the functional specialization of consonants and vowels still has to be learned at 5 months, as proposed by the "learned bias" hypotheses. In this case, a possible outcome could be that 5-month-olds perceive and process consonant and vowel changes similarly in words, in which case we could predict a preference for their correctly pronounced names over their mispronunciations in both

conditions. A further possibility is that 5-month-olds' reaction to the mispronunciations is based on the acoustic distance between the correctly versus mispronounced stimuli, independently (or not) of their consonant/vowel status. Therefore, in order to assess the contribution of the acoustic-phonetic properties of our stimuli to infants' preference responses, the consonant and vowel contrasts were characterized along three acoustic dimensions: duration, intensity and MFCCs (Mel-Frequency-Cepstrum Coefficients, a spectrally-based measure of phonetically-relevant acoustic information normalized for duration and intensity). These measures were used to evaluate the saliency and the discriminability of the different phonemes, and relate them to individual preference responses. We expected consonants to be overall less salient than vowels (Repp, 1984) but more discriminable as they are usually perceived more categorically than vowels (Fry et al., 1962).

III.1.b.i. Participants

All 120 participants were healthy French-learning monolingual 5-month-old infants (***Table III. 1.)***. For the test conditions, only infants without nicknames and who were frequently called by their own name were included. Infants in the control conditions were chosen so that they would not know anyone in their environment with the name used in the experiment. Children with names starting with a consonant were assigned to the Consonant change control group and those with names starting with a vowel to the Vowel change control group. Forty-seven additional infants were tested and their data excluded due to fussiness (36), having two consecutive trials with insufficient looking times, having more than three such trials overall (6), experimenter error (1), or being an outlier (*LT.Diff* above or below 2 *SDs* of the group Mean; 4).

Groups (all n = 30)	Stimuli (example)	Infant's name (example)	Age in days (SD)	# girls/boys
Consonant change condition				
Test (Exp. 4)	e.g. Victor vs. Zictor	Victor	164 (8)	14/16
Control (Exp. 6)	e.g. Victor vs. Zictor	Martin	164 (9)	11/19
Vowel change condition				
Test (Exp. 5)	e.g. Esther vs. Isther	Esther	164 (8)	17/13
Control (Exp. 7)	e.g. Esther vs. Isther	Adrien	165 (8)	14/16

Table III. 1. Experiments 4-7. Participant information, illustration of the 4 experimental conditions, and examples of stimuli.

III.1.b.ii. Stimuli

Each of the 60 test infants heard repetitions of a pair of stimuli corresponding to their CP (correctly pronounced) and MP (Mispronounced) names. Due to a few infants having the same names, there was a total of 28 pairs in the Consonant Change condition and 25 pairs in the Vowel Change condition. The same stimuli were used with the 60 yoked control infants. The MP of the

names always consisted of a 1-feature change (***Table III. 2***.). As mentioned earlier, all changes were native French minimal contrasts according to Dell (1985), and were expected to be discriminable by French-learning infants irrespectively of their age. Based on Mandel et al. (1995) and the fact that 3 different phonetic features were tested in each condition, the sample size was 30 infants in each test group. In the Consonant Change condition, 10 infants were tested on a place-of-articulation change, 10 on a voicing change, and 10 on a manner-of-articulation change. In the Vowel Change condition, 10 infants were tested on a place change, 10 on a roundedness change, and 10 on a height change.

CONSONANT Change Condition

PLACE changes					VOICING changes					MANNER changes				
#	CP & Test infant's name	MP	contrast CP/ MP	Control infant's name	#	CP & Test infant's name	MP	contrast CP/ MP	Control infant's name	#	CP & Test infant's name	MP	contrast CP/ MP	Control infant's name
1	Sophie	chophie	s/ʃ	Martin	1	Judith	chudith	ʒ/ʃ	Clément	1	Marie	barie	m/ b	Félix
2	Rémi	lémi	R/l	Micha	2	Pénélope	bénélope	p/b	Martin	2	Marie	barie	m/ b	Loic
3	Gaspard	daspard	g/d	Maya	3	Côme	gôme	k/g	Solène	3	Lily	zily	l/z	Gaspard
4	Camille	tamille	k/t	Serge	4	Véra	féra	v/f	Toto	4	Lucien	ducien	l/d	Marc-Antoine
5	Marin	narin	m/n	Rose	5	Jules	chules	ʒ/ʃ	Loulou	5	Nora	zora	n/z	Botho
6	Pauline	tauline	p/t	Robinson	6	Pierre	bierre	p/b	Nina	6	Bérénice	mérénice	b/ m	Diane
7	Victor	zictor	v/z	Mathieu	7	Corentin	gorentin	k/g	Maeli	7	Téa	séa	t/s	Marguerite
8	Loulou	roulou	l/R	Basil	8	Violette	fiolette	v/f	Mathilde	8	Laomé	zaomé	l/z	Mathias
9	Félix	sélix	f/s	Martin	9	Félix	vélix	f/v	Mara	9	Nouchai	zouchai	n/z	Célian
10	Félix	sélix	f/s	Paulin	10	Charles	jarles	ʃ/ʒ	Mona	10	Sanjay	tanjay	s/t	Margot

VOWEL Change Condition

PLACE changes					ROUNDEDNESS changes					HEIGHT changes				
#	CP & Test infant's name	MP	contrast CP/ MP	Control infant's name	#	CP & Test infant's name	MP	contrast CP/ MP	Control infant's name	#	CP & Test infant's name	MP	contrast CP/ MP	Control infant's name
1	Olivia	eulivia	o/ø	Ewan	1	Ellen	eullen	e/ø	Haythem	1	Alba	elba	a/ɛ	Ombelle
2	Auriane	euriane	o/ø	Anaïs	2	Elisa	eulisa	e/ø	Alice	2	Océane	ouceane	o/u	Anjali
3	Augustine	eugustine	o/ø	Alexandre	3	Iris	uris	i/y	Hédi	3	Alexandre	elexandre	a/ɛ	Héloise
4	Oriane	euriane	o/ø	Eloise	4	Eliott	euliott	e/ø	Andrée	4	Anais	enais	a/ɛ	Elio
5	Ulysse	oulysse	y/u	Hector	5	Ihsane	uhsane	i/y	Emmanuel	5	Arsène	ersene	a/ɛ	William
6	Hugo	ougo	y/u	Héloise	6	Elise	eulise	e/ø	Alexia	6	Honorine	ounorine	o/u	Ambroise
7	Augustine	eugustine	o/ø	Adèle	7	Erwan	eurwan	ɛ/œ	Abigaëlle	7	Esther	isther	e/i	Adrien
8	Hugo	ougo	y/u	Alix	8	Eliott	euliott	e/ø	Alexandre	8	Inès	enes	i/e	Adrien
9	Hugo	ougo	y/u	Aron	9	Eli	euli	e/ø	Aliénor	9	Alix	elix	a/ɛ	Ethan
10	Eugénie	ogenie	ø/o	Alienor	10	Elias	eulias	e/ø	Enzo	10	Arthur	erthur	a/ɛ	Evan

Table III. 2. Experiments 4-7. Details of stimuli presented in the Consonant change condition (upper-panel) and the Vowel change condition (lower-panel), broken down by phonetic features (in columns): for each test infant, the CP name (corresponding to the test infant's name Exp. 4-5), the MP name, the phonetic feature contrasting CP and MP (in IPA), and the control infant's name tested on the same CP/MP pair (Exp. 6-7).

For all infants, the same female native French speaker recorded 15 tokens each of CP and MP names. For both CPs and MPs, two files including the 15 tokens were made, the second file presenting the same tokens in reversed order. All sound files lasted 24 seconds.

III.1.b.iii. Procedure

Behavioral experiment.

Each infant was tested individually. The experiment was conducted inside a sound attenuated room, and in a booth made of pegboard panels (bottom) and a white curtain (top). The test booth had a red light and a loudspeaker (SONY xs-F1722) mounted at eye level on each of the side panels and a green light mounted on the center panel. Below the center light was a video camera used to monitor infants' behavior.

A PC computer terminal (Dell OptiPlex), a TV screen connected to the camera, and a response box were located outside the sound attenuated room. The response box, connected to the computer, was equipped with a series of buttons. The observer, who looked at the video of the infant on the TV screen to monitor infant's looking behavior, pressed the buttons of the response box according to the direction the infant's head, thus starting and stopping the flashing of the lights and the presentation of the sounds (see below). The observer and the infant's caregiver wore earplugs and listened to masking music over tight-fitting closed headphones, which prevented either from hearing the stimuli presented. Information about the duration of the head-turn, calculated from the observer's button-pressing, was recorded by the computer.

We used the same version of the Head-turn Preference Procedure (HPP) as Mandel et al. (1995). Each infant was held on a caregiver's lap in the center of the booth. Each trial began with the green light on the center panel blinking until the infant oriented to it. Then, the red light on one of the side panels began to flash. When the infant turned their head in that direction, the stimulus for that trial began to play. The stimuli were delivered by the loudspeakers via an audio amplifier (Marantz PM4000). Each stimulus was played to completion or stopped immediately after the infant failed to maintain the head-turn for 2 consecutive seconds. If the infant turned away from the target by 30° in any direction for less than 2 s and then turned back again, the trial continued but the time spent looking away (when the experimenter released the buttons of the response box) was automatically subtracted from the listening time by the computer program. Thus, the maximum listening time for a given trial was the duration of the entire speech sample. If a trial lasted less than 1.5 s (insufficient looking time), the trial was repeated and the original listening time was discarded.

Each session began with 2 musical trials (excerpts of classical music), one on each side to give infants an opportunity to practice one head-turn to each side. The test phase consisted of 8 trials

divided in 2 blocks, in each of which the two lists of each name were presented. Order of the different lists within each block was randomized.

Acoustic analyses of the stimuli

Three acoustic dimensions were measured to characterize the contrasted phonemes of CPs and MPs: duration, intensity and Mel Frequency Cepstral Coefficients as a measure of spectral distance (MFCCs; see below for more explanation).

For each CP/MP pair, duration and intensity were measured for the 15 tokens of the contrasted phonemes using PRAAT. This was first used to calculate mean duration and mean intensity values of the contrasted phonemes, in order to compare the relative salience and discriminability of the contrasted consonants and vowels. Second, we computed normalized duration and intensity differences (*Diff.duration*: duration difference between the contrasted phonemes of CPs and MPs divided by their mean; same for *Diff.intensity*), in order to test their link with individual performance.

MFCCs are spectrum-based features resulting from a deconvolution of the speech source (e.g. vocal fold vibrations) and speech filter (vocal tract). MFCCs were chosen rather than pitch measures because, first, pitch measures cannot be calculated for some of the consonants used in our study (unvoiced consonants). Second, MFCCs are widely used both for automatic speech and speaker recognition, as they provide a general measure of distance between two speech sounds normalized for duration and intensity that specifies well both consonant and vowel information (however though, MFCCs do not provide information regarding saliency, contrary to duration, intensity and pitch). They have been preferentially used in word and phoneme recognition studies because they retain phonetically-relevant acoustic information (e.g., Davis & Mermelstein, 1980). They involve a pre-processing of the spectral envelope of the signal with frequency bands equally spaced on the Mel scale that approximates the psychoacoustic properties of the cochlea, thus providing a better acoustic/phonetic coding than more simple measures such as spectral distance and LFCC coefficients (Linear-Frequency Cepstral Coefficients).

MFCCs were calculated using 30 ms analysis windows at a 15 ms frame rate. To do so, for each CP/MP pair, the word-initial (contrasted) phoneme of the first of the 15 tokens of each word was manually segmented using Transcriber (Barras et al., 2001). Then, manual segmentation was used to automatically locate the initial phonemes of the 14 other tokens using dynamic time warping (DTW; Sakoe & Chiba, 1978). DTW is a speech comparison method that automatically determines the optimal temporal matching between two speech patterns (detect segment similarities) independently of duration and speech rate. Then, MFCCs were computed using the 24 triangular filters mel-frequency spaced, a standard discrete cosine's transform and frequency bandwidth of 0-8

kHz. The subset of MFCCs employed in the classification to measure MFCC distances included 12 coefficients, c1 to c12, in order to best represent the envelope of the mel-spectrum. Note that coefficient c0 was not taken into account to exclude intensity differences that would affect MFCC distance measurements. MFCC distances correspond to the Euclidian distance between two tokens calculated for the 12 coefficients (i.e., the square root of the summed squared differences between the two MFCC sets). Lastly, we tested the link between individual performance and the normalized MFCC distance between CPs and MPs (*Diff.spectral)*, defined for each CP/MP pair as the ratio of the mean cross-category distance between the 15 CPs and the 15 MPs (*Dcross)* of the given pair and the mean internal variability within the 15 CPs (*DwithinCP)* and within the 15 MPs (*DwithinMP)*:

$$Diff.spectral(CP, MP) = \frac{Dcross(CP, MP)}{\sqrt{\frac{1}{2}(DwithinCP^2 + DwithinMP^2)}}$$

III. 1. c. Results

III.1.c.i. Behavioral results

Overall analysis

Mean listening times (LTs) to the CP and MP names were calculated for each infant. Group averages are presented in ***Figure III. 1***. A three-way ANOVA was conducted on LTs with a within-subjects factor of pronunciation (CP vs. MP) and between-subjects factors of group (test vs. control) and condition (Consonant Change vs. Vowel Change). Neither the effect of condition ($F(1, 116) = .26$, $p = .61$), group ($F(1, 116) = .87$; $p = .35$), nor the condition × pronunciation interaction ($F(1, 116) = .11$, $p = .74$) reached significance. The effect of pronunciation was only marginal ($F(1, 116) = 3.53$; $p = .06$), Importantly though, both the pronunciation x group interaction ($F(1, 116) = 7.23$; $p = .008$, $\eta^2_p = .06$) and the 3-way interaction between pronunciation x group x condition ($F(1, 116) = 8.64$; $p = .004$, $\eta^2_p = .07$) reached significance, establishing that infants were not behaving in the same way in both conditions. In order to understand how consonant and vowels MPs were differently processed, separate analyses were conducted for each condition.

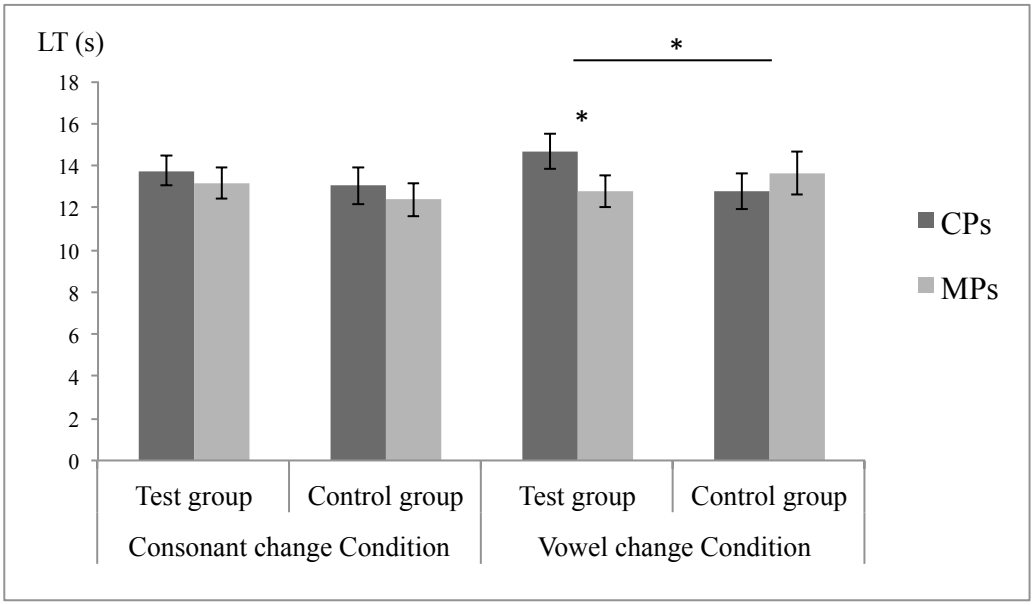


Figure III. 1. *Experiments 4-7. Mean listening times (and SEs) to the CP and MP names in test and control infants. Left panel: Consonant Change condition (test group: Exp. 4; control group: Exp. 6); Right panel: Vowel Change condition (test group: Exp. 5; control group: Exp. 7).*

Consonant change condition

For the test group (Exp. 4), the LT means were M_{CP} = 13.87 s (*SD* = 3.91 s) and M_{MP} = 13.36 s (*SD* = 3.98 s), the estimate of *LTdiff* was *Mtest* = .51 s (95% *CI* = [-.40, 1.42]) and Cohen's *d* = .21 (Cohen's *ds* are calculated with the difference of means as numerator and the standard deviation of the paired differences as denominator). For the control group (Exp. 6), the LT means were M_{CP} = 13.03 s (*SD* = 4.74 s) and M_{MP} = 12.40s (*SD* = 4.38 s), the estimate of *LTdiff* was *Mctrl* = .64 s (95% *CI* = [-.51, 1.78]) and Cohen's *d* = .21. The size of the difference between *LTdiffs* in test and control infants was -.13 (95% *CI* = [-1.63, 1.37]) with Cohen's *d* = -.03. A two-way ANOVA on LTs with the factors of pronunciation (CP vs. MP) and group (test vs. control) was conducted. The effects of pronunciation ($F(1, 58) = 2.56$, $p = .11$), group ($F(1, 58) = .75$, $p = .39$), and the pronunciation x group interaction ($F(1, 58) = .03$, $p = .86$) all failed to reach significance. Longer LTs for CPs over MPs were found for 14 of the 30 test infants (binomial test, $p = .43$), and for 16 of the 30 control infants (binomial test, $p = .43$). This pattern of results shows no effect in both the test and control groups, and no difference between the two groups.

Moreover, a second ANOVA including the between-subjects factor of feature (place, voicing or manner; all n = 10) in addition to the previous factors yielded no effect of feature ($F(2, 54) = 1.16$; $p = .32$), and no interactions involving feature (all $F's < 1$). Overall, these results show that French-learning 5-month-olds are not particularly sensitive to a consonant mispronunciation in their own name, independently of the phonetic feature contrasted.

Vowel change condition

For the test group (Exp. 5), the LT means were M_{CP} = 14.75 s (*SD* = 4.72 s) and M_{MP} = 12.89 s (*SD* = 4.35 s), the estimate of *LTdiff* was *Mtest* = 1.86 s (95% *CI* = [.74, 2.89]) and Cohen's *d* = .62. For the control group (Exp. 7), the LT means were M_{CP} = 12.76 s (*SD* = 4.24 s) and M_{MP} = 13.82 s (*SD* = 5.02 s), the estimate of *LTdiff* was *Mctrl* = -1.06 s (95% *CI* = [-2.10, -.01]) and Cohen's *d* = -.38. The size of the difference between *LTdiffs* in test and control infants was 2.92 (95% *CI* = [1.33, 4.51]) with Cohen's *d* = .68. A two-way ANOVA on LTs with the factors of pronunciation and group was conducted. The effect of pronunciation ($F(1, 58) = 1.14$, $p = .29$) and group ($F(1, 58) = .22$, $p = .64$) failed to reach significance, but the pronunciation × group interaction did ($F(1, 58) = 15.11$, $p = .0003$, $\eta^2_p = .21$). Planned comparisons revealed that Test infants significantly preferred their CP names compared to their MP versions ($F(1, 58) = 12.28$, $p = .0009$, $\eta^2_p = .17$) whereas Control infants marginally preferred MPs over CPs ($F(1, 58) = 3.97$, $p = .051$, $\eta^2_p = .06$). Note that this marginal (reversed) preference for vowel MPs in Control infants, which goes against the Test infants' preference for their correctly pronounced names, confirms the strength of this preference in the Test infants. Longer LTs for CPs over MPs were found for 21 of the 30 test infants (binomial test, $p = .02$), but only for 12 of the 30 control infants (binomial test, $p = .18$). This pattern of results shows a medium effect in the test group (preference for CPs) and a small effect in the control group (preference for MPs), and a large difference between the two groups.

Moreover, a second ANOVA including the factor feature (place, roundedness, or height; all n = 10) in addition to the previous factors yielded no effect of feature and no interaction involving feature (all *Fs* < 1), while the pronunciation × group interaction was still significant ($F(1, 54) = 14.43$, $p = .0004$, $\eta^2_p = .21$). These results establish that French-learning 5-month-olds prefer their correctly pronounced name over a 1-feature mispronunciation of that name, independently of the feature tested, as opposed to control infants who show a marginal preference for the mispronunciation. The present results thus exclude the possibility that infants in the Consonant Change condition did not have a preference for their correctly pronounced name due to difficulties with our testing procedure.

III.1.c.ii. Acoustic measures

Acoustic measures were conducted on the stimuli (see Stimuli section for details), which consisted of 28 different pairs of CPs-MPs in the Consonant Change condition and 25 pairs of CPs-MPs in the Vowel Change condition.

Duration and Intensity

Regarding salience, the consonants lasted 73.2 ms and were 71.4 dB loud on average, while the vowels lasted 106.9 ms and were 78.9 dB loud. Regarding discriminability within each pair of contrasted phoneme (e.g. the /v/ in Victor versus the /z/ in Zictor), consonant CPs were on average

9.8 ms shorter (95% *CI* = [-22.2, 2.7]) and 1.1 dB louder (95% *CI* = [-3.0, 5.4]) than consonant MPs; vowel CPs were on average 2.9 ms shorter (95% *CI* = [-7.8, 2.12]) and .5 dB softer (95% *CI* = [-1.8, 0.8]) than vowel MPs. To further explore whether there were differences in salience or discriminability between the consonant and vowel conditions, two separate ANOVAs examining duration and intensity were run, with the factors of pronunciation (CP versus MP) and condition (Consonant change vs. Vowel change). In both cases, there was only a significant effect of condition with consonants being shorter (duration: $F(1, 51) = 33.29$, $p < 10^{-6}$, $\eta^2_p = .40$) and softer than vowels (intensity: $F(1, 51) = 41.35$, $p < 10^{-6}$, $\eta^2_p = .45$), hence establishing that vowels were more salient than consonants. Moreover, there were no effect of pronunciation nor pronunciation × condition interaction, suggesting that in both conditions, CPs and MPs could not be discriminated based on duration and intensity differences (therefore, the marginal preference for MPs over CPs found in the Control group in Exp. 5 and 7 cannot be attributed to differences in terms of duration or intensity of the first phonemes).

Spectral Measures

The normalized acoustic/phonetic distance (*Diff.spectral,* normalized for duration and intensity, based on MFCCs) was used to assess discriminability (as mentioned above, MFCCs are not meaningful with respect to salience). On average, *Diff.spectral* was 1.54 (SE = .06) for consonant contrasts, and this was significantly higher than the same index for vowels (1.36, SE = .03; $t(52) = 6.45$, $p = .01$, $\eta^2_p = .11$). This establishes that consonant contrasts were acoustically more distinct than vowel contrasts, once normalized for intensity and duration.

III. 1. c. iii. Acoustic predicates of preference measures

We then explored the link between individual test infants' performance as attested by the difference in LTs between CPs and MPs (*LT.diff*), and three independent measures of acoustic distance between CPs and MPs: normalized duration and intensity differences (*Diff.duration* and *Diff.intensity*) and *Diff.spectral*. First a multiple linear regression was run on all 60 test infants (30 Consonant change, 30 Vowel change) with *LT.diff* as the dependent variable and the 3 acoustic distances as predictors (for which there was no colinearity, all Variance Inflation Factors *VIFs* < 3). The model did not explain a significant part of the variance in *LT.diff* ($R^2 = .01$, $R^2_{adjusted} = -.04$; $F(3, 56) = .31$, $p = .81$), and none of the predictors significantly predicted the difference in LTs between CPs and MPs. This analysis suggests that the variability in infants' detection of an MP in their own name is not explained by differences in any of the three acoustic dimensions, at least when analyzing the consonant and vowel conditions together. However, it is possible that infants distinguish and process consonants and vowels differently (which could be supported by their differences in salience reported above) and hence that acoustic distances have different effects within each phonemic

category. To explore this possibility, we ran the same regression as above but separately for each condition.

In the Consonant change condition (all *VIFs* < 3) this initial 3-predictor model did not explain a significant part of the variance in *LT.diff* (R^2 = .10, R^2*adjusted* = .001; $F(3, 26) = 1.00$, p = .41). For comparison purposes with the Vowel change condition (see below), we then ran a backward regression, which, not surprisingly, did not converge onto any model and in which none of the predictors contributed significantly to *LT.diff*. Therefore, the variance in infants' differential LTs towards CPs versus consonant MPs is not predicted by our 3 acoustic distances (see ***Figure III. 2***, left panel, presenting the individual data for *Diff.spectral*, which is further considered in the Vowel change condition).

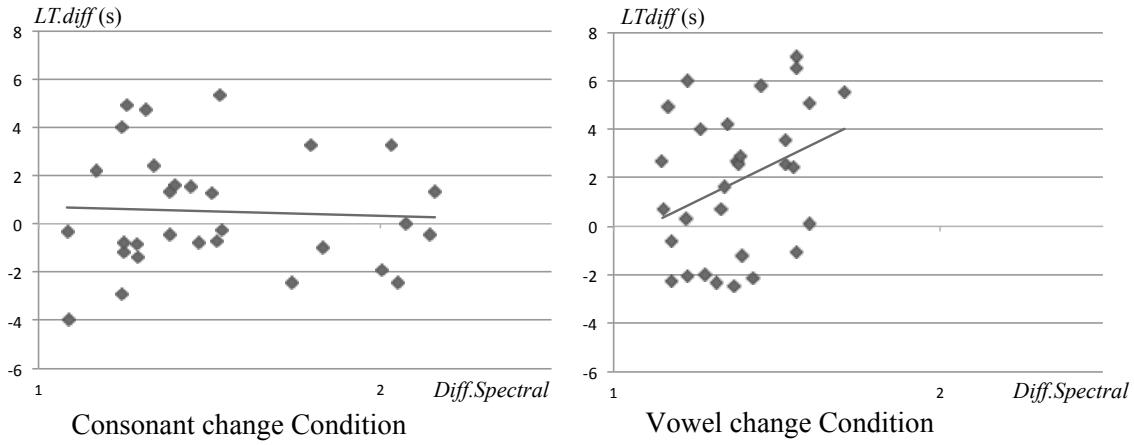


Figure III. 2. *Experiments 4-5. Link between LT.diff and Diff.Spectral in the Consonant change condition (Exp. 4, left) and Vowel change condition (Exp. 5, right).*

In the Vowel change condition (all *VIFs* < 3), the regression model with *LT.diff* and the 3 predictors yielded a marginal model ($F(3, 26) = 2.51$, p = .08) explaining 22.5% of the variance (R^2*adjusted* = .135; *standard error of estimate* = 2.80). In this model, *Diff.spectral* was the only significant predictor ($\beta_{spectral}$ = .394, p = .034). We then ran a backward regression that converged onto a significant 2-predictor model ($F(2, 27) = 3.44$, p = .047) resulting in the exclusion of *Diff.intensity*. This final model explained 20.3% of the variance (R^2*adjusted* = .144, *standard error of estimate* = 2.78) and included *Diff.spectral* and *Diff.duration* as predictors ($\beta_{spectral}$ = .369, p = .042; $\beta_{duration}$ = .294, p = .10). This shows that the variability of *Diff.spectral* and *Diff.duration* taken together accounts for a significant part of the variance in infants' differential LTs, and that *Diff.spectral* is a significant predictor of *LT.diff* even though the distribution of *Diff.spectral* values was narrower in the Vowel change condition than in the Consonant change condition (see ***Figure III. 2,*** right panel). Therefore, contrary to what was found in the Consonant change condition, the variability introduced by two acoustic distances (in particular spectral distance) in vowels partially determines French-

learning 5-month-olds' preference for their correctly pronounced name over a 1-feature vowel mispronunciation.

Lastly, the same three multiple linear regressions were run with the performance of Control infants as the dependent variable (*LT.diff_ctrl*) and the same three independent measures of acoustic distance between CPs and MPs: normalized duration and intensity differences (*Diff.duration* and *Diff.intensity*) and *Diff.spectral*. None of the 3 models explained a significant part of the variance in infants' differential LTs nor yielded any significant predictor (both conditions together: R^2 = .03, $R^2adjusted$ = -.03; $F(3, 56) = .50$, p = .69; Consonant condition: R^2 = .08, $R^2adjusted$ = -.03; $F(3, 26) = .74$, p = .54; Vowel condition: R^2 = .04, $R^2adjusted$ = -.07; $F(3, 26) = .35$, p = .79).

III. 1. d. Discussion

Previous work in adults and toddlers has shown that consonants are more important than vowels in learning and recognizing words. The present study explored whether there is an early consonant bias in one of the earliest words known by infants: their own name. French-learning 5-month-olds were tested on the impact of either a consonant or a vowel mispronunciation on their listening preference for their name. All mispronunciations involved one-feature changes (based on Dell, 1985) between two native consonants or vowels, which were expected to be discriminated by the participants. In the Consonant change condition, infants failed to show a preference for their correctly pronounced names (CPs) over consonant mispronunciations (MPs). In the Vowel change condition, infants preferred their CP names over vowel MPs, demonstrating sensitivity to vowel MPs. These findings were found independently of the consonantal (place, voicing, manner) and vocalic (place, roundedness, height) features tested. Thus, French-learning infants, tested on their sensitivity to minimal phonetic changes in one of their first words, exhibit a vowel bias at the onset of lexical processing.

Our results do not support the initial bias hypothesis (Nespor et al., 2003), which predicted a consonant bias at the onset of lexical processing. On the contrary, they suggest that the consonant bias is learned. This is in line with previous reports showing cross-linguistic differences, in particular the later emergence of the consonant bias in English-learning infants than in French-learning infants (Floccia et al., 2014; Nazzi et al., 2009). One way to interpret the English-French discrepant data has been to suppose a temporary disappearance of the initial consonant bias in English toddlers (U-shaped trajectory, see Floccia et al., 2014, for a discussion). However, the present findings contradict this idea, because even in French, a language in which the consonant bias is consistently found in toddlers (e.g., Havy & Nazzi, 2009; Nazzi, 2005; Nazzi et al., 2009), infants did not show this pattern.

Indeed, infants did not consider consonant changes in a familiar word as lexically relevant, as clear preference responses were found only for vowel changes.

The current findings support the idea that infants have to learn the differentiated functional roles of consonants and vowels. Together with previous reports of a consonant bias at 12 to 14 months of age in Italian- and French-learning infants (Hochmann et al. 2011; Zesiger & Jöhr, 2011), the present findings suggest (at least in these two languages), that the time window for the emergence of a consonant bias is the second semester of life. For this learning to take place, one possibility (the lexical hypothesis, Keidel et al., 2007) is that infants need to acquire a sizeable lexicon to discover that consonants are statistically more informative than vowels regarding the identity of the words in the lexicon. That could be done through the computation of consonant versus vowel tiers, or consonant versus vowel phonological neighbors; using this information, infants would learn that consonants are more relevant for word learning and word processing. Another possibility (the acoustic-phonetic hypothesis, Floccia et al., 2014) is that the acoustic and phonetic differences between consonants and vowels provide an early cue to infants that these types of speech sounds should be processed differently. While the current study was not designed to separate these two hypotheses, some of the present findings nevertheless contribute to this issue.

Our acoustic analyses confirm that vowels are acoustically more salient than consonants both in duration and intensity (Repp, 1984). Importantly, they also reveal that vowel contrasts are spectrally *less* distinct than consonant contrasts per time unit and controlling for intensity. In addition, infants' preference responses clearly establish that consonants and vowels are processed differently in word recognition by 5 months: infants' recognition of their name was more impaired by vowel MPs than consonant MPs, and only their performance with vowel MPs was modulated by acoustic factors. However these acoustic cues were not significant factors in the overall regression analysis, suggesting that infants might assign consonants and vowels to abstract categories independently of the acoustic factors considered here. Therefore, consonants and vowels differ on acoustic saliency (vowel advantage) and discriminability (consonant advantage), and are processed differently in a lexical task early in development.

The above observations bring support to the acoustic-phonetic hypothesis in principle; however, we are left to explain why 5-month-olds nevertheless give more relevance to vowels than to consonants in the current name recognition task. From an acoustic point of view, while it is likely that 5-month-olds can discriminate (most of) the native consonant and vowel contrasts used in the present study when presented in short speech sequences, one possibility is that 5-month-olds are still better at processing acoustic details in the most salient portions of the signal, which are usually the vocalic parts, in line with well-established findings of better perception of vowels in utero and at birth (Bertoncini et al., 1988; Benavides-Varela et al., 2012; Granier-Deferre et al., 2011). It is also

possible that infants' early preference for infant-directed speech, which is characterized by large prosodic modulations mostly carried by vowels, contributes to 5-month-olds' greater reliance on vowels in the current study (Fernald, 1985; Werker & McLeod, 1989). From a phonetic/phonological point of view, another possibility is that 5-month-olds have more advanced knowledge of native phonetic categories for vowels than for consonants (Kuhl et al., 1992; Werker & Tees, 1984), to a point where they have started learning native vocalic phonemic categories but not native consonantal phonemic categories (in line with the proposal by PRIMIR that phonemes are learned in a staggered fashion, Werker & Curtin, 2005). Finally, the finding that 5-month-olds only used the vocalic contrasts in this lexically-related task (name recognition) while being very likely able to discriminate both consonantal and vocalic contrasts in speech perception tasks, extends to this age previous reports of a possible dissociation between phonetic discrimination and use of phonetic information at the lexical level in toddlerhood as a function of level of acquisition, task requirement, or level of processing (e.g., Nazzi, 2005; Stager & Werker, 1997).

But how will the appropriate consonant bias in lexical processing be discovered in development? As they grow older, infants' temporal resolution in sounds becomes more acute (Werner, Marean, Halpin, Spetner & Gillenwater, 1992), which might allow them to perfect their ability to perform fine-grained phonetic distinctions for shorter speech sounds (i.e., consonants) in words. Acoustic/phonetic distance per time unit was found to be larger in consonants than in vowels, which could lead infants to switch attention from the vowels to the acoustically more reliable consonants. A second scenario, related to phonetic/phonological acquisition, is that the learning of native consonantal phoneme categories around 10-12 months of age (Werker & Tees, 1984) could induce a shift in cognitive resources or attention towards consonants in word processing, in line with the PRIMIR proposal (Werker & Curtin, 2005). While these explanations are compatible with the acoustic-phonetic hypothesis, a third possibility related to the lexical hypothesis is that the consonant bias emerges with the acquisition of a sizeable lexicon, allowing infants to discover that there are more phonological neighbors obtained by a consonant than a vowel change, hence that consonants are more informative at the lexical level. One good test against this latter hypothesis would be to establish that the consonant bias is already present at 8 months, an age at which infants' limited vocabularies have little or no phonological neighbors.

To sum up, these novel results break ground in specifying the relative contribution of consonants and vowels in early word recognition, directly contradicting previous accounts of the developmental origins of the consonant bias in lexical processing. Our findings are the first to report comparative evidence of mispronunciation detection of consonants and vowels in a large set of familiar words as early as 5 months, and to ground these effects in

the acoustic properties of the words being presented, which will have to be extended to more kinds of words. Importantly, 5-month-olds were found to make lexical distinctions based on minimal changes for vowels (though not for consonants) long before they have started talking in an intelligible way, supporting the notion that early lexical representations are already quite elaborate. Future studies will need to extend this finding in different ways. First, given evidence of cross-linguistic variation in the expression of the consonant bias in toddlerhood (Floccia et al., 2014), the present study will need to be extended to other languages. Second, the current study focused on infants' name, which might have a special valence and status (see Hall, 2009, for a review) and might thus be processed differently than other kinds of words such as count nouns. Recent evidence suggests that this might not be the case, since French-learning 6-month-olds were also found to be more sensitive to a vocalic than a consonantal change in a word segmentation task in which the target words were unfamiliar monosyllabic count nouns (Nishibayashi & Nazzi, 2014). Future studies will have to further explore the generality of this early vocalic bias and its link to the emerging consonant bias across languages.

III. 2. Follow-up experiments investigating sensitivity to Consonant mispronunciations

Our main study (Exp. 4-7) has established that at 5 months, monolingual French-learning infants react significantly to an initial vowel mispronunciation in their own name, but not to an initial consonant mispronunciation in this first familiar word-form. Moreover, acoustic analyses confirmed that vowel contrasts were more salient than consonant contrasts, while establishing that consonants were more distinct spectrally, thus were lexically more informative than vowels. These results suggest that at 5 months, reactions to mispronunciations are driven by acoustic saliency and the lexical specialization of consonants remains to be learned. However it does not imply that 5-month-olds have not started processing fine-grained consonant information in words altogether. Therefore follow-up Experiments 8 and 9 explore 5-month-olds' sensitivity to a consonant mispronunciation in their own name in circumstances where this change would possibly be easier to process than in the main study's consonant change condition. In particular, Experiment 8 investigates infants' reactions

to a *final* consonant mispronunciation, and Experiment 9 explores *bilingual* infants' reactions to an initial consonant mispronunciation.

III. 2. a. EXPERIMENT 8: Final change in French-learning monolinguals

III. 2. a. i. Introduction

The finding of our main study that 5-month-olds are not affected by a consonant mispronunciation in a familiar word while they are affected by a vowel change may have been specific to the position of the change tested. First, to match with the initial consonant change in Experiment 4, changes in vowels in Experiment 5 were also initial. Consequently, infants' names in Experiment 4 were all consonant-initial, while in Experiment 5 they were all vowel-initial. The side effect of this methodological precaution is that the vowel change in Experiment 5 often consisted in changing a whole syllable, whereas the consonant change in Experiment 4 involved a single phoneme at the onset of the initial CV(CC) syllable. Therefore, in our study the vowel change potentially involved a higher linguistic unit than the consonant change and may thus have been more salient.

Moreover, French has phrase-final lengthening (Delattre, 1965), i.e. final syllables are often longer and louder in words and phrases, including in isolated words (see Table 2 in Nazzi et al., 2014, for acoustic measures showing higher amplitudes and longer durations of second over first syllables in French words recorded in isolation from Polka & Sundara, 2012). If exposure to this pattern of words leads French-learning infants to attend more to phonetic details in the final than in the initial syllables of words, then a consonant change may be easier to process if it occurs at the onset of the final syllable instead of the initial syllable, as it did in Experiment 4. Indeed, several studies have found that French-learning infants seem more sensitive to word-final than to word-initial consonant changes in familiar words at 11 months (Hallé & Boysson-Bardies, 1996; and later (Zesiger & Jöhr, 2011; Jöhr, 2014).

Indeed, for word form recognition at 11 months, Hallé and de Boysson-Bardies (1996), later reanalyzed by Vihman et al. (2004), showed that French-learning infants are sensitive to a consonant change both in the initial (e.g. *canard-ganard* 'duck') and the final syllable of familiar words (*canard-kalar*), but their sensitivity was greater for the final change. Interestingly, the opposite effect was found in English-learning infants at the same age (Vihman et al., 2004). Given that the English words used in that study were had a trochaic stress pattern, which is predominant in English (Cutler & Carter, 1987), the initial syllables of disyllabic words were more salient than the final syllables. Thus, in each language group, the level of detail of consonant representations or the ability to process that information seems to be modulated by the acoustic salience of the syllable in which the consonant is embedded.

Some studies have also compared consonant and vowel processing in novel word learning tasks or familiar word recognition tasks between 14 and 30 months of age, manipulating the position of the critical phonemes within words. Depending on the method used, they have found that the consonant bias in French-learning infants was either independent in some studies (at 16, 20 and 30 months: Havy & Nazzi, 2009; Nazzi & Bertoncini, 2009; Nazzi & New, 2007; Nazzi et al., 2009) or dependent of the position of the change with an advantage for final over initial changes in other studies (at 14, 18 and 22 months: Jöhr, 2014, chapter 2-5; Zesiger & Jöhr, 2011). In the latter studies, infants were tested in an IPL paradigm to test word recognition. Infants failed to react to an initial and final vowel mispronunciation (e.g. *poussette* - *peussette* 'stroller') or consonant mispronunciation (e.g. *poussette* – *moussette*), but were crucially sensitive to a final consonant change (e.g. *chaussette* – *chaufette* 'sock'). Together with the results showing that French-learning 11-month-olds react more to mispronunciations in the final than in the initial syllable of a list of familiar words (Hallé & de Boysson-Bardies, 1996; Vihman et al., 2004), these results suggest that in word-recognition tasks French-learning infants may be more sensitive to a change occurring in the more salient final syllable than in the less salient initial syllable of a familiar word.

The sensitivity to consonant changes in words has not been compared for different syllabic positions in French-learning infants before 11 months, thus the effect of syllabic saliency on the processing of phonetic detail within word remains unknown at younger ages. However, in English-learning infants, Bortfeld et al. (2005) revealed sensitivity to a consonant change at 6 months in a word segmentation task. The change was in the initial syllable of a trochaic word (*Mommy/Tommy*), i.e. in the most salient syllable in English. Could it be that French-learning 5-month-olds react to a consonant change in word-final position, which is the most salient one in French?

To explore the effect of the acoustically more salient syllable in early processing of consonants in first familiar word-forms, the following experiment tests 5-month-old French-learning infants' preference for their correctly-pronounced versus a mispronounced version of their own name, consisting in a minimal change of the onset consonant of the final syllable of the word. As in Experiments 4-7, we used HPP. This experiment did not include a control group (to be tested in the future), thus all infants were presented with repetitions of their own correctly-pronounced name (CPs) on half of the trials, and repetitions of their own mispronounced name (MPs) on the other half of the trials.

If 5-month-olds are sensitive to a consonant change occurring at the onset of the *final* syllable of their own name, then they should exhibit a preference for CPs over MPs (e.g., a baby named *Lénie* should prefer listening to *Lénie* over *Lémie*) and thus demonstrate that in the final position, a consonant mispronunciation elicits a significant reaction in 5-month-olds. This would show that 5-month-olds are sensitive to a consonant change depending on its position within words.

II.2.a.ii. Methods

Participants

All 18 participants were healthy French-learning monolingual 5-month-old infants (***Table III. 3***.) whose names started with a consonant and had at least two syllables. Only infants without nicknames and who were frequently called by their own name were included. Three additional infants were tested and their data excluded due to fussiness (2) or experimenter error (1). The group will be completed in the future to include 30 infants.

Group (all n = 18)	Stimuli (example)	Infant's name (example)	Age in days mean (SD)	# girls/boys
	e.g. Lu**c**as vs. Lu**g**as	Lucas	166 (6.9)	8/10

Table III. 3. Experiment 8. Participant information, illustration of the experimental design, and examples of stimuli.

#	Feature	CP	MP	Contrast
1	PLACE	Juliette	Juriette	l/R
2	PLACE	Milan	Miran	l/R
3	PLACE	Tristan	Trispan	t/p
4	PLACE	Lénie	Lémie	n/m
5	PLACE	Shaïna	Shaïma	n/m
6	PLACE	Romain	Ronain	m/n
7	PLACE	Lucas	Lutas	k/t
8	PLACE	Lucile	Luchile	s/ ʃ
9	VOICING	Tennessee	Tennezee	s/z
10	VOICING	Batiste	Badiste	t/d
11	VOICING	Lucas	Lugas	k/g
12	MANNER	Suzanne	Sudanne	z/d
13	MANNER	Salomé	Salobé	m/b
14	MANNER	Louna	Louza	n/z
15	MANNER	Maxime	Maktime	s/t
16	MANNER	Maxime	Maktime	s/t
17	MANNER	Juliette	Juziette	l/z
18	MANNER	Timothée	Timossé	t/s

Table III. 4. Experiment 8. Details of stimuli presented (in columns): for each of the 18 participants, the phonetic feature contrasting the CP and MP names, the CP name, the MP name, the phonetic contrast (in IPA).

Stimuli

Each of the 18 participants heard repetitions of a pair of stimuli corresponding to their CP and MP names (***Table III. 4.***). MPs consisted in changing the onset consonant of the final syllable of the name. Only one-feature phonetic changes were used (place, voicing or manner). The use of the infants' own names yielded 17 stimuli pairs (due to two infants having the same name), allowing us to test 13

different consonant contrasts, thus yielding generalizable results. As in the main study, all changes were native French minimal contrasts according to Dell (1985), and were expected to be discriminable by French-learning infants irrespectively of their age. Eight infants were tested on a place-of-articulation change, 3 on a voicing change, and 7 on a manner-of-articulation change. Upon the completion of the study, 10 infants will have been tested in all three conditions. Fifteen names were bisyllabic and 3 trisyllabic. As in the main study, the same female native French speaker recorded 15 tokens of each CP and MP names. For both CPs and MPs, two files including the 15 tokens were made, the second file presenting the same tokens in reversed order. All sound files lasted 24 seconds.

Procedure

Same as in the main study.

II.2.a.iii. Results

Mean listening times (LTs) to the CP and MP names were calculated for each infant. Group averages are presented in ***Figure III. 3.*** The LT means were M_{CP} = 12.65 s (*SD* = 4.35 s) and M_{MP} = 13.61 s (*SD* = 4.43 s), the estimate of *LTdiff* was was *M* = -0.96 s (95% *CI* = [-3.42, 1.50]) and Cohen's *d* = .19 (Cohen's *ds* are calculated with the difference of means as numerator and the standard deviation of the paired differences as denominator). A t-test comparing LTs for CPs and MPs failed to reach significance (*t*(17) = -.82 , *p* = .42). Longer LTs for CPs over MPs were found for 7 of the 18 test infants (binomial test, *p* = .88). To further explore the effect of the different features, a two-way ANOVA on LTs with the factors of pronunciation and features (Place/Manner/Voicing) was conducted. No effect or interaction reached significance (all Fs < 1). However, these null effects need to be cautiously interpreted given the small sample size available at present (to be completed to 10 infants per feature).

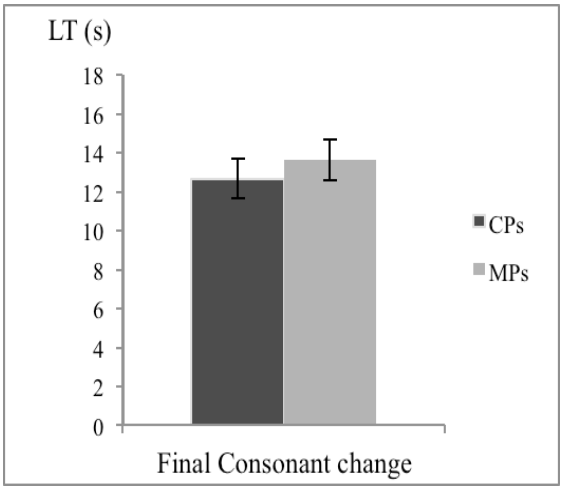


Figure III. 3. *Experiment 8. Results. Mean listening times (and SEs) to the CP and MP names differing by the onset consonant of the final syllable.*

II.2.a.iv. Discussion

In Experiment 8, French-learning 5-month-olds failed to show sensitivity to a minimal consonant mispronunciation occurring at the onset of the final syllable of their own name. Based on previous reports we assumed that the final syllable is longer and louder than the initial syllable in French words presented in isolation, thus that the consonant change is more salient in Experiment 8 than it was in Experiment 4 (i.e. initial consonant change). This assumption will have to be confirmed in a future acoustic analysis of our stimuli once the study is completed. Then, the same regression analyses as done in the main study (Exp. 4 -7) will be performed to explore the correlation between infants' performance and the size of the acoustic change in their name.

Contrary to previous data in French-learning infants in familiar word preference studies using HPP at 11 months (Hallé & de Boysson-Bardies, 1996; Vihman et al., 2004), and at 14, 18, and 22 months in word recognition studies using the IPL (Jöhr, 2014; Zesiger & Jöhr, 2011), the present preliminary results suggest that at 5 months, French-learning infants are not more sensitive to a consonant change in a familiar word when it occurs in the more salient position.

III. 2. b. EXPERIMENT 9: Initial change in French-learning bilinguals

In the last experiment of the present dissertation, we report data from a group of bilingual infants who were tested in parallel to the monolingual participants of the main study. Indeed, many parents of bilingual families are contacted through our mailing recruitment based on birth lists. Although bilingualism was not an issue at the core of the present dissertation, these bilingual infants were tested and provide the opportunity to address a topic of great interest, i.e. phonological and lexical acquisition in bilingualism, which has never been investigated through own name recognition.

II.2.b.i. Introduction

Bilingual infants exposed to two linguistic systems from birth have to cope with a greater degree of phonetic and acoustic variability in their input, which impacts significantly their linguistic achievement (Barac & Bialystok, 2012; Fabiano-Smith & Barlow, 2010; Sundara, & Scutellaro, 2011). Some infant studies explored the perceptual attunement of bilinguals to their native phonology and revealed different developmental trajectories than in monolinguals (Bosch & Sebastián-Gallés, 2003a, 2003b; Sebastián-Gallés & Bosch, 2009; Burns, Yoshida, Hill, & Werker, 2007). Some results also suggest that bilingual acquisition influences infants' ability to use phonetic detail in words (Fennell, Byers-Heinlein, & Werker, 2007; Mattock et al., 2010; Fennell & Byers-Heinlein, 2014; Havy, Bouchon & Nazzi, in revision). Here, we are interested in the effect of the bilingual dual input on phonetic processing in early word form recognition.

In monolingual infants, phonological acquisition, whereby the ability to discriminate phonetic contrasts attunes to the phonemic distinctions of the native language, starts during the first year, around the age of 6 months for vowels and 9-10 months for consonants (Kuhl, et al., 1992; Werker & Tees, 1984; Polka & Werker, 1994; Rivera-Gaxiola, Silva-Pereyra, & Kuhl, 2005). In bilingual infants, although phonological acquisition starts around the same age, the acquisition of each phoneme seems to vary depending on its presence in one or both languages, and on the way this phoneme is realized in the two languages. Indeed, some phonetic contrasts are present only in one of bilinguals' languages, as would be the case for the /r/-/l/ contrast for an English-Japanese bilingual. Other contrasts appear in both languages, e.g. /b/-/p/ exists both in French and English, and contrasts the minimal pairs *ball-poll* in English, and *balle-pâle* in French. However, in this last case, the acoustic realization of this contrast differs in the two languages. Indeed, the VOT values of stop consonants /p/ and /b/ are similar among Romance languages (French, Spanish, Italian and Portuguese), but differ between Romance and Germanic languages (German and English). Consequently, a /p/ realized with the VOT value of a Romance language could be confounded with a /b/ produced by an English speaker. Consequently, two different /p/-/b/ distinctions have to be built in the acoustic space of French-English bilinguals given that the VOT boundaries between these categories are not aligned in the two languages, whereas in French-Italian bilinguals /p/ and /b/ would be distinguished based on acoustic properties common to both their languages. Importantly, this acoustic congruence/incongruence of the contrasts in the two languages seems to have an impact on their acquisition trajectory in bilinguals (Bosch & Sebastián-Gallés, 2003a, 2003b; Sebastián-Gallés & Bosch, 2009; Burns et al., 2007).

When the contrast is present in both languages and realized with similar acoustic boundaries, the acquisition may follow the same time-course in bilinguals and in monolinguals. For instance, the /e/-/u/ vowel contrast, present in Catalan and Spanish is similarly discriminated by Catalan and Spanish monolinguals and Spanish-Catalan bilinguals at 4 and 8 months (Sebastián-Gallés & Bosch, 2009). When the contrast is present in one language and absent from the other, its pattern of development seems to follow a U-shaped curve in bilinguals, instead of a linear one in monolinguals. For instance, the /e/-/E/ vowel height contrast, present in Catalan but not in Spanish, is similarly discriminated by Catalan and Spanish monolinguals and Spanish-Catalan bilinguals at 4 months. However, at 8 months, unlike Catalan monolinguals, the bilinguals cannot discriminate it, and they succeed again at 12 months (Bosch & Sebastián-Gallés, 2003a). However, the cognitive demands of the task also appears to have an impact on these results, since Spanish-Catalan bilinguals discriminated /e/-/E/ at 8 months using an anticipatory looking paradigm (Albareda-Castellot et al., 2011), when they had failed when using the familiarization paradigm (Bosch & Sebastián-Gallés, 2003a). Similarly, for the consonant contrast /s/-/z/, which exists in Catalan but not in Spanish,

Catalan monolinguals succeeded at 6, 8 and 12 months, Spanish monolinguals only at 6 months, and Spanish-Catalan bilinguals at 6 and 12 months, but not at 8 months (Bosch & Sebastián-Gallés, 2003b). When the contrast is present in both languages but with misaligned acoustic/phonetic category boundaries, one member of the contrast may be learned later than the other (French-English bilinguals discriminate the French realization of /p-b/ before the English one, Burns et al., 2007) or follow a U-shaped development (Spanish-Catalan bilinguals discriminate /o-u/ at 4 and 12 months, not at 8 months, Sebastián-Gallés & Bosch, 2009). While more research is needed in this domain, it nevertheless appears that bilinguals learn some contrasts more easily than others depending on their distribution within the acoustic spaces of their two languages.

The acoustic congruence/incongruence of the phonological contrasts in the two languages of a bilingual infant is also crucial at the lexical level, particularly for processing phonetic detail in words (Mattock, et al., 2010; Fennell & Byers-Heinlein, 2014; Havy et al., in revision). Monolingual infants acquire their first words around 4.5 to 6 months (Bergelson & Swingley, 2012; Tincoff & Jusczyk, 1999, 2012). However, the impact of bilingualism at the lexical level has mainly been investigated in toddlers at 16-20 months in word learning tasks. By 16-17 months, monolinguals exhibit fast mapping abilities and are able to process relevant phonetic detail in computationally demanding word-learning tasks (Havy & Nazzi, 2009; Nazzi, 2005; Pater, Stager, & Werker, 2004; Stager & Werker, 1997; Werker, Fennell, Corcoran, & Stager, 2002). To explore how similarity/difference in the realization of a phonological contrast across languages affects lexical processing, Fennell et al. (2007) familiarized infants to two objects paired with two minimally different words (e.g., object A = /bih/ - object B = /dih/). Results show that before 20 months both English-French and English-Chinese bilinguals fail to react to the mismatch (e.g., object A = /dih/), while English monolinguals succeed at 17 months (Werker et al., 2002). Thus between 17 and 20 months, bilinguals still seem to have difficulties in using relevant phonetic detail during word learning. Interestingly, the acoustic realization of the contrast was misaligned in English and French, and aligned in English and Chinese. However, both bilingual groups failed in processing phonetic detail in newly learned words, thus this study failed to show an effect of similarity/difference in the realization of a phonological contrast across languages. Similarly, difficulties in processing the relevant phonetic detail of known words were found in Catalan-Spanish bilinguals at an age (18 months: Ramon-Casas, Swingley, Sebastián-Gallés, & Bosch, 2009; adults: Sebastián-Gallés, Echeverria, & Bosch, 2005), when they are not found in monolinguals anymore (e.g., Swingley & Aslin, 2000; Mani & Plunkett, 2007, 2010). However, two recent studies investigating the same issue revealed that 17-month-old English-French bilinguals were able to acquire pairs of pseudowords differing on their initial consonant (/kem/-/gem/, Fennell & Byers-Heinlein, 2014; /bos/-/gos/, Mattock et al., 2010). Therefore, the role that

similarities/differences in the phonological properties of the two languages may play at 16-17 months on bilingual infants' ability to process the detail in words is unclear.

One recent study looked at the impact of similarity/difference in the realization of phonetic contrasts on bilingual infants' word learning skills at 16 months and especially on their sensitivity to minimal consonant changes in newly learned words (Havy, Bouchon, & Nazzi, in revision). In an interactive word-learning task, French-learning bilingual infants were presented with two new objects paired with two minimally different words and then asked to take one of them (as done successfully by 16 month-old monolingual infants, Havy & Nazzi, 2009). The minimal change involved either a place of articulation (*/pyf/-/tyf/, /dul/-/gul/, /beji/-/ deji/, /tize/-/kize/)* or a voicing contrast (*/paS/-/baS/, /koet/-/goet/, /tola/-/dola/, /pivA)/-/bivA/*) in stop consonants. All contrasts existed in both languages of all infants, but depending on their L2 (i.e. second language) their acoustic realization was either relatively similar ("similar contrast" group, L2 being Italian, Spanish or European Portuguese) or clearly different ("different contrast" group, L2 being English or German). As previously mentioned, unvoiced consonants in French overlap with voiced consonants in English and German, leading to confusions between /p, t, k/ in Romance languages and /b, d, g/ in Germanic languages. Results show that as monolingual infants, the "similar contrast" bilingual infants successfully learned the words, but the "different contrast" bilinguals performed at chance, despite the same level of exposure to French in both groups. These findings suggest that at 16 months, processing the one-feature phonological contrast that distinguishes two words depends crucially on the similarity of the acoustic realization of this contrast in the infants' two languages.

However, bilinguals' ability to process phonetic detail in words has never been investigated at the onset of lexical acquisition. Therefore in the current experiment, 5-month-old bilingual infants were tested on their sensitivity to an initial consonant change in their own name, as monolingual infants were in Experiment 4 (Consonant change condition of the main study). As monolinguals, bilinguals might fail since the literature review suggests that they often perform worse than monolinguals. However, the literature reveals the presence of task effects, so a different lexical task based on own name recognition might reveal some unseen sensitivities in bilingual infants. Moreover, bilinguals' difficulties in discriminating phonetic contrasts seem to appear at 8 months for vowel contrasts (Bosch & Sebastián-Gallés, 2003a, 2003b; Sebastián-Gallés & Bosch, 2009; Albareda-Castellot et al., 2011) and between 6-12 months for consonant contrasts (6-8 months: Burns et al., 2007; 10-12 months: Bosch & Sebastián-Gallés, 2003b). Therefore, they may still be sensitive to minimal phonetic changes in a familiar word at 5 months.

As in Experiment 4 (Consonant change), we used HPP. The rest of the procedure was identical to Experiment 8. If 5-month-old bilinguals are sensitive to a consonant change occurring at

the onset of the *initial* syllable of their own name, then they should exhibit a preference for CPs over MPs and thus demonstrate that contrary to monolinguals, a consonant mispronunciation elicits a significant reaction in bilinguals at 5 months.

II.2.b.ii. Methods

Participants

All 18 participants were healthy bilingual 5-month-old infants learning French and another language (later on, L2, ***Table III. 5.***) and for 2 of them an L3. The percentage of L2 exposure (later on %L2) in the group varied between 20 and 75 %. Only infants without nicknames and who were frequently called by their own name were included. All children had names starting with a consonant. Five additional infants were tested and their data excluded due to fussiness (4), or having two consecutive trials with insufficient looking times (1).

Group (n = 18)	Stimuli (example)	Infant's name (example)	Age in days (SD)	# girls/boys
	e.g. Marius vs. Narius	Marius	167 (7.2)	9/9

Table III. 5. Experiment 9. Participant information, illustration of the experimental design, examples of stimuli and participant name.

#	Feature	CP & Test infant's name	MP	Contrast CP/MP	L2	L2%
1	PLACE	Marius	Narius	m/n	Romanian	50
2	PLACE	Chouchou	Souchou	ʃ/s	Romanian	55
3	PLACE	Rania	Lania	R/l	Tunisian Arabic	20
4	PLACE	Sarah	Farah	s/f	Koyaga (Ivorian dialect)	60
5	PLACE	Nino	Mino	n/m	Italian	25
6	PLACE	Kery	Tery	k/t	Swahili & English	30
7	VOICING	Kiara	Giara	k/g	Spanish	40
8	VOICING	Kenzo	Genzo	k/g	Kabyle	30
9	VOICING	Camila	Gamila	k/g	Farsi & Arabic	30
10	VOICING	Sacha	Zacha	s/z	Russian	30
11	VOICING	Tao	Dao	t/d	Mauritian Creole	20
12	VOICING	Jules	Chules	ʒ/ʃ	Chinese	50
13	MANNER	Lior	Zior	l/z	Hebrew	20
14	MANNER	Maxime	Baxime	m/b	West Indian Creole	30
15	MANNER	Maeden	Baeden	m/b	Madagascan	25
16	MANNER	Noham	Zoham	n/z	Bengali	30
17	MANNER	Nissai	Zissai	n/z	Swedish	75
18	MANNER	Minhchau	Binhchau	m/b	Vietnamese	45

Table III. 6. Experiment 9. Details on bilingual participants bilingualism, name and and stimuli.

Stimuli

Each of the 18 infants heard repetitions of a pair of stimuli corresponding to their CP and MP names (***Table III. 6***.)***.*** The MP name always consisted of a French one-feature consonant change in the initial position derived from the CP. All 18 infants were tested on a native French contrast (Dell, 1985). The use of infants' own names yielded 17 stimuli pairs (due to two infants having the same name), allowing us to test 12 different consonant contrasts, thus yielding generalizable results. Six infants were tested on a place-of-articulation change, 6 on a voicing change, and 6 on a manner-of-articulation change. For all infants, the same female native French speaker recorded 15 tokens of the CP and MP names. For both CPs and MPs, two files including the 15 tokens were made, the second file presenting the same tokens in reversed order. All sound files lasted 24 seconds.

Procedure

Same as in main study.

II.2.b.iii. Results

Mean listening times (LTs) to the CP and MP names were calculated for each infant. Group averages are presented in ***Figure III. 4***. The LT means were M_{CP} = 12.87 s (*SD* = 3.26 s) and M_{MP} = 13.4 s (*SD* = 3.33 s), the estimate of *LTdiff* was *M* = -4.94 s (95% *CI* = [-1.86, 8.72]) and Cohen's *d* = .18. A t-test comparing LTs for CPs and MPs failed to reach significance (*t*(17) = -.76 , *p* = .46). Longer LTs for CPs over MPs were found for 6 of the 18 test infants (binomial test, *p* = .95). To further explore the effect of different features, a two-way ANOVA on LTs with the factors of pronunciation and features (Place/Manner/Voicing, all n = 6) was conducted. A main effect of Feature was obtained (*F*(2,15) = 3.73, *p* = .048), due to longer LTs in the manner condition (*M* = 15.4s; *SE* = 0.9s) than in the place (*M* = 12.5s; *SE* = 1.4s) and voicing conditions *(M* = 11.4s; *SE* = 1.1s). However, neither the main effect of pronunciation nor the interaction reached significance. Again as for Experiment 8, caution is required when interpreting these findings due to the small sample size.

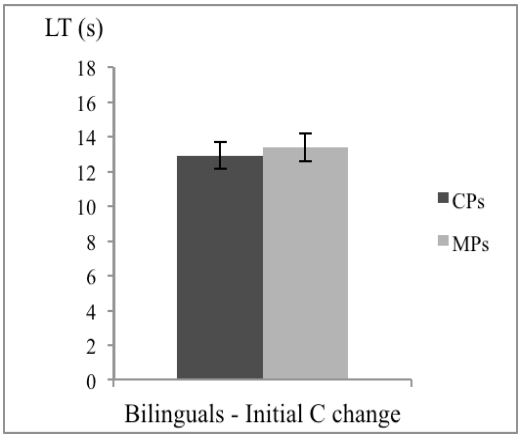


Figure III. 4. *Experiment 9. Results. Mean listening times (and SEs) to the CP and MP names differing by their initial consonant.*

II.2.b.iv. Discussion

The results of Experiment 9 fail to show that 5-month-old bilingual infants react significantly to a consonant change in their own name. Hence, these findings are similar to those of the monolinguals in Experiment 4, and show that at 5 months bilinguals do not have an advantage over monolinguals in processing phonetic details in words. To assess the effect of bilingualism on vowel changes, a future experiment will need to assess bilinguals' sensitivity to a vowel change in their own name. This would allow us to evaluate whether they succeed like the monolinguals in Experiment 5 or not, and thus have a vowel bias as the monolinguals did.

In the literature review on bilingual acquisition, we saw that the developmental trajectory of phoneme discrimination depended on several factors, firstly the presence or not of the contrast in both languages, secondly on the acoustic similarity/difference in the realization of the phonemes. In our group, based on these factors, we could distinguish 3 subgroups of bilinguals: an "L1 contrast only" group, an "L1-L2-Different" group and an "L1-L2-Similar" group. More infants should be tested to allow us to break down our data into these three categories and determine whether infants in these 3 subgroups have different name recognition abilities. In addition, the consonant and vowel categories are well represented in various phonetic contrasts tested. Havy et al. (in revision) showed that 16-month-old bilinguals learn pairs of words differing by a one-feature consonant change only if the contrast is similarly realized in both of their languages. This raises the possibility that at the onset of lexical acquisition the "L1-L2-Similar" bilingual group could be more sensitive to a phonetic change than the other two groups.

III. 3. GENERAL DISCUSSION

To explore the developmental origin of the consonant bias in the first year of life, the experiments of the current chapter investigated the role of consonants and vowels in a familiar word recognition task in French-learning infants. In our first chapter using NIRS, we failed to observe a consonant bias in lexical processing at birth contrary to Nespor et al. (2003), thus supporting the hypothesis that it emerges later, either as a by-product of early acoustic/phonological processing (Floccia et al., 2014) or lexical acquisition (Keidel et al., 2007). At 5 months, infants' phonetic discrimination abilities are only starting to attune to native categories, and are more advanced for vowels than for consonants (Kuhl et al., 1992; Werker & Tees, 1984). Their pre-lexical knowledge consists in the recognition of a few sound patterns of words, and there is some evidence of word

comprehension (Tincoff & Juszcyk, 1999; 2012; Mandel et al., 1995; Bergelson & Swingley, 2012). In Experiments 4 to 9, we used HPP to evaluate French-learning 5-month-olds' reactions to minimal phonetic changes in one of their first words, their own name (Mandel, et al., 1995).

The results of Experiments 4 to 7 first show that monolingual French-learning 5-month-olds fail to show a preference for their correctly-pronounced name over an initial consonant mispronunciation (*e.g. Remi/Lemi,* Exp. 4), while in the vowel change condition, infants preferred their CP names over initial vowel MPs (e.g. *Alix/Elix,* Exp. 5), independently of the features tested. Thus, French-learning infants recognize their name over a minimal vowel mispronunciation at 5 months, and not over a consonant mispronunciation. Although these findings would have to be confirmed using other familiar word-forms at 5 months, given the possible special status of the infant's name, they suggest that infants exhibit a vowel bias at the onset of lexical acquisition, instead of the consonant bias predicted by Nespor et al.'s (2003).

Moreover, acoustic analyses of the contrasting consonants and vowels of the stimuli in the main study (Exp. 4-7) provide a possible explanation for the infants' behavior. Firstly, consonants are softer and shorter than vowels, thus quantitatively less important in the signal, while on the spectral dimension consonant contrasts were found to be more distinct than vowel contrasts. Therefore, the advantage of vowels on quantitative dimensions may underlie infants' sensitivity to a vowel change. Secondly, regression analyses showed that infants' performance was predicted by the spectral distance between vowels in Experiment 2, but not by the spectral distance between consonants in Experiment 1. This indicates that the qualitative acoustic property that makes consonants more informative than vowels may actually not be a relevant cue for infants in case of a consonant change, while predicting infants' behavior in the case of a vowel change. This finding suggests that 5-month-olds' lack of reaction to a consonant change in familiar words may be due to their sensitivity to the acoustic saliency of the longest and loudest portions of words.

Experiment 8 further explored French-learning infants' early processing of consonants in familiar word forms by changing a consonant in the acoustically more salient position of French words. Indeed, in multisyllabic French words, final syllables are longer and louder than initial syllables, thus the phonetic detail of final syllables may be more precisely encoded or processed. Language-specific positional effects in consonant processing have been found at 11 months in word recognition tasks in French and English (Hallé & de Boysson-Bardies, 1996; Vihman et al., 2004; Jöhr, 2014; Zesiger & Jöhr, 2011), indicating that infants may preferentially process consonant information in the more salient syllable of words. French-learning 5-month-olds were thus tested on the impact of a minimal MP of the onset consonant of the final syllable of their own name (e.g. *Salomé/Salobé;*

Louna/Louza). As in Experiment 4, infants failed to show a preference for their CP name over a consonant MP name, independently of the consonantal feature (place, voicing and manner) tested. Thus, contrary to previous results in older infants, 5-month-olds failed to react to an initial as well as a final consonant in their own name. Experiment 8 thus confirms that infants seem to rely more on vowel than on consonant information in their first familiar word-forms. Whether a vowel change in the final syllable would be detected should be addressed in the future, and although it is unlikely given the results of Experiment 5, it remains possible that infants would not react to a vowel change in the final syllable of their own name. Such a finding would put in perspective the interpretation that infants process vowels preferentially over consonants in familiar words at 5 months. Indeed in Experiment 5, all names were vowel-initial words and the initial vowel was almost always also a syllable on its own (except in 5 out of 30 cases, *Arthur, Alba, Arsene, Esther*, *Erwan),* so that infants may have reacted to the vowel change because, unlike in Experiment 4, it also corresponded to a whole syllable change. However, if 5-month-olds were also sensitive to a vowel change in the final syllable of their own name, as suggested by the results of Experiment 5, it would show that the vowel bias is independent of syllable saliency in French-learning 5-month-olds.

So far, the results of Experiment 8 suggest that infants fail to react to a consonant change in their own name even when it is in an acoustically more salient position. Once Experiment 8 is completed to include more consonant changes and vowel changes as well, the stimuli should be acoustically analyzed as they were for Experiments 4 and 5, and it will then be possible to explore the link between individual performance and the acoustic properties of the change (in duration, intensity and spectral measures).

To investigate the impact of bilingual input on early phonological and lexical abilities, Experiment 9 tested bilingual infants' sensitivity to a consonant change in their own name, as done for monolinguals in Experiment 4. Bilingual infants learn some of their native contrasts more easily than others depending on the presence/absence of the phonetic contrast in their two languages, and the similarity/difference of their acoustic realization in the two languages (Bosch & Sebastián-Gallés, 2003a, 2003b; Sebastián-Gallés & Bosch, 2009; Burns et al., 2007). Moreover, the impact of bilingualism on infants' ability to use phonetic detail in lexical processing has also been shown in the second year of life (Fennell, Byers-Heinlein, & Werker, 2007; Mattock et al., 2010; Fennell & Byers-Heinlein, 2014; Havy, et al., in revision). In Experiment 9, bilingual infants failed to detect an initial consonant change in their own name. If this result were confirmed, it would support the view that 5-month-olds fail to perceive a consonant change in their own name, which is one of their first words. This experiment will be completed to further explore the early role of acoustic similarity in the realization of a phonetic contrast (Havy, et al, in revision).

To conclude, the experiments presented in this second experimental chapter fail to show a consonant bias at the onset of lexical acquisition. Rather, they suggest that vowels are preferentially processed in one of infants' first words. Indeed, 5-month-old French-learning infants are sensitive to an initial vowel change, but not to an initial (less salient) or final (more salient) consonant change, in their own name. This vowel bias, rather than the consonant bias predicted by Nespor et al. (2003), is independent of the phonetic features used. Overall we can conclude that when recognizing familiar words, 5-month-old French-learning infants are not processing consonants and vowels according to their functional specificities, as was found in French toddlers and adults. The better informativeness of consonants at the lexical level thus has to be learned between 6 and 11 months (Poltrock & Nazzi, in revision) through additional language exposure. It is not yet known how much more phonological and lexical knowledge infants have to acquire before they can preferentially use consonants over vowels to recognize and learn words, and what the mechanisms underlying this developmental change are.

CHAPTER 4 – GENERAL DISCUSSION

This dissertation has been dedicated to the investigation of the functional asymmetry between consonants and vowels in early speech processing. In particular, our experiments aimed at determining the potential bootstrapping power of the two biases proposed by Nespor et al. (2003) in their the initial C/V division of labor hypothesis. This hypothesis is based on the facts that acoustically, linguistically and psychologically consonant information is more linked to lexical aspects of language, while the grammar and structures of the language are more embedded in vowel information. Nespor et al. (2003) thus proposed that the consonant bias for lexical processing and the vowel bias for morphosyntax processing would be language-general mechanisms present from birth and thus prompting language acquisition initially. More recent infant research has provided crosslinguistic evidence that did not clearly support this hypothesis, and alternative accounts have thus been proposed for the origins of the biases (Floccia et al., 2014; Keidel et al., 2013). This dissertation therefore raised the following questions: Is the C/V functional asymmetry present at birth as a bootstrapping mechanism of language acquisition? Are infants able to use consonants preferentially when recognizing their first familiar word-forms at 5 months?

The general discussion will start by summarizing the answers that our data bring to the questions raised at the beginning of the dissertation. Further, interpretations will be drawn in the light of a model of early language acquisition, PRIMIR (Werker & Curtin, 2005). Finally, we will present specific ideas of future research to address some of the questions raised by this work.

IV. 1. Summary of findings

IV. 1. a. C/V asymmetry at birth?

The first experimental chapter attempted to answer the important question of the origins of the C/V functional asymmetry, by testing infants at birth. Although newborns do not analyze speech in terms of word constituents and syntactic rules, they have nonetheless demonstrated abilities to process speech in order to extract and memorize individual items, and to detect and learn structural patterns. In our NIRS experiments, the presence of the consonant bias for such lexical-like processes and of the vowel bias for precursors of syntactic processing were investigated directly and simultaneously for the first time in the newborn brain. To do so, adaptations of the original Gervain et al. (2008) paradigm that revealed the ability to detect repetition-based regularities carried by *syllables* (e.g. *mubaba vs. mubage*) were conducted.

However, as a preliminary study, we first needed to ensure that the results of Gervain et al. (2008) generalized to the methodological requirements of our research question. Experiment 1 thus replicated the original syllable-repetition ABB vs. ABC paradigm (Gervain et al., 2008) with a new adapted linguistic material. This material minimized the acoustic and positional differences between consonants and vowels, and used a blocked instead of an interleaved presentation of the conditions. Experiment 1 thus serves as a baseline experiment before Experiments 2 and 3 actually compare the roles that consonants and vowels play in the detection of repetition-based regularities at birth. The results of Experiment 1 confirmed that the ABB and ABC grammars elicit different dynamics in the two hemispheres. In left fronto-temporal areas, we reproduced the early perceptual discrimination of the two grammars found in Gervain et al. (2008), with ABB giving rise to a greater response at the beginning of the experiment than ABC. In addition, the ABC grammar evoked a repetition enhancement effect over time, whereas the activity evoked by the ABB grammar remained constant. Right fronto-temporal areas showed neither initial discrimination, nor change over time to either pattern contrary to Gervain et al. (2008). These results first confirm that our new material and NIRS set-up (NIRx NIRS system) successfully elicited syllable repetition detection responses in speech processing areas of the newborn brain. Additionally, the findings of Experiment 1 compared to Gervain et al. (2008) contribute to the understanding of neural repetition effects in the NIRS signal at birth and in particular highlight the importance of variability in the stimuli on the neural time course response.

Experiments 2 and 3 compare *consonant* versus *vowel* repetition detection with the same NIRS system, material and blocked-condition design as in Experiment 1. These two experiments consist in comparing the processing of three (instead of two) different artificial grammars: the unstructured control grammar (ABC, e.g. *mulevi*) is compared to two grammars carrying a segmental repetition, ABBconsonant (e.g. *muleli*) and ABBvowel (e.g. *muleve*; later on, ABBc and ABBv respectively). As in Experiment 1 and Gervain et al. (2008), the experimental design in Experiment 2 induced structural extraction mechanisms, and should thus favor vowel-repetition processing according to Nespor et al. (2003). Experiment 3 presented the same grammars and stimuli as in Experiment 2, rearranged to induce an item-based mode of processing. Indeed, Experiment 3 presented the stimuli by token and should thus favor learning over consonants according to the functional hypothesis of Nespor et al. (2003). The results of Experiment 2 show that both ABBc and ABBv grammars are eliciting larger neural responses than the ABC unstructured condition suggesting that both consonant and vowel repetition-based regularities are detected. Experiment 2 thus fails to reveal differences in the newborn brain when listening to the two ABB grammars, which was predicted by the innate C/V functional asymmetry. The results of Experiment 3 fail to show any differential processing for any of

the three grammars. Hence, they fail to show differences in processing consonants and vowels in the so-called lexical context.

Overall, this experimental investigation of the origins of the C/V functional asymmetry at birth does not reveal any bias, either for vowels or for consonants. Mainly, after confirming that our new material elicited syllabic repetition detection in Experiment 1, we found that both consonant and vowel repetitions were discriminated from ABC unstructured sequences in the rule-learning context induced in Experiment 2. This result established for the first time that segment-level repetitions are detected in the newborn brain, although probably in a weaker fashion than syllables. Indeed ABBc and ABBv patterns elicited an increase in activity as opposed to the ABC condition but this distinction did not give rise to lasting effects during the time-course of the experiment as they had in Experiment 1. Second, both consonants and vowels appear to be relevant units of repetition-based rule extraction and learning at birth. Thus, the advantage for vowels over consonants found in previous studies (Bertoncini, et al., 1988; Benavides-Varela et al., 2012) is not present in our results, which goes against a general acoustically driven vowel bias in early speech processing. However, in our study, stimuli were synthesized in such a way to minimize differences in acoustic salience between consonants and vowels (using continuous consonants, matched with vowels in length, intensity and pitch). Therefore, our results would have to be confirmed using more natural stimuli and a larger range of consonants. In relation to the origins of the C/V functional asymmetry, we found that at birth the two categories of speech-sounds may be providing equivalent input to syntax-like processes at least, contrary to the predictions of the initial bias hypothesis (Nespor et al., 2003). Although our new paradigm, adapted to track the specificities of lexical- and syntax-like processing in the newborn brain, had been validated in baseline Experiment 1, our findings however suggest unexpected methodological limitations in our material that would have to be addressed in the future, particularly by increasing variability in the material.

To conclude on our first experimental chapter, we failed to find evidence of a vowel bias in repetition-based regularity rule-learning, or a consonant bias in a lexical-like item memorization context in the newborn brain. Having failed to show the biases proposed by Nespor et al. (2003) at birth, we therefore aimed to find out whether a consonant bias in lexical processing could be observed before 6 months.

IV. 1. b. C/V asymmetry in lexical processing at 5 months?

Studies of the C/V functional asymmetry in lexical tasks reveal that a consonant bias is well attested in adults, and that cross-linguistic differences appear in infants and toddlers (Højen & Nazzi, in revision; Jöhr, 2014; Nazzi, et al., 2009; Floccia et al., 2014; Mani & Plunkett, 2007, 2010; Poltrock

& Nazzi, in revision; Zesiger & Jöhr, 2011). To determine the early presence of the consonant bias in French-learning infants, Experiments 4-7 thus explored whether the greater lexical informativeness of consonants would serve infants in processing their first familiar words. The method consisted in measuring their reactions to a mispronunciation in their own name, one of their first familiar word-forms (Mandel et al., 1995) using the HPP. Infants were presented with repetitions of their own correctly pronounced name (CPs) and repetitions of a mispronounced version of their name (MPs). Their preference for CPs over MPs was compared whether the mispronunciation was consonantal, in Experiment 4, (e.g. *Sanjay/Tanjay*), or vocalic, in Experiment 5 (e.g. *Esther/Isther*). In two additional experiments, the same sets of CPs/MPs stimuli pairs were presented to other infants having different names: in Experiment 6, baby Margot was tested on her preference for *Sanjay/Tanjay*, while in Experiment 7, baby Adrien heard *Esther/Isther.*

In Experiment 4 (Consonant change), infants did not show a preference for their CPs over *initial consonant* MPs, and behaved as infants in control Experiment 6. In Experiment 5 (Vowel change), infants preferred CPs to *initial vowel* MPs. Crucially this difference was not observed in control Experiment 7, where infants showed a small preference for MPs, thus attesting the quality of the recordings of CPs and MPs. The results of Experiment 5 and 7 suggest that 5 month-olds are sensitive to a vowel mispronunciation in one of their first familiar word-form, their own name. These findings were found independently of the consonantal (place, voicing, manner) and vocalic (place, roundedness, height) features tested. The findings of Experiments 4 to 7 together indicate that French-learning infants exhibit a vowel bias at the onset of lexical processing, thus contradicting the initial consonant bias hypothesis (Nespor et al., 2003). Taken together with evidence of a consonant bias using a similar method at 11 months (Poltrock & Nazzi, in revision), these results suggest that French-learning infants learn during the second semester of life that consonants are more informative than vowels when processing words.

Moreover acoustic analyses of the stimuli first showed that, as expected, vowels were acoustically more salient than consonants both in duration and intensity. Secondly, consonant contrasts were found to be spectrally more distinct than vowel contrasts. These findings suggest that acoustically, vowel contrasts in our stimuli were indeed more salient on the intensity and duration dimensions, but crucially consonant contrasts were more distinct on the spectral dimension. Thirdly, regression analyses were conducted on infant's performance (i.e. their looking time preference for CPs over MPs) and the three acoustic distances between contrasted phonemes of CPs and MPs (duration, intensity and spectral distances). They show that in the vowel change experiment (Exp. 5) infants' preference is best predicted by the spectral distance, and to a lesser degree by the duration distance between the contrasted vowels. In the consonant change experiment (Exp. 4), although

variables displayed similar variances as in Experiment 5, none of the acoustic predictor could be linked to infants' performance. Thus, 5 month-olds' distinction of vowel MPs is modulated by acoustic factors, first spectral distance, second duration. However, they fail to distinguish consonant MPs in spite of their higher spectral distinctiveness. These findings indicate that 5-month-olds may already be processing consonants and vowels as different categories, but show more sensitivity to a vowel change than to a less salient consonant change in one of their first familiar word (their own name), suggesting they have not yet learned the functional roles of consonants and vowels in language processing.

Following-up on the absence of sensitivity to a consonant change in Experiment 5, Experiment 8 attempted to facilitate the detection of the mispronunciation. Indeed the final syllable of French words is longer and louder than the initial syllable, and previous results revealed that French-learning infants are sometimes more sensitive to a change in the final than in the initial syllable of familiar words at 11 months (Hallé & de Boysson-Bardies, 1996; Vihman et al., 2004). Moreover, studies investigating the consonant bias from the age of 14 months could only observe it for the final syllable of words (Zesiger & Jöhr, 2011; Jöhr, 2014). Thus, in Experiment 8, the infant's own name was changed in the acoustically more salient position of French words, on the onset consonant of the final syllable (e.g. *Salomé/Salobé*). The results show that infants do not prefer CPs over final consonant MPs, independently of the consonantal feature (place, voicing and manner) tested. Together with Experiment 4, this result suggests that French-learning 5-month-olds are not more sensitive to a consonant change in a familiar word when it occurs in a more salient part of the word.

To investigate the impact of bilingualism on the role of consonants in recognizing familiar word-forms at 5 months, Experiment 9 tested bilingual infants in the same condition as monolinguals in Experiment 4. Bilingual toddlers' performance in using phonetic detail in words has yielded mixed results. They have been found to perform sometimes worse (Fennell et al., 2007; Havy, Bouchon & Nazzi, in revision; Ramon-Casas, et al. 2009) and sometimes as well as monolinguals (Mattock et al., 2010; Fennell & Byers-Heinlein, 2014; Havy et al., in revision), depending on the contrast tested, the method used and their languages. Crucially, they seem to start having difficulties in the acquisition of phonology in the first year around 6-8 months (Bosch & Sebastián-Gallés, 2003a, 2003b; Sebastián-Gallés, & Bosch, 2009; Burns et al., 2007) depending on the contrast, and its acoustic realizations in the infant's two languages. Moreover, bilinguals had never been tested at the onset of lexical acquisition in a word recognition task based on their own name. Thus, in Experiment 9, we tested bilinguals on their preference for CPs over initial consonant MPs in their own name as in Experiment 4. Results showed that, like monolinguals, bilinguals failed to detect an initial consonant change in

their own name. In the future, Experiment 9 should include more infants to allow the analysis of data broken down according to the presence/absence of the tested contrast in the infants' two languages and to the similarity/difference in the acoustic realization in the two languages. Indeed Havy et al. (in revision) showed that 16-month-old bilinguals learn pairs of words differing by a one-feature consonant change only if the contrast is present and similarly realized in both of their languages. Our paradigm provides the means to test a large number of contrasts, therefore to better assess the role of these crucial factors in the early use of phonetic detail in words for bilinguals.

To conclude on our second experimental chapter, a vowel change in a familiar word-form elicited a reaction of strong preference for the correct word-form reaction, while a consonant change did not. Whether this consonant change occurred in the initial or in the more salient final syllable of the word did not seem to impact infants' performance. Therefore, contrary to the predictions of Nespor et al. (2003) and the data in French-learning infants from 11 months onward, our studies reveal a vowel bias in lexical processing in French-learning 5-month-olds. What conclusions can we draw from this new evidence? How does it combine with other evidence from different ages and language groups?

IV. 2. Bringing the pieces together

If, as proposed by Nespor et al. (2003), the consonant/vowel division of labor was innate and language-general it may be present before substantial language experience comes in. According to this hypothesis, the consonant and vowel biases could be part of the repertoire of perceptual constraints infants are born with. As such they would help newborns deciphering the complex acoustic speech signal. Alternatively, infants may need to gain experience with language before they are able to preferentially use consonants in lexical processing and vowels in morphosyntax processing, as predicted by alternative hypotheses of a consonant bias emergence (Keidel et al., 2007; Floccia et al., 2014). To decide between the two alternatives, we will now attempt to integrate our findings with the (few and sometimes preliminary) evidence available in infants in their first year.

IV. 2. a. Language-specific trajectories of C/V functional (a)symmetry

Our results at 5 months indicate that French-learning infants relied more on vowel information when processing a familiar word-form, suggesting that the consonant bias will have to be learned through more language exposure. Two alternative hypotheses have proposed that the consonant bias needs to be learned through language exposure. According to Keidel et al. (2007), infants need

to extract enough distributional information from a substantial lexicon to pick up that consonants are more informative to code the lexicon. The attribution of specific roles to consonants and vowels would thus be modulated by lexical language-specific properties of the input. Our data are also compatible with the acoustic/phonetic hypothesis (Floccia et al., 2014), according to which consonants and vowels would already be processed categorically due to their differing language-general acoustic properties (consonants often being shorter, softer, less periodic, less steady, and perceived more categorically than vowels, Repp, 1984, Pisoni, 1973). However, the attribution of specific roles to consonants and vowels would emerge early in development and would be modulated by acoustic/phonetic language-specific properties of the input. Our data alone do not allow disentangling these two alternative explanations of the emergence of a consonant bias, and are compatible with both the lexical hypothesis (Keidel, et al., 2007) and the acoustic/phonetic hypothesis (Floccia et al., 2014). In the following, we discuss other results that may provide more information concerning the developmental trajectory of the consonant bias in the first year.

First, in his unpublished dissertation, Hochmann (2010) reports results of Italian-learning 6-month-olds on the same novel word-learning task previously described (Hochmann et al., 2011). Results at 6 months are opposite to those at 12 months: infants consider two words sharing their vowels as more similar than two words sharing their consonants. Thus, in Italian, using the same novel word-learning task, 6-month-olds exhibited a vowel bias where 12-month-olds exhibited a consonant bias.

Second, a recent study investigated the role of consonants and vowels in word segmentation in 6- and 8-month-old French-learning infants using HPP (Nishibayashi, Goyet & Nazzi, 2014; Nishibayashi & Nazzi, 2014). Infants were familiarized to passages, containing two target words (CVs or CVCs) and then tested on their preference for lists of vowel and consonant minimal mispronunciations of the target words. Results first showed that both age groups recognize and prefer the words segmented out of the passages. However the two age groups gave opposite importance to the consonants and vowels composing these words. Indeed, 8-month-olds showed a preference for the vowel mispronunciations while 6-months-olds preferred the consonant mispronunciations of the words. Thus, in French, using the same segmentation task, 6-month-olds exhibited a vowel bias while 8-month-olds exhibited a consonant bias.

Together with those of Poltrock and Nazzi (in revision) at 11 months and ours at 5 months, the previous findings suggest that French- and Italian-learning infants switch from a vowel to a consonant bias in lexical processing between 6 and 8/11 months. According to the results of the French adaptation (Kern, 2007) of the MacArthur-Bates Communicative Development (Fenson et al.,

1994), the average comprehension vocabulary score is 16 words at 8 months of age. Therefore, we can already infer that Italian- and French-learning infants do not seem to need a sizeable lexicon for the consonant bias to emerge. What about other languages?

Experiments 4-7 are being replicated with 5-month-old English-learning infants (Bouchon, Delle Luche, Floccia & Nazzi, 2012). Preliminary results suggest that they are neither sensitive to a consonant change (e.g. *Timothy-Kimothy*) nor to a vowel change (e.g. *Amelie-Umelie*) in their own name. These findings are still under investigation, in particular to determine the acoustic distance in the English stimuli. However, if these results were confirmed, they would suggest that at 5 months English-learning infants do not exhibit the vowel bias observed in French-learning infants in the same task (Experiments 4-8) and in Italian-learning 6-month-old infants in a different lexical task (Hochmann, 2010).

Overall, according to the data available so far on infant, we can distinguish three possible patterns in the development of the C/V functional asymmetry.

- First, French/Italian-learning infants at 5-6 months are relying on the more salient and periodic portions of speech, i.e. vowels, rather than on the more transient and distinct sounds, i.e. consonants, at 8-12 months. This pattern is more compatible with the acoustic/phonetic hypothesis (Floccia, et al., 2014).
- Danish-learning infants are showing a vowel bias at 20 months (Højen & Nazzi, in revision). Given that in Danish, vowels are advantaged over consonants at many levels (acoustic/phonetic, phonemic, lexical), this finding is compatible with the acoustic/phonetic and lexical hypotheses. Evidence in younger and older age groups with Danish-learning infants would be necessary to make further hypotheses.
- On the contrary, English-learning infants at 5 months may not be relying on vowels preferentially. Thus, the higher acoustic salience of vowels over consonants (to be checked for the English stimuli) may not be driving their early use of consonants and vowels. The differential use of consonants in lexical processing seems to appear only at 30 months (Nazzi et al., 2009). Thus, consistently with the lexical hypothesis (Keidel et al., 2007), advanced lexical acquisition may be required for the consonant bias to emerge in English.

What can we infer from the evidence that French- and Italian- infants seem to follow the same path in the development of the differential use of consonants and vowels, while English infants do not show a differential processing before 30 months?

IV. 2. b. Potential language-specific factors involved

These language-specific trajectories could be due to numerous language-specific factors. Floccia et al. (2014) have listed potential factors varying across languages and that could contribute to emphasize or attenuate the distinction between vowels and consonants as follows: the realization of vowels and consonants may vary as a function of the number of vowels and consonants, the presence of vocalic or consonant reduction/lenition (cf. Danish), phonological short/long vowel contrasts, ambisyllabicity (cf. English), etc. In order to isdentify some of the possible candidates, we will detail factors that French and Italian have in common and separate them from English, as they may potentially contribute to the differential developmental trajectories observed in French/Italian on the one hand, and in English on the other hand.

French and Italian are both Romance languages, while English belongs to the class of Germanic languages together with German and Dutch. Thus, French and Italian are likely to share more typological features at the phonological and lexical levels. Among these features, **syllabic structure** is a relevant factor. Indeed, Romance languages tend to favor more simple syllabic structures than Germanic languages (Delattre, 1965; Ramus, Nespor & Mehler, 1999). French and Italian syllables are less variable and have clearer boundaries, with minimum ambisyllabicity (French: Delattre, 1965). English has extremely variable syllable structure (for instance *a* and *streams* are both a single syllable. Moreover, as a result of ambisyllabicity (i.e. consonants acting simultaneously as the coda of one syllable and the onset of the following syllable), intervocalic consonants of multisyllabic words can belong both to the first and second syllables (as in 'below', bel-low). Thus, young English-learning infants are exposed to more complex alternations of consonants and vowels in their input than French- and Italian-learning infants.

Moreover, syllable structure is linked to **rhythmic properties,** which infants are sensitive to from birth (Nazzi et al., 1998; Ramus et al., 2000). However, rhythmicity is more gradient than categorical. According to the proportion of vocalic intervals as a measure of linguistic rhythm distinction between syllable-timing and stress-timing (Ramus, et al., 1999; Grabe & Low, 2002), the three critical languages can be placed along a continuum ranging from most stress-timed to most syllable-timed as follow: English – French – Italian. Thus, French and Italian are more syllable-timed languages, while English is more stress-timed. This has been found to impact lexical segmentation abilities very early on. Indeed, many studies have found that rhythm is one of the cues that infants are learning to use between 6 and 8 months to segment word forms from fluent speech (e.g., Bosch, Figueras, Teixidó, & Ramon-Casas, 2013; Goyet, de Schonen & Nazzi, 2010; Goyet, Nishibayashi & Nazzi, 2013; Jusczyk, Houston, & Newsome, 1999; Kooijman, Hagoort & Cutler, 2005, 2009; Nazzi,

Iakimova, Bertoncini, Frédonie & Alcantara, 2006; Nazzi, Mersad, Sundara, Iakimova & Polka, 2014; Nishibayashi, Goyet, & Nazzi, 2014; Polka & Sundara, 2012). Nazzi et al. (2006) thus proposed that the rhythmic unit of the native language drives early segmentation (early rhythmic bootstrapping segmentation hypothesis). Accordingly segmentation is syllable-based in French, Spanish and Catalan (Bosch et al., 2013; Goyet et al., 2013; Nazzi et al., 2006, 2013; Nishibayashi et al., 2014) and stress-based in English (Jusczyk et al., 1999) and Dutch (Kooijman et al., 2005, 2009). Moreover, infants are able to generalize their language-specific segmentation procedure to a language with the same similar rhythmic organization (English-Dutch: Houston, Jusczyk, Kuijpers, Coolen, & Cutler, 2000) but not to a language with a different rhythm (English-French: Polka & Sundara, 2012). These data show that between 6 and 12 months of age, language-specific segmentation procedures appear, and differ between more stress-timed languages like English, and more syllable-based languages like French and Italian. Thus, the identification of syllables at the phonetic level may be more favored in a French or Italian input than in English, while English would favor the identification of syllables at the prosodic level.

We have chosen to review the substantial changes that segmentation abilities are undergoing between 6 and 12 months of age for several reasons. First, they have been very well documented. Second, segmenting words implies identifying consistent strings of consonants and vowels from the stream of speech as potential word forms, thus it may have important implications for consonant and vowel processing at the lexical level. Third, they seem to differentiate French and Italian from English. And finally, they occur in the same time-window as the switch observed from a vowel to a consonant bias in lexical processing in French and Italian infants, i.e. between 6 and 8 months. Therefore they could contribute to the differential patterns of emergence of the consonant bias observed in these language groups and should be explored in future crosslinguistic investigations. Other relevant factors should be addressed, such as the proportion of consonants and vowels distinguishing words in the languages, and how this property may change as the infant' s lexicon is growing (Keidel et al., 2007).

Overall the current body of literature available on the origins of the C/V functional asymmetry is incomplete, especially concerning the vowel bias in morphosyntactic processing where data is scarce, apart from Pons and Toro (2010) at 10 months of age, and our Experiments 2-3 at birth. However, the data concerning the use of vowels and consonants in lexical processing has already uncovered language-specific patterns in the emergence of the C/V asymmetry. Indeed, the amount of language exposure and the nature of the knowledge infants have to acquire before using consonants functionally varies depending on the input infants receive. To enrich the interpretations of our results and help disentangle the acoustic/phonetic and lexical hypotheses, we will refer to a

recent theoretical framework of early language acquisition in which any information at any stage of development can be the focus of the infant's attention, which is set according to a dynamic interaction between the *initial perceptual biases,* the infant's *developmental level,* and *the task demands* (PRIMIR, Werker & Curtin, 2005).

IV. 3. The C/V functional asymmetry and the Processing Rich Information from Multidimensional Interactive Representations (PRIMIR) model (Werker & Curtin, 2005)

In the present dissertation, we explored the origins and the developmental trajectory of consonant and vowel functional specialization, in order to determine whether the two speech sound categories could elicit different speech processing mechanisms and thus provide bootstrapping effects in learning words and structures of language. This research question led us to take into consideration multiple levels of speech processing capacities (acoustic, phonetic, lexical, structural, morphosyntactic) in very young infants. The very different methodological approaches adopted in the two age groups should be taken into account when interpreting our data. This work contributes to the understanding of how various processing levels are integrated at two very early ages and in two different experimental contexts. Our data appear to fit well with this recent model based on the link between early speech perception and lexical acquisition.

IV. 3. a. The PRIMIR model (Werker & Curtin, 2005)

According to this model, infants are born with a set of biases that act as *filters* to direct infants' attention to only some of the available information in their environment, and particularly in speech. These first filters are biases brought by infants to the language-learning task and are thus determined by evolutionary and epigenetic factors. They include a preference for speech, infant-directed speech, point vowels, proper syllable form, and the ability to process rhythmical patterns in speech. These initial filters are coupled with *general learning mechanisms* that also act as filters in computing statistical analyses, and thus guarantee the acquisition of linguistically possible combinations only. The *raw physical saliency* of the information (acoustic and phonetic, but also gestural, visual, etc.) also play a determining role in speech processing by interacting with the filters and the learning mechanisms, to form representations.

Crucially, PRIMIR considers that language processing and representations cannot be understood separately: the representations built are essential for efficient and productive processing. In this perspective, interactive representations are built by organizing all the information along *three multidimensional planes*, grouped on the basis of similarity, co-occurrence, and other statistical regularities. The first multidimensional interactive representational plane is **the General Perceptual Plane**. It processes and organizes all the phonetic and indexical properties of the signal, forming and storing exemplar-like distributions of the input and their frequency of occurrence. For instance, natural classes (i.e. segments that share gestures, e.g fricatives) are clustering together and thus some phonetic categories may emerge. All the information is context-sensitive and is grouped by co-occurrence, feature similarity or by any other statistical regularity. Some information may be relevant both at the phonetic and indexical level, e.g. pitch distinguishes voices, but also voicing and vowel color, thus pitch information can contribute to both phonetic and indexical learning.

The **second plane is the Word Form and Meaningful Word Plane**. Sequences of cohesive units are extracted and stored without meaning and their representation is based on the exemplar distributions, encoding phonetic and indexical information. Different exemplars of the same word are clustered based on indexical and phonetic similarity. Through the association of a word-form to a concept, meaningful words are created in this plane. Infants hear exemplars of word-forms overlapping in phonetic and indexical properties and see different instances of the object category, again in different contexts. The difficulty is less in forming representations of each than in linking a word form to a concept, because it requires that attention focuses *just on the information that is criterial* for distinguishing similar words. Once a sufficient number of meaningful words are accumulated, a generalization of commonalities takes place and high order regularities emerge, forming a system of contrastive phonemes stored and processed in **the Phoneme Plane**.

All these planes interact with each other dynamically and contribute to focusing the attention of the language learner on a specific portion of the rich information available in the speech signal. The amount of information attended to increases or decreases depending on the *biases,* the demands of the *task* and the developmental *level* of the listener. As infants advance in development, the influence of the initial biases and of raw physical saliency on which information is attended to in the signal diminishes and the information is available across more and more planes. The consequence is that task demands and developmental level play increasingly important roles. For instance, any phonetic difference is initially discriminable at the General Perceptual Plane, and as phonemes become more robust, raw physical saliency plays a lesser role in development. Contrary to other models focusing on speech perception (Perceptual Assimilation Model, Best, 1994; Native Language Magnet, Kuhl, 1993), PRIMIR aims at integrating infant speech perception to word learning into a unified account thus taking into account their relatedness. It was initially motivated by

seemingly contradictory findings of how infants were able to use phonetic detail in word learning tasks (Stager & Werker, 1997; Werker et al., 1998; Fennell & Werker, 2003, 2004; Werker, Fennell, Corcoran, & Stager, 2002; Fennell, 2004; Pater, Stager, & Werker, 2004). In this respect, PRIMIR is appropriate and relevant to explain our data on how two classes of speech sounds, consonants and vowels, which differ in their articulatory and acoustic properties, can also become two functional classes across development.

IV. 3. b. PRIMIR and our data

At birth, the pieces of information infants attend to in speech depend on epigenetically based perceptual biases, modulated by the acoustic saliency of the various elements composing the speech signal. The method used at birth in Experiments 1 to 3, relied on the detection of repetition-based regularities in word-like units, which was previously demonstrated for syllabic repetitions (Gervain et al., 2008, 2012; Exp. 1) and tested on segmental repetition in Experiments 2 and 3. Our findings reveal that vowels are not favored over consonants for such processing (Exp. 2-3). Importantly, in our stimuli the consonant and vowel sounds were chosen and synthesized in order to minimize acoustic saliency differences. These results thus show that structural extraction mechanisms can use both consonant and vowel information when they are acoustically as close as possible. Future studies will test the possibility that a difference would appear between ABBc and ABBv grammars when using natural stimuli in which vowels are more salient than consonants. If this were found it would suggest that in order to elicit the relevant biases proposed by Nespor et al. (2003), consonants and vowels have not only to be perceived categorically, but also to differ in their acoustic saliency.

Additionally, PRIMIR assumes that the same general statistical learning mechanisms are operating over the different levels of analysis simultaneously: prosodic, syllabic, and only later phonemic. Therefore, prosodic analysis, segmentation of the speech stream, extraction of syllables, formation of phonetic categories and extraction and storage of patterns of word-like units happen simultaneously. What develops in time is the way each plane will further influence category formation and information pick-up at every other plane. This is more compatible with our data that more sequential descriptions of the development of linguistic units in speech. For instance in the Word Recognition and Phonetic Structure Acquisition model (WRAPSA, Jusczyk, 1997) infant listeners first have access to prosodic, then syllabic, and only later phonetic information, and the extraction of word-like units can only take place once speech information has been differently weighted, thus after substantial experience with the signal. However, Experiment 2 shows that segmental repetitions are detectable from birth in word-like units, suggesting that newborns are immediately

sensitive to specific patterns of word-forms implementing a repetition on syllables or even segments. This result is more in line with an integrative than with a linear view, whereby segmental information may also be accessed and structural patterns embedded in segments detected very early on.

The proposal that phonemes are learned in a staggered fashion according to PRIMIR is also compatible with the way that the consonant bias seems to emerge according to our data. Indeed, in Experiments 4-7, infants displayed different levels of sensitivity to phonetic changes occurring in a word, depending on whether they were vocalic or consonantal. Importantly, all these native phonetic changes outside of a lexical context should be discriminated at this age. This discrepancy between sensitivity to phonetic changes in a lexical context and fine-grained discrimination capacities relates both to the important notions of task requirements and developmental level in PRIMIR. Indeed, at 5 months, more native vowel categories than consonant categories are phonemically encoded (Kuhl et al., 1992; Werker & Tees, 1984). Our task requires that they focus on the critical contrasting information between two very similar versions of the same word. Therefore, our results suggest that a reaction to a mispronunciation in their own name could be observed between two phonemically differing words, which should favor performance in the vowel condition. Although acoustic saliency may still be playing a role independently of the phonemic status of a given phonetic contrast, the regression analyzes identified a link between acoustic distances and infants' performance only in the vowel-change condition, and crucially not when both groups were pulled together. Therefore, acoustic distances between CPs and MPs seemed to act differently on infants' behavior depending on whether the phonetic change involved vowels or consonants, suggesting a categorical dichotomy.

Finally one crucial aspect of PRIMIR coherent with our view is the possibility to manipulate task requirements (or experimental context) and thus draw infants' attention to different elements of information in the input. With the same stimuli, the intention behind the designs of Experiment 2 and 3 was to induce repetition rule-learning in the first case, and item memorization in the other case. Although the paradigm of Experiment 3 may still be improved by a more variable material to avoid neural attenuation, Experiment 2 already suggests that repetition-based regularity detection uses both consonant or vowel information. Hence, the two contexts led to two different types of processing.

IV. 4. Perspectives

Our results answer some of the questions addressed at the beginning of this work. However, several related questions remain unanswered and our findings have raised new issues that need to be explored in future research.

IV. 4. a. Follow-ups experiments to Experiments 2 and 3 using NIRS

To confirm the effect of segmental repetitions and confirm their weaker learning effect in Experiments 2 and 3, a future NIRS experiment should replicate Experiment 2 using more variability in the material. Experiments 2 and 3 used the same list of 216 words presented in two different experimental designs differing in the order of presentation of the words. In Experiment 2, each word was repeated 6 times in 6 different blocks, while in Experiment 3 the 6 words were repeated in a row. To increase variability in Experiment 2, the 6 consonants and 6 vowels could be combined to generate 6 times more differing stimuli than in Experiment 2 and 3, or more consonants and vowels could compose our inventory. To increase variability in Experiment 3, while keeping uniform blocks of 6 identical items and thus induce lexical context, voice changes could be used. Indeed, if the 6 repetitions of each token were synthesized with 6 different MBROLA voices, this would focus attention on phonetic content while introducing voice variability in the material.

To further assess the role of acoustic saliency effects, the least sonorant consonants (i.e. plosives) could be used in further replications of Experiments 2 and 3 (instead of the more vowel-like continuous consonants used in the current study). In the lexical experimental design, these new consonants could be favored because of their distinctiveness (Lisker & Abramson, 1964).

IV. 4. b. Follow-ups to Experiments 4 to 9 at 5 months

The following perspectives are aiming at confirming the emergence of the consonant bias in French-learning infants between 6 and 8 months (Nishibayashi & Nazzi, 2014; Experiment 4-9). The specific status of the infant's own name may have contributed to the advantage found for vowels over consonants in our results. As previously mentioned, in spite of all its methodological advantages, the use of the infants' name could have benefited indexical cues over phonetic cues processing. Moreover vowels are known to carry more indexical cues than consonants, thus a vowel change is more likely to be noticed in an emotionally charged word. Thus, our results on the sensitivity to phonetic changes may be specific to words with a special social/emotional status. The

same task may be conducted with other word-forms known to be familiar at 6 months such as body parts (Tincoff & Jusczyk, 1999; 2012), and other frequent words of the child's environment.

One limitation of our paradigm is that given the difficulty to present more than two different kinds of stimuli to young infants in HPP, we had to use a between-subject design for the experiments. Thus, we tested 5-month-old infants' preference for their correctly-pronounced over mispronounced name, either on a consonant in Experiment 4, and on a vowel in Experiment 5. Another possibility would be to use a conflict situation as in Poltrock and Nazzi (in revision). Indeed, using a similar HPP preferential measures but in a conflict situation, French-learning 11-month-olds infants preferred listening to a list of vowel-mispronunciations over consonant-mispronunciations of familiar words. Thus, in a within-subject design, infants relied more on consonant information than vowel information when recognizing familiar words. Therefore, in a future experiment using the infants own name, the consonant bias should be explored at 8 months in a familiar word recognition task similar to our own.

Another limitation of the head-turn preference procedure is that it does not provide graded measures of discrimination performance that could be ranked across sound conditions. In our paradigm, practically each infant was tested on a different set of stimuli, and thus a different phonetic contrast. Therefore, some segments and mispronunciations are expected to be more deleterious than others, which could be observed using other measures. Future studies may use anticipatory looking times to assess individual performance according to various levels of stimulus degradation. Electrophysiological methods may also be used to explore the neural mechanisms (and their development) underlying the perception of various mispronounced familiar word-forms. An oddball paradigm combined with measurements of the neural responses to mispronunciations may be appropriate to further explore the discrimination of phonetic changes in first words.

IV. 4. c. Future crosslinguistic investigations of the C/V functional asymmetry

Comparing studies in English- and French-learning infants, which have been most extensively studied, suggests very different developmental trajectories for the consonant bias in lexical processing in the second and first year of life (Nishibayashi & Nazzi, 2014; Bouchon et al., in press, Poltrock & Nazzi, in revision; Floccia et al., 2014). This raises two complementary questions: First, what language-specific factors can explain that different functions are attributed to consonants and vowels in French from the age of 8 months and not English before 30 months? Second, is the C/V functional asymmetry altogether delayed in English-learning infants, and if so how does the vowel bias in grammatical processing develop?

Cross-linguistic studies in infancy would reveal the language-specific factors (distributional, prosodic…) that might modulate how consonants and vowels are used in lexical and syntactic tasks. So far, only one study has investigated the case of Danish-learning infants and revealed that at 20 months they showed a vocalic bias in an interactive word-learning task. This result is in line with the lexical hypothesis of emergence of the consonant bias (Keidel et al., 2007) given that Danish has more vowels than consonants. However, vowels are also more prominent that consonants at the acoustic/phonetic levels (Pharao, 2011; Basbøll, 2005), thus the results are also compatible with the acoustic/phonetic hypothesis (Floccia et al., 2014). Cross-linguistic investigations should thus be conducted on French- and Danish-learning infants to assess the age of emergence of a consonant bias in Danish.

In addition, only one study investigated the relative role of vowels and consonants in repetition-based rule learning (Pons & Toro, 2010) based on Marcus et al.'s (1999) seminal study. To explore whether the division of labor is altogether weaker in English compared to French, English- and French-learning infants could be tested at 10 months on their ability to extract and generalize an ABB repetition implemented on consonants or on vowels.

IV. 5. Final Conclusions

The aim of the current dissertation was to investigate the hypothesis proposed by Nespor et al. (2003) that the consonant bias in lexical processing and the vowel bias in syntactic processing were perceptual constraints used respectively to analyze language into its constituents, i.e. words, and organizing patterns, i.e. syntactic rules. Regarding this hypothesis, our data are providing three major pieces of information concerning French-learning infants:

· Firstly, at birth, consonant and vowel information seem to be equally processed by precursory mechanisms of syntax acquisition. Indeed, repetition-based regularities carried by consonants or by vowels in words are both detected in the newborn brain. These findings first extend precursory syntax processing found previously at the syllable level (Gervain et al., 2008, 2012) to segmental information. Secondly, they are also in contradiction with the assumption made by Nespor et al. (2003), that infants are born with the ability to use vowels preferentially in syntax processing.

· Secondly, in a lexical processing task, vowels seem to be processed preferentially over consonants at 5 months, i.e. at the onset of lexical acquisition. Indeed, in a familiar word recognition task, infants were sensitive to a vowel change but not to a consonant change in one of their first

familiar word, their own name. This result contradicts the hypothesis that French-learning infants are using consonants preferentially in lexical tasks, and rather suggests that the consonant bias has to emerge through language exposure.

· Thirdly, the results of acoustic analyses of the stimuli linked to individual performances are in favor of an acoustic/phonetic emergence of the consonant bias in lexical processing. Indeed, 5-month-old infants seem to be processing consonants and vowels categorically in words according to their acoustical properties.

To sum up, these findings specify the relative contributions of consonants and vowels in lexical and syntactic precursory mechanisms at birth (Chapter 2) and in early word recognition at 5 months (Chapter 3). They fail to support the previous hypotheses formulated by Nespor et al. (2003) on the developmental origins of the C/V functional asymmetry. Together with other recent crosslinguistic reports (Hochmann et al., 2011; Hochmann, 2010; Poltrock & Nazzi, in revision; Nishibayashi & Nazzi, 2014; Højen & Nazzi, in revision; Bouchon, Delle Luche, Floccia & Nazzi, 2012), our findings contribute to defining the different developmental trajectories of the C/V division of labor in language acquisition. Some of these reports are preliminary and crosslinguistic investigations should be pursued. Such investigations will explain more precisely how language-specific input influences the early development of the C/V functional asymmetry.

BIBLIOGRAPHY

Abercrombie, D. (1967). *Elements of General Phonetics*. Edinburgh: Edinburgh University Press.

Acha, J., & Perea, M. (2010). On the role of consonants and vowels in visual-word processing Evidence with a letter search paradigm. *Language and Cognitive Processes, 25*, 423-438.

Aguirre, G. K, & D'Esposito, M. (1999). Experimental design for brain fMRI. In: P. A. Bandettini (Eds) *Functional MRI*. (P. 369–380). Berlin: Springer-Verlag;

Albareda-Castellot, B., Pons, F., & Sebastián-Gallés, N. (2011). The acquisition of phonetic categories in bilingual infants: new data from an anticipatory eye movement paradigm. *Developmental Science,* 14, 395-401.

Alderliesten, T., De Vis, J. B., Lemmers, P. M. A., van Bel, F., Benders, M. J. N. L., Hendrikse, J., & Petersen, E. T. (2014). Simultaneous quantitative assessment of cerebral physiology using respiratory-calibrated MRI and near-infrared spectroscopy in healthy adults. *Neuroimage,* 85, 255–263.

Anderson, J. L., Morgan, J. L., & White, K. S. (2003). A statistical basis for speech sound discrimination. *Language and speech,* 46(2), 155‐182.

Aslin, R. N. (2007). What's in a look? *Developmental Science*. 10, 48–53.

Aslin, R. N. (2012). Questioning the questions that have been asked about the infant brain using near-infrared spectroscopy. *Cognitive neuropsychology,* 29(1-2), 7–33.

Baldeweg, T. (2006). Repetition effects to sounds: Evidence for predictive coding in the auditory system. *Trends in Cognitive Science*, 10: 93–94.

Barac, R., & Bialystok, E. (2012). Bilingual effects on cognitive and linguistic development: Role of language, cultural background, and education. *Child Development,* 83, 413-422.

Barras, C., Geoffroy, E., Wu, Z., & Libberman, M. (2001). Transcriber: development and use of a tool for assisting speech corpora production. *Speech Communication,* 33, 5–22.

Ballem, K. D., & Plunkett, K. (2005). Phonological specificity in children at 1;2. *Journal of child language,* 32(1), 159‐173.

Bartocci, M., Winberg, J., Ruggiero, C., Bergqvist, L. L., Serra, G., & Lagercrantz, H. (2000). Activation of olfactory cortex in newborn infants after odor stimulation: a functional near-infrared spectroscopy study. *Pediatric research*, 48(1), 18‐23.

Bates, E., Dale, P., & Thal, D. (1995). Individual differences and their implications for theories of language development. In P. Fletcher & B. MacWhinney (Eds.) *The handbook of child language* (pp. 96-151). Oxford, England: Blackwell.

Belin, P., & Zatorre, R. J. (2003). Adaptation to speaker's voice in right anterior temporal lobe. *Neuroreport,* 14, 2105–2109.

Benavides-Varela, S., Gómez, D. M., Macagno, F., Bion, R. A., Peretz, I., & Mehler, J. (2011). Memory in the neonate brain. *PloS one*, 6(11), e27497.

Benavides-Varela, S., Hochmann, J.-R., Macagno, F., Nespor, M., & Mehler, J. (2012). Newborn's brain activity signals the origin of word memories. *Proceedings of the National Academy of Sciences,* 109(44), 17908-17913.

Benjamini, Y., & Yekutieli, D. (2001). The control of the false discovery rate in multiple testing under dependency. *Annals of Statistics,* 29, 1165–1188.

Berent, I., Marcus, G. F., Shimron, J., & Gafos, A. I. (2002). The scope of linguistic generalizations: evidence from Hebrew word formation. *Cognition,* 83, 113–139.

Berent, I., Vaknin, V., & Marcus, F. G. (2007). Roots, stems, and the universality of lexical representations: evidence from Hebrew. *Cognition*. 104, 254–286.

Bergelson, E., & Swingley, D. (2012). At 6–9 months, human infants know the meanings of many common nouns. *Proceedings of the National Academy of Sciences,* 109(9), 3253-3258.

Bertoncini, J., Bijeljac-Babic, R., Jusczyk, P. W., & Kennedy, L. J. (1988). An investigation of young infants' perceptual representations of speech sounds. *Journal of Experimental Psychology: General, 117*, 21-33.

Best, C. T., McRoberts, G. W., & Sithole, N. M. (1988). Examination of perceptual reorganization for nonnative speech contrasts: Zulu click discrimination by English-speaking adults and infants. *Journal of Experimental Psychology: Human Perception and Performance,* 14(3), 345-360.

Bleses, D., Basbøll, H., & Vach, W. (2011). Is Danish difficult to acquire? Evidence from Nordic past-tense studies. *Language and Cognitive Processes*, 26(8), 1193-1231.

Boas, D. A., Ellwell, C. E., Ferrari, M., & Taga, T. (2014) Twenty years of functional near-infrared spectroscopy: introduction for the special issue. *NeuroImage*, 85, 1–5

Bonatti, L. L., Peña, M., Nespor, M., & Mehler, J. (2005). Linguistic constraints on statistical computations: the role of consonants and vowels in continuous speech processing. *Psychological Science*, 16(6), 451–459.

Bonatti, L. L., Peña, M., Nespor, M., & Mehler, J. (2007). On consonants, vowels, chickens, and eggs. *Psychological Science*, 18(10), 924–925.

Bortfeld, H., Morgan, J., Golinkoff, R., & Rathbun, K. (2005). Mommy and me: Familiar names help launch babies into speech stream segmentation. *Psychological Science, 16*, 298–304.

Bortfeld, H., Fava, E., & Boas, D. A. (2009). Identifying cortical lateralization of speech processing in infants using near-infrared spectroscopy. *Developmental neuropsychology,* 34(1), 52-65.

Bortfeld, H., Wruck, E., & Boas, D. A. (2007). Assessing infants' cortical response to speech using near-infrared spectroscopy. *Neuroimage,* 34(1), 407–415.

Bosch, L., & Sebastián-Gallés, N. (2003a). Simultaneous bilingualism and the perception of a language-specific vowel contrast in the first year of life. *Language and Speech*, 46, 217–243.

Bosch, L., & Sebastián-Gallés, N. (2003b). Language experience and the perception of a voicing contrast in fricatives: Infant and adult data. In D. Recasens, M. J. Solé & J. Romero (eds.), *Proceedings of the 15th international conference of phonetic sciences*, (pp. 1987–1990). Barcelona: UAB/Casual Prods.

Bosch L., Figueras M., Teixidó, M., & Ramon-Casas M. (2013). Rapid gains in segmenting fluent speech when words match the rhythmic unit: evidence from infants acquiring syllable-timed languages. *Frontiers in Psychology,* 4, 106.

Bouchon, C., delle Luche, C., Floccia, C. & Nazzi, T. (2012, November). Consonantal specificity of name recognition in French- and English-learning 5-month-olds. Poster presented at the 37th Annual Boston University Conference On Language Development, Boston, Massachusetts, USA.

Bouchon, C., Floccia, C., Fux, T., Adda-Decker, M., & Nazzi, T. (2014). Call me Alix, not Elix: Vowels are more important than consonants in own name recognition at 5 months. *Developmental Science*, in press.

Bouchon, C., Nazzi, T., & Gervain, J. (in revision). Hemispheric asymmetries in repetition enhancement and suppression effects in the newborn brain.

Buckner, R. L., Bandettini, P. A., O'Craven, K. M., Savoy, R. L., Petersen S. E., Raichle M.E., & Rosen B. R. (1996). Detection of cortical activation during averaged single trials of a cognitive task using functional magnetic resonance imaging. *Proceedings of the National Academy of Sciences*, 93, 14878–14883.

Burns, T. C., Yoshida, K. A., Hill, K., & Werker, J. F. (2007). The development of phonetic representation in bilingual and monolingual infants. *Applied Psycholinguistics*, 28, 455–474.

Caramazza, A., Chialant, D., Capasso, R., & Miceli, G. (2000). Separable processing of consonants and vowels*. Nature,* 403(6768), 428–30.

Carreiras, M., Dunabeitia, J. A., & Molinaro, N. (2009). Consonants and vowels contribute differently to visual word recognition: ERPs of relative position priming. *Cerebral Cortex,* 19: 2659–2670.

Carreiras, M., Gillon-Dowens, M., Vergara, M., & Perea, M. (2009). Are vowels and consonants processed differently? ERP evidence with a delayed letter paradigm. *Journal of Cognitive Neuroscience*, 21: 275–288.

Carreiras, M., & Price, C. J. (2008). Brain activation for consonants and vowels. *Cerebral Cortex*, *18*(7), 1727–1735.

Carreiras, M., Vergara, M., & Perea, M. (2009). ERP correlates of transposed-letter similarity effects: Are consonants processed differently from vowels? *Neuroscience Letters*, 419, 219–224.

Chance, B., Zhuang, Z., Unah, C., Alter, C., & Lipton, L., (1993). Cognition-activated low- frequency modulation of light absorption in human brain. *Proceedings of the National Academy of Sciences*, 90, 3770–3774.

Chee, M. W. L., Venkatraman, V., Westphal, C., & Siong, S. C. (2003). Comparison of Block and Event-Related fMRI Designs in Evaluating the Word-Frequency Effect. *Human Brain Mapping,* 18, 186 –193.

Cheour, M., Haapanen, M. L., Ceponiene, R., Hukki, J., Ranta, R., & Näätänen, R. (1998). Mismatch negativity (MMN) as an index of auditory sensory memory deficit in cleft-palate and CATCH syndrome children. Neuroreport: An International Journal for the Rapid Communication of Research in Neuroscience, 9(12), 2709-2712.

Cheour-Luhtanen, M., Alho, K., Kujala, T., Sainio, K., Reinikainen, K., Renlund, M., Aaltonen, O., Eerola, O., & Näätänen, R. (1995). Mismatch negativity indicates vowel discrimination in newborns. *Hearing Research,* 82(1), 53–58.

Chomsky, N. (1957). *Syntactic Structures.* The Hague: Mouton.

Chomsky, N. (1959). A review of B. F. Skinner's Verbal Behavior. *Language,* 35(1), 26-58.

Christophe, A., Dupoux, E., Bertoncini, J., & Mehler, J. (1994). Do infants perceive word boundaries? An empirical study of the bootstrapping of lexical acquisition. *Journal of the Acoustical Society of America,* 95(3), 1570-1580.

Christophe, A., Mehler, J., & Sebastiàn-Gallés, N. (2001). Perception of prosodic boundary correlates by newborn infants*. Infancy,* 2, 358-394.

Cohen, L., Jobert, A., Le Bihan, D., & Dehaene, S. (2004). Distinct unimodal and multimodal regions for word processing in the left temporal cortex. *Neuroimage* 23, 1256 –1270.

Creel, S. C., Aslin, R. N., & Tanenhaus, M. K. (2006). Acquiring an artificial lexicon: Segment type and order information in early lexical entries. *Journal of Memory and Language,* 54(1), 1–19.

Cristià, A., Dupoux, E., Hakuno, Y., Lloyd-Fox, S., Schuetze, M., Kivits, J., Bergvelt, T., van Gelder, M., Filippin, L., Charron, S., & Minagawa-Kawai, Y., (2013). An online database of infant functional near infrared spectroscopy studies: a community-augmented systematic review. *PLoS ONE*, 8, e58906.

Cubelli, R. (1991). A selective deficit for writing vowels in acquired dysgraphia. *Nature,* 353(6341), 258-260.

Cutler, A., & Carter, D. M. (1987). The predominance of strong initial syllables in the English vocabulary. *Computer Speech and Language,* 2, 133-142.

Cutler, A., & Clifton, C. E. (1999). Comprehending spoken language: A blueprint of the listener. In C. Brown; P. Hagoort (eds.): Neurocognition of Language (p. 123-155). Oxford: Oxford University Press.

Cutler, A., Dahan, D., & Van Donselaar, W. (1997). Prosody in the comprehension of spoken language: A literature review. *Language and speech*, 40(2), 141–201.

Cutler, A., & Mehler, J. (1993). The periodicity bias. *Journal of Phonetics*, 21(1), 103-108.

Cutler, A., Mehler, J., Norris, D., & Segui, J. (1986). The Syllable's Differing Role in the Segmentation of French and English. *Journal of Memory and Language*, 25(4), 385-400.

Cutler, A., Sebastiàn-Gallés, N., Soler-Vilageliu, O., & van Ooijen, B. (2000). Constraints of vowels and consonants on lexical selection: Cross-linguistic comparisons. *Memory and Cognition*, 28(5), 746-755.

Davis, S. B. & Mermelstein, P. (1980). Comparison of parametric representations for monosyllabic word recognition in continuously spoken sentences. *IEEE Transactions on Acoustics, Speech and Signal Processing,* 28, 4, 357 – 366.

Dehaene-Lambertz, G., & Baillet, S. (1998). A phonological representation in the infant brain. *Neuroreport*, 9(8), 1885-1888.

Dehaene-Lambertz, G., Dehaene, S., Anton, J. L., Campagne, A, Ciuciu, P. Dehaene, G. P., Denghien, I., Jobert, A., LeBihan, D., Sigman, M. (2006). Functional segregation of cortical language areas by sentence repetition. *Human Brain Mapping*, 27, 360–371.

Dehaene-Lambertz, G., Hertz-Pannier, L., Dubois, J., Mériaux, S., & Roche, A., Sigman. M., & Dehaene, S. (2006). Functional organization of perisylvian activation during presentation of sentences in preverbal infants. *Proceedings of the National Academy of Sciences*, 103(38), 14240-14245.

Delattre, P. C. (1965). *Comparing the phonetic features of English, French, German and Spanish.* Heidelberg: Julius Groos Verlag.

Dell, F. (1985). *Les règles et les sons.* Second revised edition. Hermann, Paris.

Delle Luche, C., Poltrock, S., Goslin, J., New, B., Floccia, C., & Nazzi, T. (2014). Differential processing of consonants and vowels in the auditory modality: A cross-linguistic study. *Journal of Memory and Language,* 72, 1–15.

Delpy, D. T., Cope, M., van der Zee, P., Arridge, S., Wray, S., & Wyatt, J. (1988). Estimation of optical pathlength through tissue from direct time of flight measurement. *Physics in Medicine and Biology,* 33(12), 1433-1442.

Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review Neuroscience,* 18, 193–222.

Dewson, J. H. (1964). Speech sound discrimination by cats. *Science,* 141, 555-556.

Dieler, A. C., Tupak, S. V., & Fallgatter, A. J., (2012). Functional near-infrared spectroscopy for the assessment of speech related tasks. *Brain & Language*. 121, 90–109.

Dolan. R. J., Fink, G. R., Rolls, E., Booth, M., Holmes, A., Frackowiak, R. S., & Friston, K. J. (1997). How the brain learns to see objects and faces in an impoverished context. *Nature,* 389, 596–599.

Dooling, R. J., Best, C. T., & Brown, S. D. (1995). Discrimination of synthetic full-formant and sinewave /ra-la/ continua by budgerigars (Melopsittacus undulatus) and zebra finches (Taeniopygia guttata). *Journal of the Acoustical Society of America*, 97(3), 1839-1846.

Duñabeitia, J. A., & Carreiras, M. (2011). The relative position priming effect depends on whether letters are vowels or consonants. *Journal of Experimental Psychology: Learning, Memory and Cognition,* 37, 1143–1163.

Dutoit, T., Pagel, V., Pierret, N., Bataille, F., & van der Vreken, O. (1996). The MBROLA project: towards a set of high-quality speech synthesizers free of use for non-commercial purposes. In *Proceedings of the fourth international conference on spoken language processing,* 96 (3), 1393–1396.

Ehlis, A.-C., Schneider, S., Dresler, T., & Fallgatter, A. J. (2014). Application of functional near-infrared spectroscopy in psychiatry. *Neuroimage,* 85, 478–488.

Endress. A. D., & Bonatti, L. L. (2007). Rapid learning of syllable classes from a perceptually continuous speech stream. *Cognition,* 105 (2), 247-299.

Endress. A. D., Nespor, M., & Mehler, J. (2009). Perceptual and memory constraints on language acquisition. *Trends in Cognitive Sciences,* 13(8), 348–353.

Eimas, P. D., Siqueland, E. R., Jusczyk, P. W., & Vigorito, J. (1971). Speech perception in infants. *Science*, 171(968), 303-306.

Erickson, M. L. (2000). Simultaneous effects on vowel duration in American English: A covariance structure modeling approach. *The Journal of the Acoustical Society of America,* 108(6), 2980–2995.

Fabiano-Smith, L., & Barlow, J. A. (2010). Interaction in bilingual phonological acquisition: Evidence from phonetic inventories. *International Journal of Bilingual Education and Bilingualism*. 13, 81-97.

Fava, E., Hull, R., & Bortfeld, H. (2014). Dissociating Cortical Activity during Processing of Native and Non-Native Audiovisual Speech from Early to Late Infancy. *Brain Sciences,* 4(3), 471–487.

Fennell, C. T., Byers-Heinlein, K., & Werker, J. F. (2007). Using speech sounds to guide word learning: the case of bilingual infants. *Child Development*. 78, 1510–1525.

Fennell, C. T. & Byers-Heinlein, K. (in press, 2014). You sound like Mommy: Bilingual and monolingual infants learn words best from speakers typical of their language environments. *International Journal of Behavioral Development*.

Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D. J., & Pethick, S. J. (1994). Variability in Early Communicative Development. [Monograph]. *Society for Research in Child Development,* 5(242), 1-173.

Ferguson, C. A. (1977). Baby talk as a simplified register. In *Talking to children: Language input and acquisition* (p. 209-235). Cambridge: Cambridge University Press.

Fernald, A. (1985). Four-month-old infants prefer to listen to motherese. Infant *Behavior and Development*, 8, 181-195.

Fernald, A., & Kuhl, P. K. (1987). Acoustic determinants of infant preference for motherese speech. *Infant Behavior and Development,* 10, 279-293.

Ferrari, M., Mottola, L., & Quaresima, V. (2004). Principles, techniques, and limitations of near infrared spectroscopy. *Canadian journal of applied physiology-Revue canadienne de physiologie appliquée,* 29(4), 463-487.

Ferrari, M., & Quaresima, V., (2012). A brief review on the history of human functional near-infrared spectroscopy (fNIRS) development and fields of application. *Neuroimage,* 63, 921–935.

Ferreres, A. R., López, C., & China, N. N. (2003). Phonological alexia with vowel–consonant dissociation in non-word reading. *Brain & Language,* 84, 399–413.

Floccia, C., Butler, J., Goslin, J., & Ellis, L. (2009). Regional and foreign accent processing in English: Can listeners adapt? *Journal of Psycholinguistic Research, 38*, 379-412.

Floccia, C., Nazzi, T., Delle Luche, C., Poltrock, S., & Goslin, J. (2014). English-learning one-to two-year-olds do not show a consonant bias in word learning. *Journal of child language,* 1–30.

Friederici, A. D. (2005). Neurophysiological markers of early language acquisition: from syllables to sentences*. Trends in Cognitive Sciences,* 9(10), 481-488.

Friederici, A. D. (2002). Towards a neural basis of auditory sentence processing. *Trends in Cognitive Sciences*, 6, 78–84.

Friederici, A. D., Steinhauer, K., & Pfeifer, E. (2002). Brain signatures of artificial language processing: Evidence challenging the critical period hypothesis. Proc Natl Acad Sci U S A, 99(1), 529-534.

Friston, K. J., Zarahn, E., Josephs, O., Henson, R. N., & Dale, A. M. (1999) Stochastic designs in event-related fMRI. *Neuroimage,* 10, 607– 619.

Fry, D. B., Abramson, A. S., Eimas, P. D., & Liberman, A. M. (1962). The identification and discrimination of synthetic vowels. *Language and Speech*, 5(4), 171–189.

Gervain, J., Macagno, F., Cogoi, S., Peña, M., & Mehler, J. (2008). The neonate brain detects speech structure. *Proceedings of the National Academy of Sciences,* 105(37), 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. *Developmental Cognitive Neuroscience* 1: 22–46.

Gervain, J., Berent, I., & Werker, J. F. (2012). Binding at birth: The newborn brain detects identity relations and sequential position in speech. *Journal of cognitive neuroscience,* 24(3), 564–574.

Golinkoff, R. M., Hirsh-Pasek, K., Cauley, K. M., & Gordon, L. (1987). The eyes have it: lexical and syntactic comprehension in a new paradigm. *Journal of Child Language,* 14, 23-45.

Gomez, R. L., & Gerken, L. (1999). Artificial grammar learning by 1-year-olds leads to specific and abstract knowledge. *Cognition*, 70(2), 109–135.

Gomez, R. L. (2002). Variability and detection of invariant structure. *Psychological Science*, 13: 431–436.

Goyet, L., Nishibayashi, L.-L., & Nazzi, T. (2013). Early syllabic segmentation of fluent speech by infants acquiring French. *Plos One,* 8(11), e79646.

Goyet, L., de Schonen, S., & Nazzi, T. (2010). Words and syllables in fluent speech segmentation by French-learning infants: an ERP study. *Brain Research*, 1332, 75.

Grabe, E., & Low, E. L. (2002). Durational variability in speech and the rhythm class hypothesis. In *Papers in laboratory phonology*. (Vol. 7, p. 515-546). Berlin: Mouton de Gruyter.

Granier-Deferre, C., Ribeiro, A., Jacquet, A.Y., & Bassereau, S. (2011) Near-term fetuses process temporal features of speech. *Developmental Science, 14*, 336–352.

Grill-Spector, K., Kushnir, T., Hendler, T., & Malach, R. (2000). The dynamics of object-selective activation correlate with recognition performance in humans. *Nature Neuroscience,* 3, 837–843.

Grill-Spector, K., & Malach, R. (2001). fMR - adaptation: A tool for studying the functional properties of human cortical neurons. *Acta Psychologica*, 107(1–3), 293–321.

Grill-Spector, K., Henson, R., & Martin, A. (2006). Repetition and the brain: neural models of stimulus-specific effects. *Trends in Cognitive Sciences*, 10, 14–23.

Grønnum, N. (2005). DanPASS–Danish Phonetically Annotated Spontaneous Speech. In A. Eriksson & J. Lindh (Eds.), *Proceedings of FONETIK* (pp. 1578-1583). Göteborg, Sweden: Göteborg University.

Haenschel, C., Baldeweg, T., Croft, R. J., Whittington, M., & Gruzelier, J. (2000). Gamma and beta frequency oscillations in response to novel auditory stimuli: A comparison of human EEG data with in vitro models. *Proceedings of the National Academy of Sciences*, 97, 7645–7650.

Hall, D. G. (2009). Proper names in early word learning: Rethinking a theoretical account of lexical development. *Mind & Language, 24*, 404–432.

Hallé, P. A., & de Boysson-Bardies, B. D. (1994). Emergence of an early receptive lexicon: infants' recognition of words*. Infant Behavior and Development*, 17(2), 119-129.

Hallé, P. A., & de Boysson-Bardies, B. D. (1996). The format of representation of recognized words in infants' early receptive lexicon. *Infant Behavior and Development*, 19, 463‐481.

Havy, M., & Nazzi, T. (2009). Better processing of consonantal over vocalic information in word learning at 16 months of age. *Infancy,* 14(4), 439–456.

Havy, M., Bertoncini, J., & Nazzi, T. (2011). Word learning and phonetic processing in preschool age children. *Journal of Experimental Child Psychology, 108*, 25–43.

Havy, M., Bouchon, C., & Nazzi, T. (in revision). Phonetic processing when learning words: the case of bilingual infants.

Havy, M., Serres, J., & Nazzi, T. (2014). A Consonant/Vowel Asymmetry in Word-form Processing: Evidence in Childhood and in Adulthood*. Language and Speech*, 57(2), 254–281.

Hebden, J. C., Gibson, A., Austin, T., Yusof, R. M., Everdell, N., Delpy, D. T., Arridge, S. R., Meek, J. H., & Wyatt, J. S. (2004). Imaging changes in blood volume and oxygenation in the newborn infant brain using three-dimensional optical tomography. *Physics in Medicine and Biology,* 49(7), 1117‐1130.

Hebden, J. C., Gibson, A., Yusof, R. M., Everdell, N., Hillman, E. M., Delpy, D. T., Arridge, S. R., Meek, J. H., Austin, T., & Wyatt, J. S. (2002). Three-dimensional optical tomography of the premature infant brain. *Physics in Medicine and Biology,* 47(23), 4155‐4166.

Henson, R., Shallice, T., & Dolan, R. (2000). Neuroimaging evidence for dissociable forms of repetition priming. *Science,* 287, 1269–1272.

Henson, R. N. (2003). Neuroimaging studies of priming*. Progress in Neurobiology*, 70, 53–81.

Hochmann, J.-R. (2010). Unpublished Categories, Words and Rules in Language Acquisition [Doctoral dissertation]. Trieste – Italy: SISSA.

Hochmann, J.-R., Benavides-Varela, S., Nespor, M., & Mehler, J. (2011). Consonants and vowels: different roles in early language acquisition. *Developmental Science*, 14(6), 1445–1458.

Højen, A., & Nazzi, T. (in revision). Vowel-bias in Danish word-learning: Processing biases are language-specific.

Homae, F., (2014). A Brain of Two Halves: Insights into interhemispheric organization pro- vided by Near-infrared Spectroscopy. *Neuroimage,* 85, 354–362.

Hoshi, Y., & Tamura, M., (1993). Detection of dynamic changes in cerebral oxygenation coupled to neuronal function during mental work in man. *Neuroscience Letters*, 150, 5–8.

Houston, D. M., Jusczyk, P. W., Kuijpers, C. T. L., Coolen, R., & Cutler, A. (2000). Cross-language word segmentation by 9-month-olds. *Psychonomic Bulletin and Review*, 7(3), 504‑509.

Huppert, T. J., Diamond, S. G., Franceschini, M. A., & Boas, D.A., (2009). HomER: a review of time-series analysis methods for near-infrared spectroscopy of the brain. *Applied Optics*, 48 (10), D280–D298.

James, T. W., & Gauthier, I. (2006). Repetition-induced changes in BOLD response reflect accumulation of neural activity. *Hum Brain Mapping,* 27, 37–46.

Jöbsis, F. F. (1977). Noninvasive, infrared monitoring of cerebral and myocardial oxygen sufficiency and circulatory parameters. *Science*, 198, 1264–1267.

Jöhr J. L, (2014). Les consonnes du biberon sous influences: Rôle du contexte dans les compétences phonologiques précoces. [Doctoral dissertation]. Genève, Switzerland: Université de Genève.

Jusczyk, P. W., Cutler, A., & Redanz, N. J. (1993a). Infants' preference for the predominant stress patterns of English words. *Child development*, 64(3), 675-687.

Jusczyk, P. W., Friederici, A. D., Wessels, J. M., Svenkerud, V. Y., & Jusczyk, A. M. (1993b). Infants′ sensitivity to the sound patterns of native language words. *Journal of Memory and Language*, 32(3), 402-420.

Jusczyk, P. W., & Aslin, R. N. (1995). Infants' detection of the sound patterns of words in fluent speech. *Cognitive Psychology,* 29(1), 1–23.

Jusczyk, P. W. (1997). The discovery of spoken language. Cambridge, MA: MIT Press.

Jusczyk, P. W., Houston, D. M., & Newsome, M. R. (1999). The beginnings of word segmentation in English-learning infants. *Cognitive psychology*, 39(3), 159‑207.

Kato, T., Kamei, A., Takashima, S., & Ozaki, T. (1993). Human visual cortical function during photic stimulation monitoring by means of near-infrared spectroscopy. *Journal of Cerebral Blood Flow & Metabolism*. 13, 516–520.

Keidel, J. L., Jenison, R. L., Kluender, K. R., & Seidenberg, M. S. (2007). Does grammar constrain

statistical learning? Commentary on Bonatti, Peña, Nespor, and Mehler (2005). *Psychological Science, 18,* 922–923.

Kern, S. (2007). Lexicon development in French-speaking infants. *First Language*, 27(3), 227–250.

Kewley-Port, D., Burkle, T. Z., & Lee, J. H. (2007). Contribution of consonant versus vowel information to sentence intelligibility for young normal-hearing and elderly hearing-impaired listeners. *The Journal of the Acoustical Society of America*, 122(4), 2365–2375.

Kleinschmidt, A., Obrig, H., Requardt, M., Merboldt, K. D., Dirnagl, U., Villringer, A., & Frahm, J. (1996). Simultaneous recording of cerebral blood oxygenation changes during human brain activation by magnetic resonance imaging and near-infrared spectroscopy. *Journal of Cerebral Blood Flow & Metabolism*, 16(5), 817–826.

Kluender, K. R., Lotto, A. J., Holt, L. L., & Bloedel, S. L. (1998). Role of experience for language-specific functional mappings of vowel sounds. *Journal of the Acoustic Society of America,* 104(6), 3568-3582.

Kooijman, V.M., Hagoort, P., & Cutler, A. (2005). Electrophysiological evidence for prelinguistic infants' word recognition in continuous speech. *Cognitive Brain Research,* 24, 109–116.

Kooijman, V., Hagoort, P., & Cutler, A. (2009). Prosodic structure in early word segmentation: ERP evidence from Dutch ten-month-olds. *Infancy,* 14, 591–612.

Kourtzi, Z., & Huberle, E. (2005). Spatiotemporal characteristics of form analysis in the human visual cortex revealed by rapid event-related fMRI adaptation. *Neuroimage* 28: 440–452.

Kuhl, P. K. (1991). Human adults and human infants show a « perceptual magnet effect » for the prototypes of speech categories, monkeys do not. *Perception & psychophysics,* 50(2), 93-107.

Kuhl, P. K., & Miller, J. D. (1978). Speech perception by the chinchilla: Identification functions for synthetic VOT stimuli. *The Journal of the Acoustical Society of America,* 63(3), 905–917.

Kuhl, P. K. (1981). Discrimination of speech by nonhuman animals: Basic auditory sensitivities conducive to the perception of speech-sound categories. *The Journal of the Acoustical Society of America*, 70(2), 340–349.

Kuhl, P. K., & Padden, D. M. (1982). Enhanced discriminability at the phonetic boundaries for the voicing feature in macaques. *Perception & psychophysics*, 32(6), 542-550.

Kuhl, P. K., & Padden, D. M. (1983). Enhanced discriminability at the phonetic boundaries for the place feature in macaques. *Journal of the Acoustical Society of America,* 73(3), 1003-1010.

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.

Kuhl, P. K. (2004). Early language acquisition: cracking the speech code. *Nature Reviews. Neuroscience*, 5(11), 831–43

Kuhl, P. K., Conboy, B. T., Coffey-Corina, S., Padden, D., Rivera-Gaxiola, M., & Nelson, T. (2008). Phonetic learning as a pathway to language: new data and native language magnet theory expanded (NLM-e). *Philosophical Transactions of the Royal Society B: Biological Sciences. 363*, 979–1000.

Kusaka, T., Kawada, K., Okubo, K., Nagano, K., Namba, M., Okada, H., Imai, T., Isobe, K., & Itoh, S. (2004). Noninvasive Optical Imaging in the Visual Cortex in Young Infants. *Human brain mapping,* 22(2), 122‐132.

Ladefoged, P. (1993). A Course in Phonetics. Harcourt Brace Jovanovich College Publishers.

Ladefoged, P. (2001). Vowels and consonants: An introduction to the sounds of languages. Oxford, United Kingdom: Blackwell.

Lee, H. W., Rayner, K., & Pollatsek, A. (2001). The relative contribution of consonants and vowels to word identification during reading. *Journal of Memory and Language,* 44, 189-205.

Liberman, A. M., Delattre, P. C., Cooper, F. S., & Gerstman, L. J. (1954). The role of consonant-vowel transitions in the perception of the stop and nasal consonants. *Psychological Monographs: General and Applied,* 68(8), 1-13.

Liberman, A. M., Harris, K. S., Hoffman, H. S., & Griffith, B. C. (1957). The discrimination of speech sounds within and across phoneme boundaries. *Journal of Experimental Psychology,* 54(5), 358-368.

Liberman, A. M., Cooper, F. S., Shankweiler, D. P., & Studdert-Kennedy, M. (1967). Perception of the speech code. *Psychological review,* 74(6), 431-461.

Lisker, L., & Abramson, A. S (1964). A cross-language study of voicing in initial stops: Acoustical measurements. *Word*, 20, 384-422.

Lloyd-Fox, S., Blasi, A., Volein, A., Everdell, N., Elwell, C. E., & Johnson, M. H. (2009). Social perception in infancy: a near infrared spectroscopy study. *Child development*, 80(4), 986‐999.

Lloyd-Fox, S., Blasi, A., Elwell, C. E., (2010). Illuminating the developing brain: the past, present and future of functional near infra red spectroscopy. *Neuroscience and Biobehavioral Reviews,* 34 (3), 269–284.

Maddieson, I. (1984). Phonetic cues to syllabification. *University of California Los Angeles Working Papers in Phonetics,* 59, 85-101

Maddieson, I. (2011). Lateral consonants. The World Atlas of Language Structures Online', Max Planck Digital Library, Munich. (Available online at http://wals.info/chapter/3).

Mandel, D. R., Jusczyk, P. W., & Pisoni, D. B. (1995). Infants' recognition of the sound patterns of their own names. *Psychological Science, 6*, 314–317.

Mani, N., & Plunkett, K. (2007). Phonological specificity of vowels and consonants in early lexical representations. *Journal of Memory and Language*, 57(2), 252‐272.

Mani, N., & Plunkett, K. (2008). Fourteen-month-olds pay attention to vowels in novel words. *Developmental Science,* 11, 53–59.

Mani, N., & Plunkett, K. (2010). Twelve-month-olds know their cups from their keps and tups. *Infancy.* 15(5), 445–470

Marcus, G. F., Vijayan, S., Rao, S. B., & Vishton, P. M. (1999). Rule learning by seven-month-old infants. *Science,* 283(5398), 77–80.

Mattock, K., Polka, L., Rvachew, S., & Krehm, M., (2010). The first steps in word learning are easier when the shoes fit: comparing monolingual and bilingual infants. *Developmental Science*, 13, 229-243.

May, P., Tiitinen H., Ilmoniemi R. J., Nyman G., Taylor J. G., & Näätänen, R. (1999). Frequency change detection in human auditory cortex. *Journal of Computational Neuroscience*, 6, 99–120.

Meek, J. (2002). Basic principles of optical imaging and application to the study of infant development. *Developmental Science*, 5(3), 371–380.

Mehler, J., Jusczyk, P., Lambertz, G., Halsted, N., Bertoncini, J., & Amiel-Tison, C. (1988). A precursor of language acquisition in young infants. *Cognition,* 29(2), 143-178.

Mehler, J., Peña, M., Nespor, M., & Bonatti, L. (2006). The « soul » of language does not use statistics: Reflections on vowels and consonants. *Cortex,* 42(6), 846–854.

Mersad, K., & Nazzi, T. (2012). When Mommy comes to the rescue of statistics: Infants combine top-down and bottom-up cues to segment speech. *Language Learning and Development,* 8(3), 303‐315.

Miller, E. K., & Desimone, R. (1994). Parallel neuronal mechanisms for short-term memory. *Science,* 263, 520–522.

Minagawa-Kawai, Y., Mori, K., Hebden, J. C., & Dupoux, E. (2008). Optical imaging of infants' neurocognitive development: recent advances and perspectives. *Developmental neurobiology,* 68(6), 712‐728.

Minagawa-Kawai, Y., Cristia, A., & Dupoux E. (2011). Cerebral lateralization & early speech acquisition: a developmental scenario. *Developmental Cognitive Neuroscience*, 1, 217-232.

Morgan, J. L., & Demuth, K. (1996). Signal to Syntax: an overview. In J.L. Morgan & K. Demuth (Eds.), *Signal to syntax: Bootstrapping from speech to grammar in early acquisition* (pp. 1–22). Mahwah, NJ: Lawrence Erlbaum Associates.

Morse, P. A., & Snowdon, C. T. (1975). An investigation of categorical speech discrimination by rhesus monkeys. *Perception & psychophysics,* 17(1), 9-16.

Näätänen, R (1984) In search of a short-duration memory trace of a stimulus in the human brain. In L Pulkkinen &P Lyytinen (Eds), *Human action and personality. Essays in honour of Martii Takala* (pp 29–43) Jyvaeskylae: University of Jyvaeskylae.

Näätänen, R. (2001). The perception of speech sounds by the human brain as reflected by the mismatch negativity (MMN) and its magnetic equivalent (MMNm). *Psychophysiology,* 38(1), 1-21.

Näätänen, R., & Rinne, T. (2002). Electric brain response to sound repetition in humans: an index of long-term-memory – trace formation? *Neuroscience Letters*, 318, 49–51.

Naccache, L., & Dehaene, S. (2001). The priming method: imaging unconscious repetition priming reveals an abstract representation of number in the parietal lobes. *Cerebral Cortex*, 11, 966 –974.

Nation, P. & Waring, R. (1997). Vocabulary size, text coverage, and word lists. In Schmitt, N. and McCarthy, M., (Eds). *Vocabulary: description, acquisition, and pedagogy.* Cambridge: Cambridge University Press.

Nazzi, T. (2005). Use of phonetic specificity during the acquisition of new words: differences between consonants and vowels. *Cognition*, 98(1), 13-30.

Nazzi, T., Bertoncini, J., & Mehler, J. (1998). Language discrimination by newborns: Toward an understanding of the role of rhythm. *Journal of Experimental Psychology: Human Perception and Performance,* 24(3), 756-766.

Nazzi, T., Floccia, C., & Bertoncini, J. (1998). Discrimination of pitch contours by neonates. *Infant Behavior and Development*, 21(4), 779–784.

Nazzi, T., & Bertoncini, J. (2009). Phonetic specificity in early lexical acquisition: New evidence from consonants in coda positions. *Language and Speech*, 52, 463–480.

Nazzi, T., Floccia, C., Moquet, B., & Butler, J. (2009). Bias for consonantal information over vocalic information in 30-month-olds: cross-linguistic evidence from French and English. *Journal of Experimental Child Psychology,* 102(4), 522–537.

Nazzi, T., & Gopnik, A. (2001). Linguistic and cognitive abilities in infancy: When does language become a tool for categorization? *Cognition,* 80(3), B22-B20.

Nazzi, T., Iakimova, G., Bertoncini, J., Frédonie, S., & Alcantara, C. (2006). Early segmentation of fluent speech by infants acquiring French: Emerging evidence for crosslinguistic differences. *Journal of Memory and Language,* 54(3), 283–299.

Nazzi, T., Mersad, K., Sundara, M., Iakimova, G., & Polka, L. (2014). Early word segmentation in infants acquiring Parisian French: task-dependent and dialect-specific aspects. *Journal of Child Language,* 41(03), 600-633

Nespor, M., Peña, M., & Mehler, J. (2003). On the different roles of vowels and consonants in speech processing and language acquisition. *Lingue e linguaggio,* 2(2), 203–230.

Nespor, M., & Vogel, I. (1986). *Prosodic phonology* (Vol. 28). Dordrecht, Holland ; Riverton, N.J., U.S.A.: Foris.

New, B., Araújo, V., & Nazzi, T. (2008). Differential processing of consonants and vowels in lexical access through reading. *Psychological Science*, 19(12), 1223–7.

New, B., & Nazzi, T. (2014). The time course of consonant and vowel processing during word recognition. *Language, Cognition and Neuroscience*, 29(2), 147–157

New, B., Pallier, C., Ferrand, L., & Matos, R. (2001). Une base de données lexicales du français contemporain sur internet: LEXIQUETM. *L'année Psychologique*, 101(3), 447-462.

Nguyen, D.K., Tremblay, J., Pouliot, P., Vannasing, P., Florea, O., Carmant, L., Lepore, F., Sawan, M., Lesage, F., & Lassonde, M., (2011). Non-invasive continuous EEG-fNIRS recording of temporal lobe seizures*. Epilepsy Research,* 99(1-2), 112-126.

Nishibayashi, L.-L., Goyet, L., & Nazzi, T. (2014). Early speech segmentation in French-learning infants: monosyllabic words versus embedded syllables. *Language & Speech,* in press.

Nishibayashi, L. & Nazzi, T. (2014, July). *Asymmetry in consonant/vowel processing in recognizing segmented word forms: evidence of a switch between 6 and 8 months of age.* Poster presented at 19th Biennal Meeting of the International Society for Infant Studies, Berlin, Germany.

Obrig, H., & Villringer, A. (2003). Beyond the visible-imaging the human brain with light*. Journal of Cerebral Blood Flow & Metabolism*, 23(1), 1–18.

Obrig, H., Wenzel, R., Kohl, M., Horst, S., Wobst, P., Steinbrink, J., Thomas, F., & Villringer, A. (2000). Near-infrared spectroscopy: does it function in functional activation studies of the adult brain? *International journal of psychophysiology,* 35(2-3), 125‐142.

Obrig, H., Rossi, S., Telkemeyer, S., & Wartenburger, I. (2010). From acoustic segmentation to language processing: evidence from optical imaging*. Frontiers in Neuroenergetics,* 2(13).

Obrig, H., (2014). NIRS in clinical neurology–a "promising" tool? *Neuroimage* 85, 535–546.

Okada, E., Firbank, M., & Delpy, D. T. (1995). The effect of overlying tissue on the spatial sensitivity profile of near-infrared spectroscopy. *Physics in Medecine and Biology*, 40, 2093–2108.

Otsuka, Y., Nakato, E., Kanazawa, S., Yamaguchi, M. K., Watanabe, S., & Kakigi, R. (2007). Neural activation to upright and inverted faces in infants measured by near infrared spectroscopy. *NeuroImage,* 34(1), 399‐406.

Owren, M. J., & Cardillo, G. C. (2006). The relative roles of vowels and consonants in discriminating talker identity versus word meaning. *The Journal of the Acoustical Society of America*, 119(3), 1727–1739.

Pater, J., Stager, C. L., & Werker, J. F. (2004). The perceptual acquisition of phonological contrasts. *Language,* 80, 384-402.

Peña, M., Maki, A., Kovacic, D., Dehaene-Lambertz, G., Koizumi, H., Bouquet, F., & Mehler, J. (2003). Sounds and silence: An optical topography study of language recognition at birth. *Proceedings of the National Academy of Sciences,* 100(20), 11702-11705.

Pharao, N. (2011). *Plosive reduction at the group level and in the individual speaker.* Paper presented at the International Congress of Phonetic Sciences XVII 2011, Hong Kong.

Pike, K. L. (1945*). The intonation of American English*. Ann Arbor: University of Michigan Press.

Pisoni, D. B. (1973). Auditory and phonetic memory codes in the discrimination of consonants and vowels. *Perception & psychophysics*, 13(2), 253-260.

Polka, L., & Bohn, O. S. (1996). A cross-language comparison of vowel perception in English-learning and German-learning infants. *The Journal of the Acoustical Society of America,* 100(1), 577-592.

Polka, L., & Sundara, M. (2012). Word segmentation in monolingual infants acquiring Canadian-English and Canadian-French: Native language, cross-language and cross-dialect comparisons. *Infancy,* 17, 198-223.

Polka, L., & Werker, J. F. (1994). Developmental changes in perception of nonnative vowel contrasts. *Journal of Experimental Psychology: Human Perception and Performance,* 20(2), 421.

Poltrock, S., & Nazzi, T. (in revision). Consonant/vowel asymmetry in early word form recognition.

Pons, F. (2006). The effects of distributional learning on rats' sensitivity to phonetic information. *Journal of Experimental Psychology: Animal Behavior Processes*, 32, 97–101.

Pons, F., & Toro, J. M. (2010). Structural generalizations over consonants and vowels in 11-month-old infants. *Cognition,* 116(3), 361-367.

Quaresima, V., Bisconti, S., & Ferrari, M., (2012). A brief review on the use of functional near-infrared spectroscopy (fNIRS) for language imaging studies in human newborns and adults. *Brain & Language*, 121, 79–89.

Ramon-Casas, M., Swingley, D., Bosch, L., & Sebastián-Gallés, N. (2009). Vowel categorization during word recognition in bilingual toddlers. *Cognitive Psychology,* 59, 96-121.

Ramus, F., Hauser, M. D., Miller, C., Morris, D., & Mehler, J. (2000). Language discrimination by human newborns and by cotton-top tamarin monkeys. *Science,* 288(5464), 349-351.

Ramus, F., & Mehler, J. (1999). Language identification with suprasegmental cues: a study based on speech resynthesis. *Journal of the Acoustical Society of America*, 105(1), 512‑521.

Ramus, F., Nespor, M., & Mehler, J. (1999). Correlates of linguistic rhythm in the speech signal. *Cognition*, 73(3), 265‑292.

Repp, B. H. (1984). Categorical perception: issues, methods, findings. In N. J. Lass (ed.), *Speech and language : advances in basic research and practice*, (Vol. 10, p. 243–335). New York : Academic Press.

Rivera-Gaxiola, M., Silva-Pereyra, J., & Kuhl, P. K., (2005). Brain potentials to native and non-native speech contrasts in 7- and 11-month-old American infants. *Developmental Science,* 8, 162–172.

Rossi, S., Jurgenson, I.B., Hanulikova, A., Telkemeyer, S., Wartenburger, I., Obrig, H., (2011). Implicit processing of phonotactic cues: evidence from electrophysiological and vascular responses. *Journal of Cognitive Neuroscience,* 23, 1752–1764.

Rossi, S., Telkemeyer, S., Wartenburger, I., Obrig, H., (2012). Shedding light on words and sentences: near-infrared spectroscopy in language research. *Brain & Language*, 121, 152–163.

Saffran, J., Newport, E. L., & Aslin, R. N. (1996a). Word Segmentation: The Role of Distributional Cues. *Journal of Memory and Language,* 35(4), 606–621.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996b). Statistical learning by 8-month-old infants. *Science,* 274(5294), 1926‑1928.

Saffran, J. R., Werker, J. F., & Werner, L. A. (2007). *The infant's auditory world: Hearing, speech, and the beginnings of language*. Handbook of child psychology.

Sakatani, K., Chen, S., Lichty, W., Zuo, H., & Wang, Y. P. (1999). Cerebral blood oxygenation changes induced by auditory stimulation in newborn infants measured by near infrared spectroscopy. *Early human development*, 55(3), 229‑236.

Sakoe, H. & Chiba, S. (1978). Dynamic programming algorithm optimization for spoken word recognition. *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, 26*, 33-49.

Sansavini, A., Bertoncini, J., & Giovanelli, G. (1997). Newborns discriminate the rhythm of multisyllabic stressed words. *Developmental psychology*, 33(1), 3‑11.

Sandman, C. A., Wadhwa, P., Hetrick, W., Porto, M., & Peeke, H. V. S. (1997). Human Fetal Heart Rate Dishabituation between Thirty and Thirty-Two Weeks Gestation. *Child Development*, 68 (6), 1031-1040.

Sato, H., Hirabayashi, Y., Tsubokura, H., Kanai, M., Ashida, T., Konishi, I., Ushida-Ota, M., Konishi, Y., & Maki, A. (2012). Cerebral hemodynamics in newborn infants exposed to speech sounds: A whole-head optical topography study. *Human brain mapping,* 33(9), 2092‑2103.

Scholkmann, F., Kleiser, S., Metz, A. J., Zimmermann, R., Mata Pavia, J., Wolf, U., & Wolf, M., (2014). A review on continuous wave functional near-infrared spectroscopy and imaging instrumentation and methodology. *NeuroImage* 85, 6–27.

Schwytay J., & Höhle, B. (2014, July). *Different task demands as a filter for the sensitivity to consonants in early word learning: Evidence from German 20-month-olds.* Poster presented at the 13th International Congress for the Study of Child Language, Amsterdam, The Netherlands.

Sebastián-Gallés, N., & Bosch, L. (2009). Developmental shift in the discrimination of vowel contrasts in bilingual infants: is the distributional account all there is to it? *Developmental Science*, 12, 874-887.

Sebastián-Gallés, N., Echeverria S., & Bosch L. (2005). The influence of initial exposure on lexical representation: Comparing early and simultaneous bilinguals. *Journal of Memory and Language*. 52, 240-255.

Segaert, K., Weber. K., de Lange F. P., Petersson, K. M., & Hagoort, P. (2013). The suppression of repetition enhancement: A review of fMRI studies. *Neuropsychologia,* 51, 59-66.

Selkirk, E. (1974). French liaison and the X notation*. Linguistic Inquiry,* 573–590.

Shapiro, A. C., Rogers, T. T., Cordova, N. I., Turk-Browne, N. B., & Botvinick, M. M. (2013). Neural representations of events arise from temporal community structure. *Nat Neuroscience,* 16(4), 486–492.

Sharp, D. J., Scott, S. K., Cutler, A., & Wise, R. J. S. (2005). Lexical retrieval constrained by sound structure: The role of the left inferior frontal gyrus. *Brain & Language*, 92, 309–319.

Shibata, H., Suzuki, M., Gyoba, J., (2007). Cortical activity during the recognition of cooperative actions. *Neuroreport,* 18 (7), 697–701.

Shimada, S., & Hiraki, K. (2006). Infant's brain responses to live and televised action. *NeuroImage,* 32(2), 930‐939.

Sinnott, J. M. (1989). Detection and discrimination of synthetic English vowels by Old World monkeys (Cercopithecus, Macaca) and humans*. The Journal of the Acoustical Society of America,* 86, 557‐565.

Sinnott, J. M., & Kreiter, N. A. (1991). Differential sensitivity to vowel continua in Old World monkeys (Macaca) and humans*. The Journal of the Acoustical Society of America,* 89(5), 2421–2429.

Soares, A. P., Perea, M., & Comesaña, M. (2014). Tracking the emergence of the consonant bias in visual-word recognition: Evidence with developing readers. *PloS one,* 9(2), e88580.

Southgate, V., Begus, K., Lloyd-Fox, S., di Gangi, V., Hamilton, A., (2014). Goal representation in the infant brain. *Neuroimag*e, 85, 294–301.

Stager, C. L. & Werker, J. F. (1997). Infants listen for more phonetic detail in speech perception than in word-learning tasks. *Nature, 388*, 381-382.

Stevens, K. N., & Blumstein, S. E. (1978). Invariant cues for place of articulation in stop consonants. *The Journal of the Acoustical Society of America,* 64(5), 1358-1368.

Streeter, L. A. (1976). Language perception of two-month-old infants shows effects of both innate mechanisms and experience. *Nature,* 259, 39-41.

Sundara, M., & Scutellaro, A. (2011). Rhythmic distance between languages affects the development of speech perception in bilingual infants. *Journal of Phonetics,* 39, 505-513.

Swain, I. U., Zelazo, P. R., & Clifton, R. K. (1993). Newborn infants' memory for speech sounds retained over 24 hours. *Developmental psychology,* 29(2), 312-323.

Swingley, D., & Aslin, R. N. (2000). Spoken word recognition and lexical representation in very young children. *Cognition,* 76, 147-166.

Swingley, D. (2005). 11-Month-Olds' Knowledge of how familiar words sound. *Developmental Science,* 8(5), 432-443.

Taga, G., & Asakawa, K. (2007). Selectivity and localization of cortical response to auditory and visual stimulation in awake infants aged 2 to 4 months. *NeuroImage,* 36(4), 1246-1252.

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. *The Journal of Neuroscience*, 29 (47), 14726–14733.

Thiessen, E. D., & Saffran, J. R. (2003). When cues collide: Use of stress and statistical cues to word boundaries by 7- to 9-month-old infants. *Developmental psychology*, 39(4), 706-716.

Tincoff, R., & Jusczyk, P. W. (1999). Some beginnings of word comprehension in 6-month-olds. *Psychological Science,* 10(2), 172-175.

Tincoff, R., & Jusczyk, P. W. (2012). Six-Month-Olds Comprehend Words That Refer to Parts of the Body. *Infancy,* 17(4), 432-444.

Toro, J. M., Nespor, M., Mehler, J., & Bonatti, L. L. (2008). Finding words and rules in a speech stream functional differences between vowels and consonants. *Psychological Science,* 19(2), 137–144.

Toro, J. M., Shukla, M., Nespor, M., & Endress, A. D. (2008). The quest for generalizations over consonants: asymmetries between consonants and vowels are not the by-product of acoustic differences. *Perception & Psychophysics,* 70(8), 1515-1525.

Torricelli, A., Contini, D., Pifferi, A., Spinelli, L., & Cubeddu, R. (2008). Functional brain imaging by multi-wavelength time-resolved near infrared spectroscopy. *Opto-Elecetronics Review,* 16(2), 131-135.

Trecca, F. Bleses, D., Christiansen, M. H., Basbøll, H. Højen, A., Madsen, T. O. & Andersen, S. R. (2014, july). *The Effect of Vocalic vs. Consonantal Phonetic Structure on Language Segmentability: the Case of Danish.* Poster presented at the 13th International Congress for the Study of Child Language, Amsterdam, The Netherlands.

Turk-Browne, N. B., Yi, D. J., & Chun, M. M. (2006). Linking implicit and explicit memory: common encoding factors and shared representations. *Neuron,* 49, 917–927.

Turk-Browne, N. B., Yi, D. J., Leber, A. B., & Chun, M. M. (2007). Visual quality determines the direction of neural repetition effects. *Cerebral Cortex,* 17, 425–43.

Turk-Browne, N. B., Scholl, B. J., & Chun, M. M. (2008). Babies and brains: habituation in infant cognition and functional neuroimaging. *Frontiers of Psychology,* 2, 1-11.

Valiante, A. G., Barr, R. G., Zelazo, P. R., Papageorgiou, A. N., & Young, S. N. (2006). A typical feeding enhances memory for spoken words in healthy 2-to 3-day-old newborns. *Pediatrics*, 117(3), e476–e486.

Valiante, A. G., Barr, R. G., Zelazo, P. R., Brant, R., & Young, S. N. (2013). Effects of Familiarity and Feeding on Newborn Speech–Voice Recognition. *Infancy,* 18(4), 443‐461.

Van Ooijen, B. (1996). Vowel mutability and lexical selection in English: evidence for a word reconstruction task. *Memory and Cognition*, 24, 573-583.

Vergara-Martinez, M., Perea, M., Marin, A., & Carreiras, M. (2011). The processing of consonants and vowels during letter identity and letter position assignment in visual-word recognition: An ERP study. *Brain & Language* 118: 105–117.

Vihman, M. M., Nakai, S., DePaolis, R. A., & Hallé, P. (2004). The role of accentual pattern in early lexical representation. *Journal of Memory and Language*, 50, 336-353.

Villringer, A., Planck, J., Hock, C., Schleinkofer, L., Dirnagl, U., 1993. Near infrared spectroscopy (NIRS): a new tool to study hemodynamic changes during activation of brain function in human adults. *Neuroscience Letters.* 154, 101–104.

Vuilleumier, P., Schwartz, S., Duhoux, S., Dolan, R. J., & Driver, J. (2005). Selective attention modulates neural substrates of repetition priming and ''implicit'' visual memory: Suppressions and enhancements revealed by fMRI. *Journal of Cognitive Neuroscience*, 17(8), 1245–1260.

Wagner, J. B., Fox, S. E, Tager-Flusberg, H., & Nelson, C. A. (2011). Neural processing of repetition and non-repetition grammars in 7- and 9-month-old infants. *Frontiers in Psychology,* 2, 168.

Wallois, F., Mahmoudzadeh, M., Patil, A., & Grebe, R. (2012). Usefulness of simultaneous EEG–NIRS recording in language studies. *Brain and language*, 121(2), 110–123.

Watanabe, E., Nagahori, Y., & Mayanagi, Y., (2002). Focus diagnosis of epilepsy using near-infrared spectroscopy. *Epilepsia*, 43 (9), 50–55.

Wartenburger, I., Steinbrink, J., Telkemeyer, S., Friedrich, M., Friederici, A. D., & Obrig, H. (2007). The processing of prosody: Evidence of interhemispheric specialization at the age of four. *NeuroImage,* 34(1), 416-425.

Waters, R. A., & Wilson, W. A. (1976). Speech perception by rhesus monkeys: the voicing distinction in synthesized labial and velar stop consonants. *Perception & psychophysics,* 19, 285-289.

Weissenborn, J., & Höhle, B. (2001). *Approaches to Bootstrapping: Phonological, Lexical, Syntactic and Neurophysiological Aspects of Early Language Acquisition*. John Benjamins Publishing Company.

Werner, L. A., Marean, G. C., Halpin, C. F., Spetner, N. B., & Gillenwater, J. M. (1992). Infant auditory temporal acuity: Gap detection. *Child Development,* 63, 260-272.

Werker, J. F., & Tees, R. C. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development,* 7, 49-63.

Werker, J. F., & Lalonde, C. E. (1988). Cross-language speech perception: initial capabilities and developmental change. *Developmental psychology*, 24(5), 672.

Werker, J.F., & McLeod, P.J. (1989). Infant preference for both male and female infant-directed talk: A developmental study of attentional and affective responsiveness. *Canadian Journal of Psychology, 43*, 230–246.

Werker, J. F. (1994). Cross-language speech perception: development change does not involve loss*.* In J. C. Goodman & H. C. Nusbaum (Eds.) *The development of speech perception: the transition from speech sounds to spoken words* (pp. 93–120). Cambridge, MA: MIT Press.

Werker, J. F., & Curtin, S. (2005). PRIMIR: A developmental framework of infant speech processing. *Language Learning and Development*, 1, 197-234.

Werker, J. F., Fennell, C. T., Corcoran, K.M., & Stager, C.L. (2002). Infants' ability to learn phonetically similar words: effects of age and vocabulary. *Infancy,* 3, 1–30.

Werker, J. F., & Yeung, H. H. (2005). Infant speech perception bootstraps word learning. *Trends in Cognitive Sciences*, 9(11), 519-527.

White, K. S., & Morgan, J. L. (2008). Sub-segmental detail in early lexical representations. *Journal of Memory and Language*, 59(1), 114-132.

White, L., & Mattys, S. L. (2007). Calibrating rhythm: First language and second language studies. *Journal of Phonetics,* 35(4), 501–522.

Wiggs, C. L., Martin, A. (1998). Properties and mechanisms of perceptual priming. *Current Opinon in Neurobiolology*, 8, 227–233.

Wolf, M., Ferrari, M., & Quaresima, V. (2007). Progress of near-infrared spectroscopy and topography for brain and muscle clinical applications. *Journal of biomedical optics,* 12(6), 062104-062104.

Zesiger, P., & Jöhr, J. (2011). Les représentations phonologiques des mots chez le jeune enfant. Phonological representations of words in young infants. *Enfance,* 3, 293–309.

Zesiger, P., Dupuis Lozeron, E., Lévy, A., & Frauenfelder, U., (2012). Phonological specificity in 12- and 17- month old French-speaking infants*. Infancy*, 17, 591-609.

APPENDICES

DESC	12242	WILEY	Dispatch: 29.9.14	CE: Wiley
Journal Code	Manuscript No.		No. of pages: 12	PE: Sharanya

Developmental Science (2014), pp 1–12 DOI: 10.1111/desc.12242

PAPER

Call me Alix, not Elix: vowels are more important than consonants in own-name recognition at 5 months

Camillia Bouchon,[1,2] Caroline Floccia,[3] Thibaut Fux,[4] Martine Adda-Decker[4] and Thierry Nazzi[1,2]

1. *Université Paris Descartes, Sorbonne Paris Cité, France*
2. *CNRS – Laboratoire de Psychologie de la Perception (UMR 8242), Paris, France*
3. *School of Psychology, Plymouth University, UK*
4. *CNRS – Laboratoire de Phonétique et de Phonologie (UMR 7018), Paris, France*

Abstract

Consonants and vowels differ acoustically and articulatorily, but also functionally: Consonants are more relevant for lexical processing, and vowels for prosodic/syntactic processing. These functional biases could be powerful bootstrapping mechanisms for learning language, but their developmental origin remains unclear. The relative importance of consonants and vowels at the onset of lexical acquisition was assessed in French-learning 5-month-olds by testing sensitivity to minimal phonetic changes in their own name. Infants' reactions to mispronunciations revealed sensitivity to vowel but not consonant changes. Vowels were also more salient (on duration and intensity) but less distinct (on spectrally based measures) than consonants. Lastly, vowel (but not consonant) mispronunciation detection was modulated by acoustic factors, in particular spectrally based distance. These results establish that consonant changes do not affect lexical recognition at 5 months, while vowel changes do; the consonant bias observed later in development does not emerge until after 5 months through additional language exposure.

Research highlights

- Consonants are more important in distinguishing words in the lexicon and are accordingly used preferentially by toddlers and adults in word processing.
- Using HPP and a controlled design, we measured French-learning 5-month-old infants' preferences for their correctly pronounced over a minimally mispronounced version of their own name, one of the first familiar words. Behavioral results indicated sensitivity to vowel changes, and not to consonant changes.
- Moreover, detailed acoustic analyses linked individual performance to spectrally based distance only for vowels.
- This shows that 5-month-old infants do not yet attribute a privileged role to consonants when recognizing their name, thus highlighting how the complex links between speech processing and lexical acquisition in the first months of life will necessarily change with phonological and lexical development to allow the acquisition of the biases found in toddlerhood/adulthood.

Introduction

Consonants and vowels are the two basic sound categories central to the structure of speech in all languages (Ladefoged, 1993). They differ in many respects: For example, vowels tend to be longer and louder than consonants (Repp, 1984), and are thus perceived more clearly in utero (Granier-Deferre, Ribeiro, Jacquet & Bassereau, 2011) and at birth (Bertoncini, Bijeljac-Babic, Jusczyk & Kennedy, 1988; Benavides-Varela, Hochmann, Macagno, Nespor & Mehler, 2012). Vowels are also less numerous than consonants in most languages (Maddieson, 1984). Furthermore, consonants are

Address for correspondence: Camillia Bouchon, Laboratoire Psychologie de la Perception, 45 rue des Saint-Pères, 75006 PARIS, France; e-mail: camillia.bouchon@gmail.com

processed more categorically than vowels (Fry, Abramson, Eimas & Liberman, 1962), and by partly different brain areas as shown by neuropsychological (Caramazza, Chialant, Capasso & Miceli, 2000) and electrophysiological/brain imaging studies (Carreiras & Price, 2008). During development, native vowel categories are learned earlier (6 months: Kuhl, Williams, Lacerda, Stevens & Lindblom, 1992) than consonant categories (10–12 months; Werker & Tees, 1984).

These differences between consonants and vowels have led to the proposal of a 'division of labor' that could help infants learn their native language (Nespor, Peña & Mehler, 2003). Specifically, two complementary functional biases have been proposed, such that '. . .consonants, rather than vowels, are most relevant to build the lexicon, and vowels, rather than consonants, are most relevant for grammatical [and prosodic] information' (p. 224). This hypothesis has been extremely influential in the field (see below), but very few studies have explored the origins of these hypothesized learning biases. Here we explore the respective role of consonants and vowels in the earliest steps of lexical acquisition, focusing on the emergence of the consonant bias for lexical processing.

Two kinds of hypotheses regarding the origin of the consonant bias have been offered. The 'initial bias' hypothesis states that infants start processing consonants and vowels as distinctive linguistic categories from birth, ascribing a limited role to input characteristics and thus predicting no developmental or cross-linguistic differences (Nespor *et al.*, 2003; Bonatti, Peña, Nespor & Mehler, 2007; Pons & Toro, 2010). Alternatively, 'learned bias' hypotheses propose that this bias emerges during development, as a result of infants' acquisition of the acoustic-phonetic (Floccia, Nazzi, Delle Luche, Poltrock & Goslin, 2014) or lexical (Keidel, Jenison, Kluender & Seidenberg, 2007) properties of their native language.

Many adult studies, using various tasks, have shown that consonants are privileged over vowels in lexical processing in French, Spanish, Italian, English, and Dutch. This was found in tasks measuring lexical access in both auditory (Cutler, Sebastián-Gallés, Soler-Vilageliu & Van Ooijen, 2000; Delle Luche, Poltrock, Goslin, New, Floccia & Nazzi, 2014) and written modalities (Acha & Perea, 2010; New, Araùjo & Nazzi 2008), detection of word–forms from continuous speech (Toro, Nespor, Mehler & Bonatti, 2008), and novel word learning (Havy, Serres & Nazzi, 2014).

Several studies also tested the consonant bias in toddlers and children. These studies, initially conducted in French, revealed a consonant bias in novel word learning between 16 months and 5 years of age (Nazzi, 2005; Havy, Bertoncini & Nazzi, 2011; Havy & Nazzi, 2009; Havy *et al.*, 2014), and a consonant bias in familiar word recognition at 14 months (Zesiger & Jöhr, 2011). These findings establish that the consonant bias in (French) lexical processing is consistently present from early in development, and converging evidence has been found at 12 months in Italian-learning infants (Hochmann, Benavides-Varela, Nespor & Mehler, 2011). However, studies on English-learning infants offer a different picture. A consonant bias was found in word learning tasks at 30 months but not 16 and 23 months (Floccia *et al.*, 2014; Nazzi, Floccia, Moquet & Butler, 2009) and in word recognition tasks at 15 months but not 12, 18 or 24 months (Mani & Plunkett, 2007, 2010). These latter findings appear to go against the initial bias hypothesis, which predicts an early and language-independent consonant bias. However, it remains possible that the consonant bias would only be momentarily masked in English-learning toddlers.

Therefore, the current study aimed to specify its origin and early trajectory by testing whether the consonant bias is present from the very beginning of lexical acquisition, by 5 months of age. This age was chosen given data establishing recognition of some familiar word forms (Mandel, Jusczyk & Pisoni, 1995; Bortfeld, Morgan, Golinkoff & Rathbun, 2005; Mersad & Nazzi, 2012) and comprehension of some words (Tincoff & Jusczyk, 1999, 2012; Bergelson & Swingley, 2012) at 5–6 months. Moreover, while some of these earlier studies had examined the effects of mispronunciations on early recognition of familiar words (consonant mispronunciation in English: Bortfeld *et al.*, 2005; vowel mispronunciation in French: Mersad & Nazzi, 2012), none had directly compared the effects of consonant and vowel mispronunciations. Here, we tested whether recognition of their own names by French-learning 5-month-olds is affected by a consonant change or a vowel change. Lastly, we also conducted acoustic analyses on the stimuli used in order to determine whether infants' performance is related to acoustic properties of the stimuli presented, in order to evaluate the acoustic/phonetic learning hypothesis (Floccia *et al.*, 2014).

Methods

Following Mandel *et al.* (1995), we used the Headturn Preference Procedure (HPP) to test the sensitivity of French-learning 5-month-olds to a consonant change (Consonant change condition) versus a vowel change (Vowel change condition) in their own name. Infants in the test groups were presented with repetitions of their

own correctly pronounced name (CPs) on half of the trials, and repetitions of their own mispronounced name (MPs) on the other half of the trials. We only used one-feature phonetic changes (Consonant change condition: place, voicing or manner; Vowel change condition: place, roundedness or height). Since all phonetic changes were native French contrasts (Dell, 1985), most, if not all, contrasts were expected to be distinguishable by our French-learning participants, given evidence that most native contrasts can be discriminated by infants from birth (for reviews, see Werker, 1994; Kuhl, Conboy, Coffey-Corina, Padden, Rivera-Gaxiola & Nelson, 2008). Importantly, the use of infants' own names allowed us to test many different consonant and vowel changes (25 in the Consonant change condition and 28 in the Vowel change condition, due to some infants having the same names), providing generalizable results. A preference for CPs over MPs (indicated by the difference in looking times towards CPs and MPs, referred to as *LT.diff* later on) would indicate sensitivity to the mispronunciation (e.g. in the Consonant change condition, a baby named *Victor* should prefer listening to *Victor* over *Zictor*). As a precaution, to rule out effects due to pure acoustic preferences, yoked control infants were tested on the same stimuli as test infants with the main constraints that they had a different name, did not know anyone with the critical name, and had a name starting with the same phoneme category (consonant or vowel) as the critical name (e.g. a baby named *Martin* was presented with *Victor* vs. *Zictor*). In this way, CPs and MPs were equally unfamiliar to control infants, and no preference should be observed for CPs over MPs (i.e. a baby named *Martin* should not prefer *Victor* over *Zictor*).

According to the initial bias hypothesis (Nespor *et al*., 2003), 5-month-olds should already be more sensitive to a consonant change than to a vowel change in their own name. Therefore, test infants should exhibit a larger preference for CPs over MPs in the Consonant change condition compared to the Vowel change condition, and no preference should be found in the yoked control groups of both experiments. Alternatively, if this prediction were not confirmed, it would suggest that the functional specialization of consonants and vowels still has to be learned at 5 months, as proposed by the 'learned bias' hypotheses. In this case, a possible outcome could be that 5-month-olds perceive and process consonant and vowel changes similarly in words, in which case we could predict a preference for their correctly pronounced names over their mispronunciations in both conditions. A further possibility is that 5-month-olds' reaction to the mispronunciations is based on the acoustic distance between the correctly versus mispronounced stimuli, independently (or not) of their consonant/vowel status. Therefore, in order to assess the contribution of the acoustic–phonetic properties of our stimuli to infants' preference responses, the consonant and vowel contrasts were characterized along three acoustic dimensions: duration, intensity and MFCCs (Mel-Frequency-Cepstrum Coefficients, a spectrally based measure of phonetically relevant acoustic information normalized for duration and intensity). These measures were used to evaluate the saliency and the discriminability of the different phonemes, and relate them to individual preference responses. We expected consonants to be overall less salient than vowels (Repp, 1984) but more discriminable as they are usually perceived more categorically than vowels (Fry *et al*., 1962).

Participants

All 120 participants were healthy French-learning monolingual 5-month-old infants (Table 1). For the test conditions, only infants without nicknames and who were frequently called by their own name were included. Infants in the control conditions were chosen so that they would not know anyone in their environment with the name used in the experiment. Children with names starting with a consonant were assigned to the Consonant change control group and those with names starting with a vowel to the Vowel change control group. Forty-seven additional infants were tested and their data excluded due to fussiness (36), having two consecutive trials with insufficient looking times, having more than three such trials overall (6), experimenter error (1), or being an outlier (*LT.Diff* above or below 2 *SDs* of the group Mean; 4).

Stimuli

Each of the 60 test infants heard repetitions of a pair of stimuli corresponding to their CP (Correctly

Table 1 *Participant information, illustration of the four experimental conditions, and examples of stimuli*

Groups (all n = 30)	Stimuli (example)	Infant's name (example)	Age in days (*SD*)	# girls/ boys
Consonant change condition				
Test	e.g. Victor vs. Zictor	Victor	164 (8)	14/16
Control	e.g. Victor vs. Zictor	Martin	164 (9)	11/19
Vowel change condition				
Test	e.g. Esther vs. Isther	Esther	164 (8)	17/13
Control	e.g. Esther vs. Isther	Adrien	165 (8)	14/16

pronounced) and MP (Mispronounced) names. Due to a few infants having the same names, there was a total of 28 pairs in the Consonant change condition and 25 pairs in the Vowel change condition. The same stimuli were used with the 60 yoked control infants. The MP of the names always consisted of a one-feature change (Table 2). As mentioned earlier, all changes were native French minimal contrasts according to Dell (1985), and were expected to be discriminable by French-learning infants irrespective of their age. Based on Mandel *et al*. (1995) and the fact that three different phonetic features were tested in each condition, the sample size was 30 infants in each test group. In the Consonant change condition, 10 infants were tested on a place-of-articulation change, 10 on a voicing change, and 10 on a manner-of-articulation change. In the Vowel change condition, 10 infants were tested on a place change, 10 on a roundedness change, and 10 on a height change.

For all infants, the same female native French speaker recorded 15 tokens each of CP and MP names. For both CPs and MPs, two files including the 15 tokens were made, the second file presenting the same tokens in reversed order. All sound files lasted 24 seconds.

Procedure

Behavioral experiment

Each infant was tested individually. The experiment was conducted inside a sound attenuated room, and in a booth made of pegboard panels (bottom) and a white curtain (top). The test booth had a red light and a loudspeaker (SONY xs-F1722) mounted at eye level on each of the side panels and a green light mounted on the center panel. Below the center light was a video camera used to monitor infants' behavior.

A PC computer terminal (Dell OptiPlex), a TV screen connected to the camera, and a response box were located outside the sound attenuated room. The response box, connected to the computer, was equipped with a series of buttons. The observer, who looked at the video of the infant on the TV screen to monitor the infant's looking behavior, pressed the buttons of the response box according to the direction the infant's head, thus starting and stopping the flashing of the lights and the presentation of the sounds (see below). The observer and the infant's caregiver wore earplugs and listened to masking music over tight-fitting closed headphones, which prevented either from hearing the stimuli presented. Information about the duration of the head-turn, calculated from the observer's button-pressing, was recorded by the computer.

We used the same version of the Head-turn Preference Procedure (HPP) as Mandel *et al*. (1995). Each infant was held on a caregiver's lap in the center of the booth. Each trial began with the green light on the center panel blinking until the infant oriented to it. Then, the red light on one of the side panels began to flash. When the infant turned their head in that direction, the stimulus for that trial began to play. The stimuli were delivered by the loudspeakers via an audio amplifier (Marantz PM4000). Each stimulus was played to completion or stopped immediately after the infant failed to maintain the head-turn for 2 consecutive seconds. If the infant turned away from the target by 30° in any direction for less than 2 s and then turned back again, the trial continued but the time spent looking away (when the experimenter released the buttons of the response box) was automatically subtracted from the listening time by the computer program. Thus, the maximum listening time for a given trial was the duration of the entire speech sample. If a trial lasted less than 1.5 s (insufficient looking time), the trial was repeated and the original listening time was discarded.

Each session began with two musical trials (excerpts of classical music), one on each side to give infants an opportunity to practice one head-turn to each side. The test phase consisted of eight trials divided into two blocks, in each of which the two lists of each name were presented. Order of the different lists within each block was randomized.

Acoustic analyses of the stimuli

Three acoustic dimensions were measured to characterize the contrasted phonemes of CPs and MPs: duration, intensity and Mel Frequency Cepstral Coefficients as a measure of spectral distance (MFCCs; see below for more explanation).

For each CP/MP pair, duration and intensity were measured for the 15 tokens of the contrasted phonemes using PRAAT. This was first used to calculate mean duration and mean intensity values of the contrasted phonemes, in order to compare the relative salience and discriminability of the contrasted consonants and vowels. Second, we computed normalized duration and intensity differences (*Diff.duration*: duration difference between the contrasted phonemes of CPs and MPs divided by their mean; same for *Diff.intensity*), in order to test their link with individual performance.

MFCCs are spectrum-based features resulting from a deconvolution of the speech source (e.g. vocal fold

Table 2 *Details of stimuli presented in the Consonant change condition (upper-panel) and the Vowel change condition (lower-panel), broken down by phonetic features (in columns): for each test infant, the correctly pronounced name (CP, corresponding to the test infant's name), the mispronunciation (MP), the phonetic feature contrasting the CP and MP (in International Phonetic Alphabet), and the control infant's name tested on this CP/MP pair are given*

Consonant change condition

Place changes					Voicing changes					Manner changes				
#	CP & Test infant's name	MP	contrast CP/MP	Control infant's name	#	CP & Test infant's name	MP	contrast CP/MP	Control infant's name	#	CP & Test infant's name	MP	contrast CP/MP	Control infant's name
1	Sophie	chophie	s/ʃ	Martin	1	Judith	chudith	ʒ/ʃ	Clément	1	Marie	barie	m/b	Félix
2	Rémi	lémi	R/l	Micha	2	Pénélope	bénélope	p/b	Martin	2	Marie	barie	m/b	Loic
3	Gaspard	daspard	g/d	Maya	3	Côme	gôme	k/g	Solène	3	Lily	zily	l/z	Gaspard
4	Camille	tamille	k/t	Serge	4	Véra	féra	v/f	Toto	4	Lucien	ducien	l/d	Marc-Antoine
5	Marin	narin	m/n	Rose	5	Jules	chules	ʒ/ʃ	Loulou	5	Nora	zora	n/z	Botho
6	Pauline	tauline	p/t	Robinson	6	Pierre	bierre	p/b	Nina	6	Bérénice	mérénice	b/m	Diane
7	Victor	zictor	v/z	Mathieu	7	Corentin	gorentin	k/g	Maeli	7	Téa	séa	t/s	Marguerite
8	Loulou	roulou	l/R	Basil	8	Violette	fiolette	v/f	Mathilde	8	Laomé	zaomé	l/z	Mathias
9	Félix	sélix	f/s	Martin	9	Félix	vélix	f/v	Mara	9	Nouchai	zouchai	n/z	Célian
10	Félix	sélix	f/s	Paulin	10	Charles	jarles	ʃ/ʒ	Mona	10	Sanjay	tanjay	s/t	Margot

Vowel change condition

Place changes					Roundedness changes					Height changes				
#	CP & Test infant's name	MP	contrast CP/MP	Control infant's name	#	CP & Test infant's name	MP	contrast CP/MP	Control infant's name	#	CP & Test infant's name	MP	contrast CP/MP	Control infant's name
1	Olivia	eulivia	o/ø	Ewan	1	Ellen	eullen	e/ø	Haythem	1	Alba	elba	a/ɛ	Ombelle
2	Auriane	euriane	o/ø	Anaïs	2	Elisa	eulisa	e/ø	Alice	2	Océane	ouceane	o/u	Anjali
3	Augustine	eugustine	o/ø	Alexandre	3	Iris	uris	i/y	Hédi	3	Alexandre	elexandre	a/ɛ	Héloise
4	Oriane	euriane	o/ø	Eloise	4	Eliott	euliott	e/ø	Andrée	4	Anais	enais	a/ɛ	Elio
5	Ulysse	oulysse	y/u	Hector	5	Ihsane	uhsane	i/y	Emmanuel	5	Arsène	ersene	a/ɛ	William
6	Hugo	ougo	y/u	Héloise	6	Elise	eulise	e/ø	Alexia	6	Honorine	ounorine	o/u	Ambroise
7	Augustine	eugustine	o/ø	Adèle	7	Erwan	eurwan	ɛ/œ	Abigaëlle	7	Esther	isther	e/i	Adrien
8	Hugo	ougo	y/u	Alix	8	Eliott	euliott	e/ø	Alexandre	8	Inès	enes	i/e	Adrien
9	Hugo	ougo	y/u	Aron	9	Eli	euli	e/ø	Aliénor	9	Alix	elix	a/ɛ	Ethan
10	Eugénie	ogenie	ø/o	Alienor	10	Elias	eulias	e/ø	Enzo	10	Arthur	erthur	a/ɛ	Evan

vibrations) and speech filter (vocal tract). MFCCs were chosen rather than pitch measures because, first, pitch measures cannot be calculated for some of the consonants used in our study (unvoiced consonants). Second, MFCCs are widely used both for automatic speech and speaker recognition, as they provide a general measure of distance between two speech sounds normalized for duration and intensity that specifies well both consonantal and vocalic information (however, MFCCs do not provide information regarding saliency, contrary to duration, intensity and pitch). They have been preferentially used in word and phoneme recognition studies because they retain phonetically relevant acoustic information (e.g. Davis & Mermelstein, 1980). They involve a pre-processing of the spectral envelope of the signal with frequency bands equally spaced on the Mel scale that approximates the psychoacoustic properties of the cochlea, thus providing a better acoustic/phonetic coding than more simple measures such as spectral distance and LFCC coefficients (Linear-Frequency Cepstral Coefficients).

MFCCs were calculated using 30 ms analysis windows at a 15 ms frame rate. To do so, for each CP/MP pair, the word-initial (contrasted) phoneme of the first of the 15 tokens of each word was manually segmented using Transcriber (Barras, Geoffroy, Wu & Libberman, 2001). Then, manual segmentation was used to automatically locate the initial phonemes of the 14 other tokens using dynamic time warping (DTW; Sakoe & Chiba, 1978). DTW is a speech comparison method that automatically determines the optimal temporal matching between two speech patterns (detect segment similarities) independently of duration and speech rate. Then, MFCCs were computed using the 24 triangular filters mel-frequency spaced, a standard discrete cosine's transform and frequency bandwidth of 0–8 kHz. The subset of MFCCs employed in the classification to measure MFCC distances included 12 coefficients, c1 to c12, in order to best represent the envelope of the mel-spectrum. Note that coefficient c0 was not taken into account to exclude intensity differences that would affect MFCC distance measurements. MFCC distances correspond to the Euclidian distance between two tokens calculated for the 12 coefficients (i.e. the square root of the summed squared differences between the two MFCC sets). Finally, we tested the link between individual performance and the normalized MFCC distance between CPs and MPs (*Diff.spectral*), defined for each CP/MP pair as the ratio of the mean cross-category distance between the 15 CPs and the 15 MPs (*Dcross*) of the given pair and the mean internal variability within the 15 CPs (*DwithinCP*) and within the 15 MPs (*DwithinMP*):

$$Diff.spectral(CP,MP) = \frac{Dcross(CP,MP)}{\sqrt{\frac{1}{2}(DwithinCP^2 + DwithinMP^2)}}$$

Results

Behavioral results

Overall analysis

Mean listening times (LTs) to the CP and MP names were calculated for each infant. Group averages are presented in Figure 1. A three-way ANOVA was conducted on LTs with a within-subjects factor of pronunciation (CP vs. MP) and between-subjects factors of group (test vs. control) and condition (Consonant change vs. Vowel change). Neither the effect of condition ($F(1, 116) = .26$, $p = .61$), group ($F(1, 116) = .87$; $p = .35$), nor the condition × pronunciation interaction ($F(1, 116) = .11$, $p = .74$) reached significance. The effect of pronunciation was only marginal ($F(1, 116) = 3.53$; $p = .06$), Importantly though, both the pronunciation × group interaction ($F(1, 116) = 7.23$; $p = .008$, $\eta^2_p = .06$) and the three-way interaction between pronunciation × group × condition ($F(1, 116) = 8.64$; $p = .004$, $\eta^2_p = .07$) reached significance, establishing that infants were not behaving in the same way in both conditions. In order to understand how consonant and vowels MPs were differently processed, separate analyses were conducted for each condition.

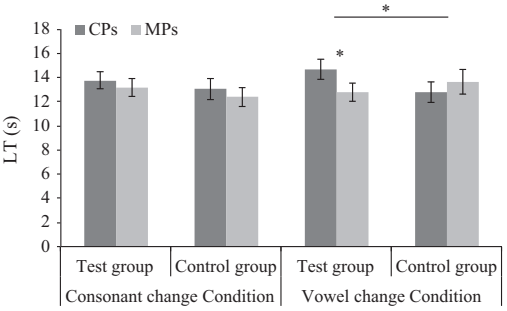


Figure 1 *Mean listening times (and SEs) to the CP and MP names in test and control infants. Left panel: Consonant change condition; Right panel: Vowel change condition.*

Consonant change condition

For the test group, the LT means were M_{CP} = 13.87 s (*SD* = 3.91 s) and M_{MP} = 13.36 s (*SD* = 3.98 s), the estimate of *LT.diff* was *Mtest* = .51 s (95% *CI* = [−.40, 1.42]) and Cohen's *d* = .21 (Cohen's *ds* are calculated with the difference of means as numerator and the standard deviation of the paired differences as denominator). For the control group, the LT means were M_{CP} = 13.03 s (*SD* = 4.74 s) and M_{MP} = 12.40s (*SD* = 4.38 s), the estimate of *LT.diff* was *Mctrl* = .64 s (95% *CI* = [−.51, 1.78]) and Cohen's *d* = .21. The size of the difference between *LT.diffs* in test and control infants was −.13 (95% *CI* = [−1.63, 1.37]) with Cohen's *d* = −.03. A two-way ANOVA on LTs with the factors of pronunciation (CP vs. MP) and group (test vs. control) was conducted. The effects of pronunciation (F(1, 58) = 2.56, p = .11), group (F(1, 58) = .75, p = .39), and the pronunciation × group interaction (F(1, 58) = .03, p = .86) all failed to reach significance. Longer LTs for CPs over MPs were found for 14 of the 30 test infants (binomial test, p = .43), and for 16 of the 30 control infants (binomial test, p = .43). This pattern of results shows no effect in either the test or control groups, and no difference between the two groups.

Moreover, a second ANOVA including the between-subjects factor of feature (place, voicing or manner; all n = 10) in addition to the previous factors yielded no effect of feature (F(2, 54) = 1.16; p = .32), and no interactions involving feature (all Fs < 1). Overall, these results show that French-learning 5-month-olds are not particularly sensitive to a consonant mispronunciation in their own name, independently of the phonetic feature contrasted.

Vowel change condition

For the test group, the LT means were M_{CP} = 14.75 s (*SD* = 4.72 s) and M_{MP} = 12.89 s (*SD* = 4.35 s), the estimate of *LT.diff* was *Mtest* = 1.86 s (95% *CI* = [.74, 2.89]) and Cohen's *d* = .62. For the control group, the LT means were M_{CP} = 12.76 s (*SD* = 4.24 s) and M_{MP} = 13.82 s (*SD* = 5.02 s), the estimate of *LT.diff* was *Mctrl* = −1.06 s (95% *CI* = [−2.10, −.01]) and Cohen's *d* = −.38. The size of the difference between *LT.diffs* in test and control infants was 2.92 (95% *CI* = [1.33, 4.51]) with Cohen's *d* = .68. A two-way ANOVA on LTs with the factors of pronunciation and group was conducted. The effect of pronunciation (F(1, 58) = 1.14, p = .29) and group (F(1, 58) = .22, p = .64) failed to reach significance, but the pronunciation × group interaction did (F(1, 58) = 15.11, p = .0003, η^2_p = .21). Planned comparisons revealed that Test infants significantly preferred their CP names compared to their MP versions (F(1, 58) = 12.28, p = .0009, η^2_p = .17) whereas Control infants marginally preferred MPs over CPs (F(1, 58) = 3.97, p = .051, η^2_p = .06). Note that this marginal (reversed) preference for vocalic MPs in Control infants, which goes against the Test infants' preference for their correctly pronounced names, confirms the strength of this preference in the Test infants. Longer LTs for CPs over MPs were found for 21 of the 30 test infants (binomial test, p = .02), but only for 12 of the 30 control infants (binomial test, p = .18). This pattern of results shows a medium effect in the test group (preference for CPs) and a small effect in the control group (preference for MPs), and a large difference between the two groups.

Moreover, a second ANOVA including the factor feature (place, roundedness, or height; all n = 10) in addition to the previous factors yielded no effect of feature and no interaction involving feature (all Fs < 1), while the pronunciation × group interaction was still significant (F(1, 54) = 14.43, p = .0004, η^2_p = .21). These results establish that French-learning 5-month-olds prefer their correctly pronounced name over a one-feature mispronunciation of that name, independently of the feature tested, as opposed to control infants who show a marginal preference for the mispronunciation. The present results thus exclude the possibility that infants in the Consonant change condition did not have a preference for their correctly pronounced name due to difficulties with our testing procedure.

Acoustic measures

Acoustic measures were conducted on the stimuli (see Stimuli section for details), which consisted of 28 different pairs of CPs-MPs in the Consonant change condition and 25 pairs of CPs-MPs in the Vowel change condition.

Duration and intensity

Regarding salience, the consonants lasted 73.2 ms and were 71.4 dB loud on average, while the vowels lasted 106.9 ms and were 78.9 dB loud. Regarding discriminability within each pair of contrasted phoneme (e.g. the /v/ in Victor vs. the /z/ in Zictor), consonant CPs were on average 9.8 ms shorter (95% *CI* = [−22.2, 2.7]) and 1.1 dB louder (95% *CI* = [−3.0, 5.4]) than consonant MPs; vowel CPs were on average 2.9 ms shorter (95% *CI* = [−7.8, 2.12]) and .5 dB softer (95% *CI* = [−1.8, 0.8]) than vowel MPs. To further explore whether there were differences in salience or discriminability between the consonant and vowel conditions, two separate ANOVAs

examining duration and intensity were run, with the factors of pronunciation (CP versus MP) and condition (Consonant change vs. Vowel change). In both cases, there was only a significant effect of condition with consonants being shorter (duration: $F(1, 51) = 33.29$, $p < 10^{-6}$, $\eta^2_p = .40$) and softer than vowels (intensity: $F(1, 51) = 41.35$, $p < 10^{-6}$, $\eta^2_p = .45$), hence establishing that vowels were more salient than consonants. Moreover, there were no effects of pronunciation nor a pronunciation × condition interaction, suggesting that in both conditions, CPs and MPs could not be discriminated based on duration and intensity differences (therefore, the marginal preference for MPs over CPs found in the Control group in Experiment 2 cannot be attributed to differences in terms of duration or intensity of the first phonemes).

Spectral measures

The normalized acoustic/phonetic distance (*Diff.spectral*, normalized for duration and intensity, based on MFCCs) was used to assess discriminability (as mentioned above, MFCCs are not meaningful with respect to salience). On average, *Diff.spectral* was 1.54 ($SE = .06$) for consonant contrasts, and this was significantly higher than the same index for vowels (1.36, $SE = .03$; $t(52) = 6.45$, $p = .01$, $\eta^2_p = .11$). This establishes that consonant contrasts were acoustically more distinct than vowel contrasts, once normalized for intensity and duration.

Acoustic predicates of preference measures

We then explored the link between individual test infants' performance as attested by the difference in LTs between CPs and MPs (*LT.diff*), and three independent measures of acoustic distance between CPs and MPs: normalized duration and intensity differences (*Diff.duration* and *Diff.intensity*) and *Diff.spectral*. First, a multiple linear regression was run on all 60 test infants (30 Consonant change, 30 Vowel change) with *LT.diff* as the dependent variable and the three acoustic distances as predictors (for which there was no colinearity, all Variance Inflation Factors, *VIF*s < 3). The model did not explain a significant part of the variance in *LT.diff* ($R^2 = .01$, $R^2adjusted = -.04$; $F(3, 56) = .31$, $p = .81$), and none of the predictors significantly predicted the difference in LTs between CPs and MPs. This analysis suggests that the variability in infants' detection of an MP in their own name is not explained by differences in any of the three acoustic dimensions, at least when analyzing the consonant and vowel conditions together. However, it is possible that infants distinguish and process consonants and vowels differently (which could be supported by their differences in salience reported above) and hence that acoustic distances have different effects within each phonemic category. To explore this possibility, we ran the same regression as above but separately for each condition.

In the Consonant change condition (all *VIF*s < 3) this initial three-predictor model did not explain a significant part of the variance in *LT.diff* ($R^2 = .10$, $R^2adjusted = .001$; $F(3, 26) = 1.00$, $p = .41$). For comparison purposes with the Vowel change condition (see below), we then ran a backward regression, which, not surprisingly, did not converge onto any model and in which none of the predictors contributed significantly to *LT.diff*. Therefore, the variance in infants' differential LTs towards CPs versus consonant MPs is not predicted by our three acoustic distances (see Figure 2, left panel, presenting the individual data for *Diff.spectral*, which is further considered in the Vowel change condition).

In the Vowel change condition (all *VIF*s < 3), the regression model with *LT.diff* and the three predictors yielded a marginal model ($F(3, 26) = 2.51$, $p = .08$) explaining 22.5% of the variance ($R^2adjusted = .135$; *standard error of estimate* = 2.80). In this model, *Diff.spectral* was the only significant predictor ($\beta_{spectral}$

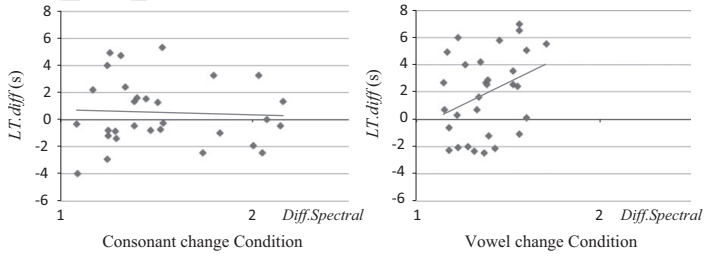


Figure 2 *Link between LT.diff and Diff.spectral in the Consonant change condition (left) and Vowel change condition (right).*

= .394, p = .034). We then ran a backward regression that converged onto a significant two-predictor model (F(2, 27) = 3.44, p = .047) resulting in the exclusion of *Diff.intensity*. This final model explained 20.3% of the variance (R^2*adjusted* = .144, *standard error of estimate* = 2.78) and included *Diff.spectral* and *Diff.duration* as predictors ($\beta_{spectral}$ = .369, p = .042; $\beta_{duration}$ = .294, p = .10). This shows that the variability of *Diff.spectral* and *Diff.duration* taken together accounts for a significant part of the variance in infants' differential LTs, and that *Diff.spectral* is a significant predictor of *LT.diff* even though the distribution of *Diff.spectral* values was narrower in the Vowel change condition than in the Consonant change condition (see Figure 2, right panel). Therefore, contrary to what was found in the Consonant change condition, the variability introduced by two acoustic distances (in particular spectral distance) in vowels partially determines French-learning 5-month-olds' preference for their correctly pronounced name over a one-feature vowel mispronunciation.

Lastly, the same three multiple linear regressions were run with the performance of Control infants as the dependent variable (*LT.diff_ctrl*) and the same three independent measures of acoustic distance between CPs and MPs: normalized duration and intensity differences (*Diff.duration* and *Diff.intensity*) and *Diff.spectral*. None of the three models explained a significant part of the variance in infants' differential LTs nor yielded any significant predictor (both conditions together: R^2 = .03, R^2*adjusted* = −.03; F(3, 56) = .50, p = .69; Consonant condition: R^2 = .08, R^2*adjusted* = −.03; F(3, 26) = .74, p = .54; Vowel condition: R^2 = .04, R^2*adjusted* = −.07; F(3, 26) = .35, p = .79).

Discussion

Previous work with adults and toddlers has shown that consonants are more important than vowels in learning and recognizing words. The present study explored whether there is an early consonant bias in one of the earliest words known by infants: their own name. French-learning 5-month-olds were tested on the impact of either a consonant or a vowel mispronunciation on their listening preference for their name. All mispronunciations involved one-feature changes (based on Dell, 1985) between two native consonants or vowels, which were expected to be discriminated by the participants. In the Consonant change condition, infants failed to show a preference for their correctly pronounced names (CPs) over consonant mispronunciations (MPs). In the Vowel change condition, infants preferred their CP names over vowel MPs, demonstrating sensitivity to vowel MPs. These findings were found independently of the consonantal (place, voicing, manner) and vocalic (place, roundedness, height) features tested. Thus, French-learning infants, tested on their sensitivity to minimal phonetic changes in one of their first words, exhibit a vowel bias at the onset of lexical processing.

Our results do not support the initial bias hypothesis (Nespor *et al*., 2003), which predicted a consonant bias at the onset of lexical processing. On the contrary, they suggest that the consonant bias is learned. This is in line with previous reports showing cross-linguistic differences, in particular the later emergence of the consonant bias in English-learning infants than in French-learning infants (Floccia *et al*., 2014; Nazzi *et al*., 2009). One way to interpret the English–French discrepant data has been to suppose a temporary disappearance of the initial consonant bias in English toddlers (U-shaped trajectory, see Floccia *et al*., 2014, for a discussion). However, the present findings contradict this idea, because even in French, a language in which the consonant bias is consistently found in toddlers (e.g. Havy & Nazzi, 2009; Nazzi, 2005; Nazzi *et al*., 2009), infants did not show this pattern. Indeed, infants did not consider consonant changes in a familiar word as lexically relevant, as clear preference responses were found only for vowel changes.

The current findings support the idea that infants have to learn the differentiated functional roles of consonants and vowels. Together with previous reports of a consonant bias at 12 to 14 months of age in Italian- and French-learning infants (Hochmann *et al*., 2011; Zesiger & Jöhr, 2011), the present findings suggest (at least in these two languages) that the time window for the emergence of a consonant bias is the second semester of life. For this learning to take place, one possibility (the lexical hypothesis, Keidel *et al*., 2007) is that infants need to acquire a sizeable lexicon to discover that consonants are statistically more informative than vowels regarding the identity of the words in the lexicon. That could be done through the computation of consonant versus vowel tiers, or consonant versus vowel phonological neighbors; using this information, infants would learn that consonants are more relevant for word learning and word processing. Another possibility (the acoustic–phonetic hypothesis, Floccia *et al*., 2014) is that the acoustic and phonetic differences between consonants and vowels provide an early cue to infants that these types of speech sounds should be processed differently. While the current study was not designed to separate these two hypotheses, some of the present findings nevertheless contribute to this issue.

Our acoustic analyses confirm that vowels are acoustically more salient than consonants both in duration and intensity (Repp, 1984). Importantly, they also reveal that vowel contrasts are spectrally *less* distinct than consonant contrasts per time unit and controlling for intensity. In addition, infants' preference responses clearly establish that consonants and vowels are processed differently in word recognition by 5 months; infants' recognition of their name was more impaired by vowel MPs than consonant MPs, and only their performance with vowel MPs was modulated by acoustic factors. However, these acoustic cues were not significant factors in the overall regression analysis, suggesting that infants might assign consonants and vowels to abstract categories independently of the acoustic factors considered here. Therefore, consonants and vowels differ on acoustic saliency (vowel advantage) and discriminability (consonant advantage), and are processed differently in a lexical task early in development.

The above observations give support to the acoustic–phonetic hypothesis in principle; however, we are left to explain why 5-month-olds nevertheless give more relevance to vowels than to consonants in the current name recognition task. From an acoustic point of view, while it is likely that 5-month-olds can discriminate (most of) the native consonant and vowel contrasts used in the present study when presented in short speech sequences, one possibility is that 5-month-olds are still better at processing acoustic details in the most salient portions of the signal, which are usually the vocalic parts, in line with well-established findings of better perception of vowels in utero and at birth (Bertoncini *et al*., 1988; Benavides-Varela *et al*., 2012; Granier-Deferre *et al*., 2011). It is also possible that infants' early preference for infant-directed speech, which is characterized by large prosodic modulations mostly carried by vowels, contributes to 5-month-olds' greater reliance on vowels in the current study (Fernald, 1985; Werker & McLeod, 1989). From a phonetic/phonological point of view, another possibility is that 5-month-olds have more advanced knowledge of native phonetic categories for vowels than for consonants (Kuhl *et al*., 1992; Werker & Tees, 1984), to a point where they have started learning native vocalic phonemic categories but not native consonantal phonemic categories (in line with the proposal by PRIMIR that phonemes are learned in a staggered fashion; Werker & Curtin, 2005). Finally, the finding that 5-month-olds only used the vocalic contrasts in this lexically related task (name recognition), while being very likely able to discriminate both consonantal and vocalic contrasts in speech perception tasks, extends to this age previous reports of a possible dissociation between phonetic discrimination and use of phonetic information at the lexical level in toddlerhood as a function of level of acquisition, task requirement, or level of processing (e.g. Nazzi, 2005; Stager & Werker, 1997).

But how will the appropriate consonant bias in lexical processing be discovered in development? As they grow older, infants' temporal resolution in sounds becomes more acute (Werner, Marean, Halpin, Spetner & Gillenwater, 1992), which might allow them to perfect their ability to perform fine-grained phonetic distinctions for shorter speech sounds (i.e. consonants) in words. Acoustic/phonetic distance per time unit was found to be larger in consonants than in vowels, which could lead infants to switch attention from the vowels to the acoustically more reliable consonants. A second scenario, related to phonetic/phonological acquisition, is that the learning of native consonantal phoneme categories around 10–12 months of age (Werker & Tees, 1984) could induce a shift in cognitive resources or attention towards consonants in word processing, in line with the PRIMIR proposal (Werker & Curtin, 2005). While these explanations are compatible with the acoustic–phonetic hypothesis, a third possibility related to the lexical hypothesis is that the consonant bias emerges with the acquisition of a sizeable lexicon, allowing infants to discover that there are more phonological neighbors obtained by a consonant than a vowel change, hence that consonants are more informative at the lexical level. One good test against this latter hypothesis would be to establish that the consonant bias is already present at 8 months, an age at which infants' limited vocabularies have little or no phonological neighbors.

To sum up, these novel results break ground in specifying the relative contribution of consonants and vowels in early word recognition, directly contradicting previous accounts of the developmental origins of the consonant bias in lexical processing. Our findings are the first to report comparative evidence of mispronunciation detection of consonants and vowels in a large set of familiar words as early as 5 months, and to ground these effects in the acoustic properties of the words being presented, which will have to be extended to more kinds of words. Importantly, 5-month-olds were found to make lexical distinctions based on minimal changes for vowels (though not for consonants) long before they have started talking in an intelligible way, supporting the notion that early lexical representations are already quite elaborate. Future studies will need to extend this finding in different ways. First, given evidence of cross-linguistic variation in the expression of the consonant bias in toddlerhood (Floccia *et al*., 2014), the present study will need to be extended to other languages. Second, the current study focused on infants' name, which might

have a special valence and status (see Hall, 2009, for a review) and might thus be processed differently from other kinds of words such as count nouns. Recent evidence suggests that this might not be the case, since French-learning 6-month-olds were also found to be more sensitive to a vocalic than a consonantal change in a word segmentation task in which the target words were unfamiliar monosyllabic count nouns (Nishibayashi & Nazzi, 2014). Future studies will have to further explore the generality of this early vocalic bias and its link to the emerging consonant bias across languages.

Author contributions

CB, TN and CF designed the research, CB carried out the research, TF and MAD provided the acoustic analyses, CB and TN analyzed data, and CB, TN and CF wrote the paper.

Acknowledgements

This work was funded by a joint ANR-ESRC grant to TN (ANR-09-FRBR-015) and CF (RES-062-33-0001), by LABEX EFL (ANR-10-LABX-0083) to TN, MAD and CB (partial dissertation funding). Thanks to Lionel Granjon, Viviane Huet, Henny Yeung, Judit Gervain and Claire Delle Luche, and all the infants and parents who participated.

References

Acha, J., & Perea, M. (2010). On the role of consonants and vowels in visual-word processing: evidence with a letter search paradigm. *Language and Cognitive Processes*, **25**, 423–438.

Barras, C., Geoffroy, E., Wu, Z., & Libberman, M. (2001). Transcriber: development and use of a tool for assisting speech corpora production. *Speech Communication*, **33**, 5–22.

Benavides-Varela, S., Hochmann, J.-R., Macagno, F., Nespor, M., & Mehler, J. (2012). Newborn's brain activity signals the origin of word memories. *Proceedings of the National Academy of Sciences, USA*, **109**, 17908–17913.

Bergelson, E., & Swingley, D. (2012). At 6–9 months, human infants know the meanings of many common nouns. *Proceedings of the National Academy of Sciences, USA*, **109**, 3253–3258.

Bertoncini, J., Bijeljac-Babic, R., Jusczyk, P.W., & Kennedy, L.J. (1988). An investigation of young infants' perceptual representations of speech sounds. *Journal of Experimental Psychology: General*, **117**, 21–33.

Bonatti, L.L., Penã, M., Nespor, M., & Mehler, J. (2007). On consonants, vowels, chickens, and eggs. *Psychological Science*, **18**, 924–925.

Bortfeld, H., Morgan, J., Golinkoff, R., & Rathbun, K. (2005). Mommy and me: familiar names help launch babies into speech stream segmentation. *Psychological Science*, **16**, 298–304.

Caramazza, A., Chialant, D., Capasso, D., & Miceli, G. (2000). Separable processing of consonants and vowels. *Nature*, **403**, 428–430.

Carreiras, M., & Price, C.J. (2008). Brain activation for consonants and vowels. *Cerebral Cortex*, **18**, 1727–1735.

Cutler, A., Sebastián-Gallés, N., Soler-Vilageliu, O., & Van Ooijen, B. (2000). Constraints of vowels and consonants on lexical selection: cross-linguistic comparisons. *Memory & Cognition*, **28**, 746–755.

Davis, S.B., & Mermelstein, P. (1980). Comparison of parametric representations for monosyllabic word recognition in continuously spoken sentences. *IEEE Transactions on Acoustics, Speech and Signal Processing*, **28** (4), 357–366.

Dell, F. (1985). *Les règles et les sons (2nd* (revised edn.). Paris: Hermann.

Delle Luche, C., Poltrock, S., Goslin, J., New, B., Floccia, C., & Nazzi, T. (2014). Differential processing of consonants and vowels in the auditory modality: a cross- linguistic study. *Journal of Memory and Language*, **72**, 1–15.

Fernald, A. (1985). Four-month-old infants prefer to listen to motherese. *Infant Behavior and Development*, **8**, 181–195.

Floccia, C., Nazzi, T., Delle Luche, C., Poltrock, S., & Goslin, J. (2014). English-learning one- to two-year-olds do not show a consonant bias in word learning. *Journal of Child Language*, **41** (5), 1085–1114.

Fry, D.B., Abramson, A.S., Eimas, P.D., & Liberman, A.M. (1962). The identification and discrimination of synthetic vowels. *Language and Speech*, **5**, 171–188.

Granier-Deferre, C., Ribeiro, A., Jacquet, A.Y., & Bassereau, S. (2011). Near-term fetuses process temporal features of speech. *Developmental Science*, **14**, 336–352.

Hall, D.G. (2009). Proper names in early word learning: rethinking a theoretical account of lexical development. *Mind & Language*, **24**, 404–432.

Havy, M., Bertoncini, J., & Nazzi, T. (2011). Word learning and phonetic processing in preschool age children. *Journal of Experimental Child Psychology*, **108**, 25–43.

Havy, M., & Nazzi, T. (2009). Better processing of consonantal over vocalic information in word learning at 16 months of age. *Infancy*, **14**, 439–456.

Havy, M., Serres, J., & Nazzi, T. (2014). A consonant/vowel asymmetry in word-form processing: eye-tracking evidence in childhood and in adulthood. *Language and Speech*, **57**, 254–281.

Hochmann, J.-R., Benavides-Varela, S., Nespor, M., & Mehler, J. (2011). Consonants and vowels: different roles in early language acquisition. *Developmental Science*, **14**, 1445–1458.

Keidel, J.L., Jenison, R.L., Kluender, K.R., & Seidenberg, M.S. (2007). Does grammar constrain statistical learning? Com-

mentary on Bonatti, Peña, Nespor, and Mehler (2005). *Psychological Science*, **18**, 922–923.
Kuhl, P.K., Conboy, B.T., Coffey-Corina, S., Padden, D., Rivera-Gaxiola, M., & Nelson, T. (2008). Phonetic learning as a pathway to language: new data and native language magnet theory expanded (NLM-e). *Philosophical Transactions of the Royal Society B: Biological Sciences*, **363**, 979–1000.
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.
Ladefoged, P. (1993). *A course in phonetics* (3rd edn.). New York: Harcourt Brace Jovanovich.
Maddieson, I. (1984). *Patterns of sounds*. Cambridge: Cambridge University Press.
Mandel, D.R., Jusczyk, P.W., & Pisoni, D.B. (1995). Infants' recognition of the sound patterns of their own names. *Psychological Science*, **6**, 314–317.
Mani, N., & Plunkett, K. (2007). Phonological specificity of vowels and consonants in early lexical representations. *Journal of Memory and Language*, **57**, 252–272.
Mani, N., & Plunkett, K. (2010). Twelve-month-olds know their cups from their keps and tups. *Infancy*, **15**, 445–470.
Mersad, K., & Nazzi, T. (2012). When mommy comes to the rescue of statistics: infants combine top-down and bottom-up cues to segment speech. *Language Learning and Development*, **8**, 303–315.
Nazzi, T. (2005). Use of phonetic specificity during the acquisition of new words: differences between consonants and vowels. *Cognition*, **98**, 13–30.
Nazzi, T., Floccia, C., Moquet, B., & Butler, J. (2009). Bias for consonantal over vocalic information in French- and English-learning thirty-month-olds: crosslinguistic evidence in early word learning. *Journal of Experimental Child Psychology*, **102**, 522–537.
Nespor, M., Peña, M., & Mehler, J. (2003). On the different roles of vowels and consonants in speech processing and language acquisition. *Lingue e Linguaggio*, **2**, 203–229.
New, B., Araùjo, V., & Nazzi, T. (2008). Differential processing of consonants and vowels in lexical access through reading. *Psychological Science*, **19**, 1223–1227.
Nishibayashi, L., & Nazzi, T. (2014). Asymmetry in consonant/vowel processing in recognizing segmented word forms: evidence of a switch between 6 and 8 months of age. Poster presented at the 19th Biennal Meeting of the International Society for Infant Studies, Berlin, Germany, 3–5 July.
Pons, F., & Toro, J.M. (2010). Structural generalizations over consonants and vowels in 11-month-old infants. *Cognition*, **116**, 361–367.
Repp, B.H. (1984). Categorical perception: issues, methods, findings. In N.J. Lass (Ed.), *Speech and language: Advances in basic research and practice*, Vol. **10** (pp. 243–335). New York: Academic Press.
Sakoe, H., & Chiba, S. (1978). Dynamic programming algorithm optimization for spoken word recognition. *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, **26**, 33–49.
Stager, C.L., & Werker, J.F. (1997). Infants listen for more phonetic detail in speech perception than in word-learning tasks. *Nature*, **388**, 381–382.
Tincoff, R., & Jusczyk, P.W. (1999). Some beginnings of word comprehension in 6-month-olds. *Psychological Science*, **10**, 172–175.
Tincoff, R., & Jusczyk, P.W. (2012). Six-month-olds comprehend words that refer to parts of the body. *Infancy*, **17**, 432–444.
Toro, J.M., Nespor, M., Mehler, J., & Bonatti, L.L. (2008). Finding words and rules in a speech stream: functional differences between vowels and consonants. *Psychological Science*, **19**, 137–144.
Werker, J.F. (1994). Cross-language speech perception: development change does not involve loss. In J.C. Goodman & H.C. Nusbaum (Eds.), *The development of speech perception: The transition from speech sounds to spoken words* (pp. 93–120). Cambridge, MA: MIT Press.
Werker, J.F., & Curtin, S. (2005). PRIMIR: a developmental model of speech processing. *Language Learning and Development*, **1**, 197–234.
Werker, J.F., & McLeod, P.J. (1989). Infant preference for both male and female infant-directed talk: a developmental study of attentional and affective responsiveness. *Canadian Journal of Psychology*, **43**, 230–246.
Werker, J.F., & Tees, R.C. (1984). Cross-language speech perception: evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development*, **7**, 49–63.
Werner, L.A., Marean, G.C., Halpin, C.F., Spetner, N.B., & Gillenwater, J.M. (1992). Infant auditory temporal acuity: gap detection. *Child Development*, **63**, 260–272.
Zesiger, P., & Jöhr, J. (2011). Les représentations phonologiques des mots chez le jeune enfant. *Enfance*, **3**, 293–309.

Received: 25 April 2014
Accepted: 28 July 2014