The Developmental Origins of the Consonant Bias in Lexical Processing

Thierry Nazzi1, Silvana Poltrock1,2, and Katie Von Holzen1
1Laboratoire Psychologie de la Perception, Centre National de la Recherche Scientifique (CNRS) and Université Paris Descartes, and 2Faculty of Cognitive Sciences, University of Potsdam

Abstract
Consonants have been proposed to carry more of the weight of lexical processing than vowels. This consonant bias has consistently been found in adults and has been proposed to facilitate early language acquisition. We explore the origins of this bias over the course of development and in infants learning different languages. Although the consonant bias was originally thought to be present at birth, evidence suggests that it arises from the early stages of phonological and (pre-)lexical acquisition. We discuss two theories that account for the acquisition of the consonant bias: the lexical and acoustic-phonetic hypotheses.

Keywords
language acquisition, lexical development, consonant bias, cross-linguistic differences

Imagine reading this manuscript if all of the consonants were removed: Comprehension would prove a challenge. In contrast, if all of the vowels were removed, comprehension would still proceed relatively smoothly. This example illustrates the importance of consonants in lexical processing, which was put forth by Nespor, Peña, and Mehler (2003). Nespor and colleagues proposed that consonants carry more information about the lexicon, whereas vowels play a more important role in syntactic and prosodic processing. For example, when auditorily transforming a nonword into a real word, English, Dutch, and Spanish listeners preserve consonantal over vocalic information, changing kebra into cobra rather than zebra (Cutler, Sebastián-Gallés, Soler-Vilageliu, & van Ooijen, 2000; van Ooijen, 1996). This bias for consonantal information in lexical processing (hereafter, C-bias) has been supported by studies with adults across several languages (Dutch, English, French, Italian, Portuguese, Spanish) and a variety of different methods such as word learning (Bonatti, Peña, Nespor, & Mehler, 2005; Creel, Aslin, & Tanenhaus, 2006; Escudero, Mulak, & Vlach, 2015; Havy, Serres, & Nazzi, 2014; Mehler, Peña, Nespor, & Bonatti, 2006) and lexical access (Carreiras, Dunabeitia, & Molinaro, 2009; Delle Luche et al., 2014; Lee, Rayner, & Pollatsek, 2002; New, Araújo, & Nazzi, 2008; New & Nazzi, 2014; Scares, Perea, & Comesana, 2014).

Provided that this pattern generalizes to further languages (an issue for future research), consonants appear to hold a privileged status in adult lexical processing and may even be a feature of a mature, linguistic processing system. Although consonants and vowels have been found to be used differently in many domains of auditory processing (Kolinsky, Lidji, Peretz, Besson, & Morais, 2009; Owren & Cardillo, 2006), Nespor and colleagues (2003) proposed that for language processing, consonants and vowels have specialized biases: a consonant bias for lexical processing and a vowel bias for syntactic/prosodic processing. The question is, then, whether we are born with these biases or whether they must be learned over time. This is a nontrivial question, since knowing these biases early may facilitate language acquisition. In this article, we focus on the C-bias in lexical processing and explore two related questions: What is the developmental trajectory of the C-bias, and what are the mechanisms driving it?

Lexical Processing Biases in Early Development

Nazzi (2005) was the first study to directly examine a bias for consonants or vowels in lexical processing in young...
infants. In that study, 20-month-old French-learning infants were presented with triads of novel objects, comprising two objects given the same name and a third object given a name that differed by one phonetic feature, and were asked to group together the two objects with the same name. Infants successfully learned the object labels when they differed by one of their consonants (e.g., /pize/ vs. /tize/) but not when they differed by one of their vowels (/pize/ vs. /pyze/). This study was therefore the first to extend the C-bias found in adults to infants as young as 20 months.

Subsequent studies with French-learning infants strengthened the argument that the C-bias is a feature of lexical processing in early development (Havy & Nazzi, 2009; Zesiger & Jöhr, 2011) and can be generalized across different consonants (e.g., /nuk/ vs. /muk/, /rize/ vs. /lize/; Nazzi & New, 2007) and positions within a word (Nazzi & Bertoncini, 2009). The C-bias is even in place by the first birthday in French-learning infants. In a test of word-form recognition, Poltrock and Nazzi (2015) confirmed that 11-month-olds prefer to listen to familiar words compared to nonwords and then showed that when infants are presented with vowel and consonant mispronunciations, they prefer to listen to the vowel mispronunciations. This suggests that changing consonant information in a word interferes with its auditory recognition to a greater extent than changing vowel information. Although French-learning toddlers do eventually use vocalic information when learning new words, older children and adults continue to give more weight to consonantal than to vocalic information (Havy et al., 2014; Nazzi, Floccia, Moquet, & Butler, 2009).

While French-learning infants consistently demonstrated a C-bias by their first birthday, support for an innate C-bias became less clear as infants at even younger ages and from other language backgrounds were tested (see Fig. 1). Regarding cross-linguistic evidence, English-learning infants were compared on their use of consonant and vowel information when learning words and did not show a C-bias until 30 months (Floccia, Nazzi, Luche, Poltrock, & Goslin, 2014; Nazzi et al., 2009). One experiment found evidence that English-learning infants are more sensitive to a consonant than a vowel change at 15 months, but similar levels of sensitivity were found at 12, 18, and 24 months (Mani & Plunkett, 2007, 2010). Although these studies on English-learning infants have revealed symmetry for consonants and vowels during lexical processing at younger ages, a C-bias emerges by at least 30 months, continuing on into adulthood (Delle Luche et al., 2014).

Results from studies on Danish-learning infants have provided further evidence that the C-bias differs across languages in infancy. Danish provides an interesting linguistic case, as it is a language in which there are more vowels than consonants (contrary to most other languages) and in which consonants are often underarticulated (i.e., extensive consonant reduction), which increases the relative salience of vowels—both phenomena that should disfavor acquiring a C-bias. When learning new words, Danish-learning 20-month-olds were found to rely more on vowels than consonants, providing the first evidence of a vowel bias in lexical processing (V-bias; Højen & Nazzi, 2016). Future studies will be needed to investigate whether the V-bias in Danish-learning 20-month-olds extends to other ages and to Danish-speaking adults, or whether Danish-speaking adults show a C-bias in lexical processing similar to adults of other language backgrounds tested so far (e.g., Cutler et al., 2000; New et al., 2008).

These studies demonstrated cross-linguistic differences in the biases infants ascribe to consonant and vowel information during lexical processing, showing a rather late emergence of the C-bias in English and a V-bias in Danish. While cross-linguistic differences in the C-bias suggest that it is not language independent, there is still the possibility that all infants are born with a C-bias that is later modulated by language exposure. Alternatively, the C-bias might develop, if appropriate, as infants acquire the properties of their native language. In a study investigating brain activity using near-infrared spectroscopy (NIRS), Italian newborns were found to have better memory for vowel over consonant information in words they had been presented with several minutes prior (Benavides-Varela, Hochmann, Macagn, Nesp, & Mehler, 2012). This placed a V-bias in lexical processing at the onset of development. By their first birthday, however, this initial V-bias changed to a C-bias for Italian-learning infants; Italian-learning 12-month-olds relied on consonant rather than vowel information to anticipate the appearance of a toy (Hochmann, Benavides-Varela, Nesp, & Mehler, 2011). Therefore, during the first year of life, Italian-learning infants appear to develop a C-bias.

Studies investigating young French-learning infants also suggest that the C-bias may develop during the infant’s first year. Bouchon and colleagues (Bouchon, Floccia, Fux, Adda-Decker, & Nazzi, 2015) presented 5-month-olds with correct and mispronounced versions of their name. Vowel, but not consonant, mispronunciations impaired name recognition, demonstrating a V-bias. Furthermore, Nishibayashi and Nazzi (2016) directly investigated the age at which the C-bias emerges using a segmentation task exploring infants’ ability to extract word forms from fluent speech. French-learning 6- and 8-month-olds were first familiarized with two short stories in which two target words were repeated and then tested on their preference for consonant versus vowel mispronunciations of the target words, with the assumption that infants should listen longer to mispronunciations that keep more important segmental information intact. At 8 months, infants preferred to listen to vowel-mispronunciations...
Developmental Origins of the Consonant Bias

The C-bias for lexical processing therefore appears between 6 and 8 months even in French.

Mechanisms Underlying the Development of the C-Bias

Taken together, studies investigating the C-bias in lexical processing across different languages and throughout development have supported the notion that the C-bias is learned, possibly from an initial V-bias. The question remains, however, as to what mechanisms are driving the acquisition of the C-bias; we discuss below two hypotheses proposed to account for its acquisition: the lexical hypothesis and the acoustic-phonetic hypothesis.

The lexical hypothesis was proposed in response to a series of experiments by Bonatti and colleagues, who found that French adults use statistical information carried by consonants, but not vowels, to segment words from a continuous speech stream in an artificial grammar (Bonatti et al., 2005). Keidel and colleagues (Keidel, Jenison, Kluender, & Seidenberg, 2007) argued that the C-bias is directly related to adult listeners’ implicit and acquired knowledge about the structure of the lexicon of their native language. Analyses of the French and English lexicons showed that words in both languages are better distinguished by consonants than vowels. For example, in French, knowing the identity of the consonant sequence of a CVCVCV word (e.g., C-S-N in casino) restricts the average number of possible candidates to 6.03 words, while knowing the vowel sequence (A-I-O) leaves an average of 8.8 possible candidates—that is, 1.46 times as many (Delle Luche et al., 2014; Keidel et al., 2007). This acquired knowledge would result in preferential processing of consonant information during lexically related tasks in these languages.

The lexical hypothesis was originally proposed to account for an adult C-bias and therefore does not specify whether infants must have meaning attached to the words in their lexicon or if a C-bias can develop based on word forms memorized in a pre-lexicon (see Poltrock & Nazzi, 2015). This distinction between lexical and pre-lexical acquisition is important, as it has implications regarding the point in development when the C-bias develops. At present, we know little regarding these issues. For example, the C-bias found by Poltrock and Nazzi (2015) at 11 months, an age when infants have already arguably developed a lexicon with links between word forms and concepts, was not found to be related to lexicon size (but see Mani & Plunkett, 2010, for a relationship between vocabulary size and sensitivity to vowel mispronunciations). Nishibayashi and Nazzi (2016) found a C-bias by 8 months, an age at which infants may have an emerging lexicon too small to elicit a C-bias (even if at 6 months infants can already recognize the meaning of a handful of words; Bergelson & Swingley, 2012; Tincoff &
Jusczyk, 1999, 2012), but might have a large enough pre-lexicalized to allow such emergence. A better understanding of the relationship between the C-bias and lexical versus pre-lexical development is needed to address this issue. Therefore, to directly evaluate the validity of the lexical hypothesis, experiments must examine the biases of infants at the cusp of developing a C-bias (8 months or younger) and relate these biases to the number of words and the structure of the lexicon that infants know.

In response to Keidel et al. (2007), Bonatti and colleagues (Bonatti, Peña, Nespor, & Mehler, 2007) argued that the C-bias cannot be driven by distributional information alone. This is further articulated by the acoustic-phonetic hypothesis (Floccia et al., 2014), which relates the emergence of the C-bias to the acquisition of the acoustic-phonetic properties of the native language during the first year of life. Initially, infants would rely on salient acoustic properties and thus preferentially process vowels, which are longer, more periodic, and steadier than consonants (Benavides-Varela et al., 2012; Bouchon et al., 2015). Already at birth, however, consonant contrasts are perceived more categorically than vowel contrasts, and as development proceeds, this would help signal to infants that consonants are more reliable and faster processing cues to word identity (Hochmann et al., 2011). In addition, this switch may be related to increases in fine temporal resolution of infants’ auditory system or the acquisition of the native phonemic categories (see Poltrock & Nazzi, 2015). As a result, infants would develop phonologically distinct categories for consonants and vowels, which would each act as preferential cues for the processing of different kinds of information (consonants for the lexicon; vowels for prosody/syntax). In contrast to the lexical hypothesis, the acoustic-phonetic hypothesis is related not to lexical size or structure but instead to phonological development and sensitivity to acoustic-phonological characteristics. For example, sensitivity to vowel mispronunciations was modulated by acoustic factors for French-learning 5-month-olds (Bouchon et al., 2015), suggesting that acoustic information may be driving the V-bias at this age. To better evaluate the acoustic-phonetic hypothesis, studies investigating the C-bias throughout development will need to relate similar acoustic measures with infant biases for vowels and consonants.

Both the lexical and the acoustic-phonetic hypothesis present partial explanations for cross-linguistic differences observed in toddlers. According to the lexical hypothesis, the V-bias in Danish (Højen & Nazzi, 2016) could be due to the fact that Danish (unlike French or English) has more vowels than consonants, which might result in vowels being more informative than consonants at the lexical level. In contrast, from an acoustic-phonetic perspective, extensive consonant reduction in Danish could make consonants harder to process and identify than in other languages. As for the late emergence of the C-bias in English, vowels might play a greater lexical role in English than in French as a result of several phenomena in English, such as vocalic reduction and contrastive lexical stress (e.g., *insight* vs. *incite*) and the tense/lax vowel distinction (e.g., *bin* vs. *bean*). In future studies, acoustic-phonological and lexical analyses will be essential for understanding the mechanisms that modulate the consonant bias cross-linguistically.

Indeed, it is possible that the lexical and the acoustic-phonetic hypotheses are not mutually exclusive. For the lexical role of consonants to emerge, consonants must first be perceived as a phonological category (Hochmann et al., 2011). A bias for consonant information based on the distribution of consonants and vowels in an infant’s first words and speech input is necessarily precluded by the knowledge that consonants and vowels are different, which is the result of their acoustic-phonetic properties. In addition, clarification is needed for whether the lexical hypothesis operates on a lexicon with word-concept links or whether the C-bias can develop before infants have access to word meaning. To further evaluate their relative contributions, future work will need to examine both proposals within the same experimental design, in different languages, linking measures of both lexical size and acoustic salience and variability to infants’ use of consonant and vowel information in lexically related tasks. These studies will also bring new data regarding the links between phonological and lexical acquisitions, a topic that has gained interest in the past years (e.g., Yeung, Chen, & Werker, 2014; Yeung & Nazzi, 2014).

**Recommended Reading**

Floccia, C., Nazzi, T., Delle Luche, C., & Poltrock, S. (2014). (See References). Provides a full discussion of the acoustic-phonetic hypothesis for readers who wish to learn more about theories for the mechanisms supporting the development of the consonant bias.

Keidel, J. L., Jenison, R. L., Kluender, K. R., & Seidenberg, M. S. (2007). (See References). Provides a full discussion of the lexical hypothesis for readers who wish to learn more about theories for the mechanisms supporting the development of the consonant bias.

Nazzi, T. (2005). (See References). The first article to examine the existence of a consonant bias during language acquisition.


**Declaration of Conflicting Interests**

The authors declared that they had no conflicts of interest with respect to their authorship or the publication of this article.
Funding
This study was supported by grants from the Agence Nationale de la Recherche (ANR-09-FRBR-015 and ANR-13-BSH2-0004) and from the Labex “Empirical Foundations of Linguistics” (ANR-10-LABX-0083) to T. Nazzi.

References


