Developmental Science

Developmental Science 14:6 (2011), pp 1445–1458 DOI: 10.1111/j.1467-7687.2011.01089.x

PAPER

Consonants and vowels: different roles in early language acquisition

Jean-Rémy Hochmann,[1] Silvia Benavides-Varela,[1] Marina Nespor[2] and Jacques Mehler[1]

1. *Cognitive Neuroscience Department, SISSA, International School for Advanced Studies, Trieste, Italy*
2. *Facoltà di Psicologia, University of Milano Bicocca, Italy*

Abstract

Language acquisition involves both acquiring a set of words (i.e. the lexicon) and learning the rules that combine them to form sentences (i.e. syntax). Here, we show that consonants are mainly involved in word processing, whereas vowels are favored for extracting and generalizing structural relations. We demonstrate that such a division of labor between consonants and vowels plays a role in language acquisition. In two very similar experimental paradigms, we show that 12-month-old infants rely more on the consonantal tier when identifying words (Experiment 1), but are better at extracting and generalizing repetition-based srtuctures over the vocalic tier (Experiment 2). These results indicate that infants are able to exploit the functional differences between consonants and vowels at an age when they start acquiring the lexicon, and suggest that basic speech categories are assigned to different learning mechanisms that sustain early language acquisition.

Introduction

To acquire language, infants need both to learn words and to extract and generalize structural regularities that play a role in learning grammar, e.g. syntax. For instance, when hearing the utterance 'the girl kicks the boy', we need more than just the meaning of each word to understand the whole sentence. The listener has to understand the relation between the verb and both the subject and the object. Learning words requires memorizing specific elements of the input (e.g. *girl*, *boy*) and representing them in a format that allows their recognition and distinction from other words (e.g. *boy* vs. *toy*), whereas learning regularities about syntactic structures implies the ability to extract relations between elements of the input (e.g. whether the verb precedes or follows the object) and generalize them to new sentences. Young language learners could thus profit from a (partial) 'division of labor', such that one speech category might preferentially support the acquisition of the lexicon, whereas another might be more dedicated to the identification of structural regularities, in particular those signaling relations between constituents. Here, we will present a series of observations that hint at different roles for consonants and vowels in language acquisition. Namely, we propose that consonants are more involved with word identification and encoding, because they are better suited than vowels for categorical perception. Vowels, in contrast, carry prosodic variations and provide cues to determine the boundaries and the organization of syntactic constituents (Nespor & Vogel, 1986; Selkirk, 1984). This functional difference between consonants and vowels constitutes the Consonant-Vowel hypothesis, hereafter referred to as the CV hypothesis (Nespor, Peña & Mehler, 2003).

The CV hypothesis

Part I: The consonantal bias in lexical processes

The role of consonants in the CV hypothesis originates from the observation that across languages, consonants allow more quality distinctions than vowels, where by quality we mean distinctions in terms of articulatory features.[1] With only a few exceptions, consonants are indeed cross-linguistically more numerous than vowels. For example, in Malay the proportion is 20C: 5V; in Italian 24C: 7V; in Hausa 32C: 5V; in Arabic 29C: 3V; in Igbo 27C: 5V; in Sindhi 46C: 10V (see Maddieson, 2008).

[1] There are systems where consonants may be distinguished for duration (e.g. Italian, *caro–carro;* [káro]–[kár:o]; 'dear' – 'cart') and systems where vowels may be distinguished because of suprasegmentals, such as nasality or tones (e.g. French, *beau–bon*; [bo]–[bɔ̃]; 'beautiful' – 'good'). Also in these types of systems, by and large, consonant distinctions outnumber vowel distinctions.

Address for correspondence: Jean-Rémy Hochmann, SISSA, International School for Advanced Studies, Cognitive Neuroscience Department, Via Bonomea, 265, 34136, Trieste, Italy; e-mail: jr.hochmann@gmail.com

Most importantly, across languages, five vowel systems are the most common and most systems have over 20 consonants (Nespor *et al.*, 2003). Because consonants are more numerous than vowels, they are relatively more informative for lexical distinctions.

A second observation motivating our belief that consonants have a predominant role to play in lexical learning is that, in Semitic languages such as Arabic and Hebrew, lexical roots are only represented on the consonantal tier (Berent, Vaknin & Marcus, 2007; McCarthy, 1985; Prunet, Beland & Adrissi, 2000). For example, in Arabic, the root *ktb* has the lexical meaning related to 'write'. The different vowels intervening between the consonants serve to form different words and word-forms that are related in meaning (e.g. *katib*: writer, *kataba*: he wrote, *kitab*: book, *maktaba*: library, etc.). In contrast, there is no documented language that has lexical roots based only on vowels. Furthermore, the case of consonantal lexical roots may be an extreme case of the situation observed in other languages, where consonants are in general more informative than vowels for lexical distinctions (see Keidel, Jenison, Kluender & Seidenberg, 2007, for the analysis of the French adult lexicon). In the Supplementary Material, we analyze the lexicon of infants for French and Italian, two languages that differ in the ratio of consonants and vowels (French has 19 consonants and 13 vowels; Italian has 24 consonants and seven vowels). For both languages, we verify that the sequence of consonants is more informative than the sequence of vowels for word identification. This suggests that the lexical role of consonants does not change as a function of the ratio of consonants and vowels.

Part II: The role of vowels in signaling syntactic organization

The CV hypothesis attributes a specific role to vowels for the acquisition of syntax. This hypothesis is based on three observations, which we develop below. First, vowels are the main carriers of prosodic information, including rhythm (Lehiste, 1970; Ramus, Nespor & Mehler, 1999). Second, pre-lexical infants and neonates are sensitive to rhythm and to prosodic phrase boundaries (Christophe, Mehler & Sebastián-Gallés, 2001; Christophe, Nespor, Dupoux, Guasti & van Ooyen, 2003; Nazzi, Bertoncini & Mehler, 1998; Ramus, Hauser, Miller, Morris & Mehler, 2000). Third, prosodic and rhythmic information provides cues that correlate with important morphosyntactic properties (Morgan & Demuth, 1996; Nespor, Shukla & Mehler, 2011; Nespor & Vogel, 1986; Selkirk, 1984). Syntax acquisition may thus start with inferences from prosodic cues carried by vowels.

In fact, vowels more than consonants can vary in terms of pitch, intensity and duration in relation to their sentential position. That is, vowels are more affected than consonants by prosody, which provides signals to syntactic constituency (Gleitman & Wanner, 1982; Morgan & Demuth, 1996; Nespor & Vogel, 2008; Selkirk, 1984). In particular, a phonological phrase boundary always coincides with a syntactic boundary. These boundaries are available to infants (Kemler Nelson, Hirsh-Pasek, Jusczyk & Wright-Cassidy, 1989; Gerken, Jusczyk & Mandel, 1994) as well as newborns (Christophe *et al.*, 2001).

Moreover, certain syntactic properties correlate with prosodic cues across languages. For example, complements are marked mainly by pitch and intensity prominence in head final languages such as Turkish, and mainly by duration prominence in head initial languages such as French (Nespor, Shukla, van de Vijver, Avesani, Schraudolf & Donati, 2008). Those prosodic cues remain valid within languages that allow both types of constructions. For example, in German, the complement-head structures are mainly marked by an initial pitch and intensity prominence, while the head-complement structures are mainly marked by a final duration prominence (Nespor *et al.*, 2008; Shukla & Nespor, 2010). Observing this and other correlations, prosodic bootstrapping theories (Morgan & Demuth, 1996) of language acquisition have proposed the existence of a bridge between prosodic and syntactic properties.

Finally, vowels carry a second type of rhythmic information. The proportion of time occupied by vowels in the speech input determines the rhythmic class of languages. Ramus *et al.* (1999) showed that vowels occupy about 45% of the speech stream in stress-timed languages (e.g. Dutch, English), about 50% in syllable-timed languages (e.g. French, Italian) and 55% in mora-timed languages (e.g. Japanese). Newborns can use this information to discriminate between two languages that belong to different rhythmic classes (Ramus *et al.*, 2000). The rhythmic class to which a language belongs correlates with important morphosyntactic properties (Nazzi *et al.*, 1998; Nespor *et al.*, 2011; Ramus *et al.*, 1999). In particular, the percentage of the speech stream that vowels occupy is indicative of the complexity of the syllabic repertoire of a given language. In turn, syllabic complexity correlates with the length of common words (Mehler & Nespor, 2004). In addition, typological studies have shown that languages with simple syllabic structures tend to be verb final, to use post-positions and have a rich case system (Donegan & Stampe, 1983; Gil, 1986; Fenk-Oczlon & Fenk, 2005; Nespor *et al.*, 2011). Infants may therefore infer certain morphosyntactic properties of their language from the identification of its rhythmic class.

In sum, the prosodic – in particular rhythmic – properties of vowels provide infants with information about the syllabic repertoire, signal important syntactic boundaries, and provide cues to fundamental syntactic properties such as the relative order of heads and complements. These observations inspired the second part of the CV hypothesis: infants may focus on vowels when extracting structural information from their input, especially when learning about the relation between different sentence constituents.

Experimental evidence in favor of the CV hypothesis

A number of experimental results support the CV hypothesis. Infants are able to use statistical information, such as dips in transition probabilities (TPs) between syllables to identify word boundaries in a continuous speech stream (Saffran, Aslin & Newport, 1996). Newport and Aslin (2004) argued that adults could also compute TPs both between successive consonants and between successive vowels. However, assuming that participants use TPs to identify potential words, the CV hypothesis predicts that they should perform better at computing TPs over consonants than over vowels. In fact, in their experiments, Newport and Aslin (2004) allowed adjacent repetitions of the consonantal or vocalic frames to segment. Bonatti, Peña, Nespor and Mehler (2005) showed that, avoiding the adjacent repetition of both kinds of frames, adult participants are able to segment a continuous speech stream using the dependencies between consonants, but they fail to segment the stream using the dependencies between vowels. This suggests that participants can compute TPs over consonants but not over vowels. Moreover, Mehler, Peña, Nespor and Bonatti (2006) showed that when, within a single stream, TPs between consonants and TPs between vowels predict different segmentations, the statistics over consonants are favored. Thus, consonants appear to be a privileged category for discovering words in a continuous speech stream.

Word learning experiments in infants and toddlers further confirmed the advantage of consonants in encoding lexical items. Nazzi and colleagues (Nazzi, Floccia, Moquet & Butler, 2009) showed that in a word learning situation where 30-month-olds must ignore either a consonantal one-feature change or a vocalic one-feature change (e.g. match a /duk/ with either a /guk/ or a /dɔ̃k/), both French- and English-learning infants choose to neglect the vocalic change rather than the consonantal change. This preference was observed for word-initial (/guk/-/duk/-/dɔ̃k/), word-final (/pib/-/pid/-/pɛd/) and word-internal consonants (/gito/-/gipo/-/gupo/), and did not depend on an inability to process fine vocalic information. In agreement with these results, 16- to 20-month-old infants could acquire simultaneously two words differing only in one consonant, whereas they could not do so for minimal pairs differing in one vowel (Havy & Nazzi, 2009; Nazzi, 2005; Nazzi & Bertoncini, 2009). Furthermore, these findings are not restricted to one specific consonantal class (Nazzi *et al*., 2009; Nazzi & New, 2007). Similar results were also obtained with 3-year-old French children (Havy, Bertoncini & Nazzi, 2011). Finally, even 4- and 5-year-old children, who are able to process fine phonetic information for both vowels and consonants, still show a consonantal bias for word learning (Havy, 2009).

The second prediction of the CV hypothesis is that infants should focus on vowels when extracting structural information. Acquiring the syntax of natural languages requires abstraction abilities. Chomsky (1957) showed that linguistic productions rely on structures that cannot be described by statistical dependencies between words (modeled as Markov chains or finite-state automatons). Rather, the description of syntax must contain hierarchical structures defined over abstract categories.

To show generalization abilities in infancy, researchers must provide evidence that infants transfer a response learned for a set of stimuli to a new set of stimuli. The representation that has been generalized can be inferred by looking at what is common and what differs between the two sets. From the perspective of the acquisition of structural syntactic information, we want to test infants' ability to generalize an abstract structure that cannot be reduced to statistical dependencies.

Recently, researchers have tackled this matter by using structures defined by the relative positions of identical elements (e.g. ABB, ABA). Indeed, since the seminal work of Marcus and colleagues (Marcus, Vijayan, Bandi Rao & Vishton, 1999) repetition-based structures have been extensively used to test abstract generalization abilities in infants and adults (see Endress, Dehaene-Lambertz & Mehler, 2007; Endress, Scholl & Mehler, 2005; Johnson, Fernandes, Frank, Kirkham, Marcus, Rabagliati & Slemmer, 2009; Kovács & Mehler, 2009b; Saffran, Pollak, Seibel & Shkolnik, 2007). Marcus and colleagues used the following structures: ABB (instantiated by words like *pukiki*, *mesasa*, etc.) and ABA (instantiated by words like *pukipu*, *mesame*, etc.). They showed that, after being habituated to exemplars of one of the structures, infants could discriminate between novel exemplars of both structures, suggesting that they had extracted and generalized the structures, rather than memorized a series of speech sequences.

There is a current debate about the precise mechanisms that underlie the extraction and generalization of repetition structures (Endress, Nespor & Mehler, 2009; Endress *et al*., 2005; Marcus *et al*., 1999). In the current paper, however, we will not discuss the mechanisms involved, but rather the constraints that may apply to these mechanisms. What matters to our present investigation of the respective roles of consonants and vowels in language acquisition is that the generalization of a repetition structure requires abstraction abilities that cannot be reduced to either memory or statistical dependencies between syllables or other constituents. Thus, in that respect, we take the generalization of a repetition structure as being comparable to the extraction of syntactic relations. Repetition structures can be viewed as the relation between two elements: infants display a generalization ability when they can recognize a repetition structure in novel syllable sequences that they have not encountered before (Kovács & Mehler, 2009b; Marcus *et al*., 1999). The generalization of repetition structures can thus be used as a simplified model that might clarify how infants extract syntactic information.

The CV hypothesis – claiming that vowels are favored in the extraction of structural information – predicts that repetition-based structures should be easier to detect and generalize when they are implemented over vowels than over consonants. Experimental work corroborated this prediction for adults. Toro, Nespor, Mehler and Bonatti (2008a) showed that adult participants could easily learn the ABA regularity over vowels and generalize it to words using new vowels and consonants but respecting the same structure. In contrast, they were unable to learn the same regularity over consonants. Adults remain unable to generalize ABA over consonants even when vowel duration was reduced to one-third of the duration of consonants, while they could generalize ABA over barely audible vowels (Toro, Shukla, Nespor & Endress, 2008b). Thus, the reliance on vowels for extracting repetition-based regularities is not solely due to a major acoustic salience. Rather, vowels and consonants are involved in different types of processes, as suggested by the existence of a different neural substrate for each category (Caramazza, Chialant, Capasso & Miceli, 2000; Knobel & Caramazza, 2007).

Our study

In the experimental work presented below, we ask whether the documented functional difference between consonants and vowels can play a role in early steps of language acquisition. Evidence for the lexical role of consonants exists for participants older than 16 months (Havy & Nazzi, 2009), who already have a sizable receptive vocabulary (about one hundred and sixty words according to the French version of the MacArthur-Bates Communicative Development Inventory developed by S. Kern; personal communication). Moreover, evidence for the specialized role of vowels was reported for adults (Toro *et al.*, 2008a, 2008b). Recently, Pons and Toro (2010) suggested that 11-month-olds were already better at learning the AAB structure over vowels than over consonants. In a preferential looking paradigm, infants habituated to words respecting the AAB structure over vowels (e.g. *dabale*, *tolodi*, *tibilo*) could discriminate between novel words respecting the same structure (e.g. *nadato*, *lotoba*, *dilite*), and words that did not respect the AAB structure (e.g. *dutone*, *lanude*, *bitado*). In contrast, infants habituated to the AAB structure instantiated over consonants (e.g. *dadeno*, *lulabo*, *nunide*) did not discriminate novel words respecting the structure (e.g. *dedulo*, *lulina*, *nunobi*) from words that did not respect the structure (e.g. *dutani*, *litedo*, *bilune*). However, Pons and Toro (2010) used the same vowels and consonants in the test phase as in the familiarization phase. In particular, test and familiarization words shared the repeated vowels (e.g. *dabale* as a familiarization word, and *batalo* as a test word). Thus, an alternative explanation of Pons and Toro's results is that infants process mainly vowels and learned the repeated tokens (i.e. *aa*, *oo*, *uu*, *ii*, *ee*). They did not necessarily generalize the AAB structure, but may rather remember repeated vowel sequences. Showing generalization would require using novel vowels and consonants to form novel words in the test phase.

Taken together, the results reviewed above do not fully demonstrate the functional specialization of consonants and vowels in infancy. In fact, all the results could be explained by a switch in a general processing bias from vowels (at 11 months) to consonants (by 16 months) in infancy, while functional specialization would emerge only in adulthood. Thus, the aims of our study are, first, to assess whether the consonantal bias for lexical processes is already functional by 12 months, an age at which infants just start building the lexicon (Lock, 1980; McShane, 1979; Stager & Werker, 1997); second, to assess 12-month-olds' ability to generalize repetition structures over consonants and vowels. Moreover, by testing 12-month-olds in two very similar paradigms, which vary only in those details that differentiate a word learning experiment from a situation promoting the discovery of structural relations, we directly assess the CV hypothesis of a functional specialization of consonants and vowels.

Twelve-month-olds' word learning abilities are still immature (Lock, 1980; McShane, 1979; Stager & Werker, 1997). Infants begin to learn words in a very fast manner only towards the end of the second year of life (Carey & Bartlett, 1978; Golinkoff, Church Jacquet, Hirsh-Pasek & Nandakumar, 1996; Golinkoff, Hirsh-Pasek, Bailey & Wenger, 1992; Heibeck & Markman, 1987) after the emergence of a series of conceptual constraints that support word learning (Halberda, 2003; Hochmann, Endress & Mehler, 2010; Mervis, 1987; Markman, 1990; Markman & Hutchinson, 1984; Markman, Wasow & Hansen, 2003; Soja, Carey & Spelke, 1985). If the lexical role of consonants is already set before this stage, it can constrain and shape infants' vocabulary development. Similarly, even if numerous studies have shown that infants are sensitive to the boundaries of syntactic constituents (Gout, Christophe & Morgan, 2004; Soderstrom, Seidl, Kemler Nelson & Jusczyk, 2003), it remains unknown how much 12-month-olds know about the relation among these constituents. If we can establish that vowels are privileged for the extraction of such relational information, we may increase our understanding of the mechanisms present in syntax acquisition.

Experiment 1 tests the first part of the CV hypothesis (see Figure 1A). We ask whether infants rely more on consonants or on vowels when distinguishing among words. Adapting the paradigm developed by Kovács (2008) and Kovács & Mehler (2009b), we teach infants that one word predicts a toy's appearance on one side of the screen (e.g. *dudu*), while another word predicts a toy's appearance on the other side (e.g. *keke*). Infants are then presented with an ambiguous word, composed of the consonants of the former word and the vowels of the latter (e.g. *dede*) (or vice versa, e.g. *kuku*). On this occasion, no toy appears. Reliance on consonants would lead infants to search for the toy on the location pre-

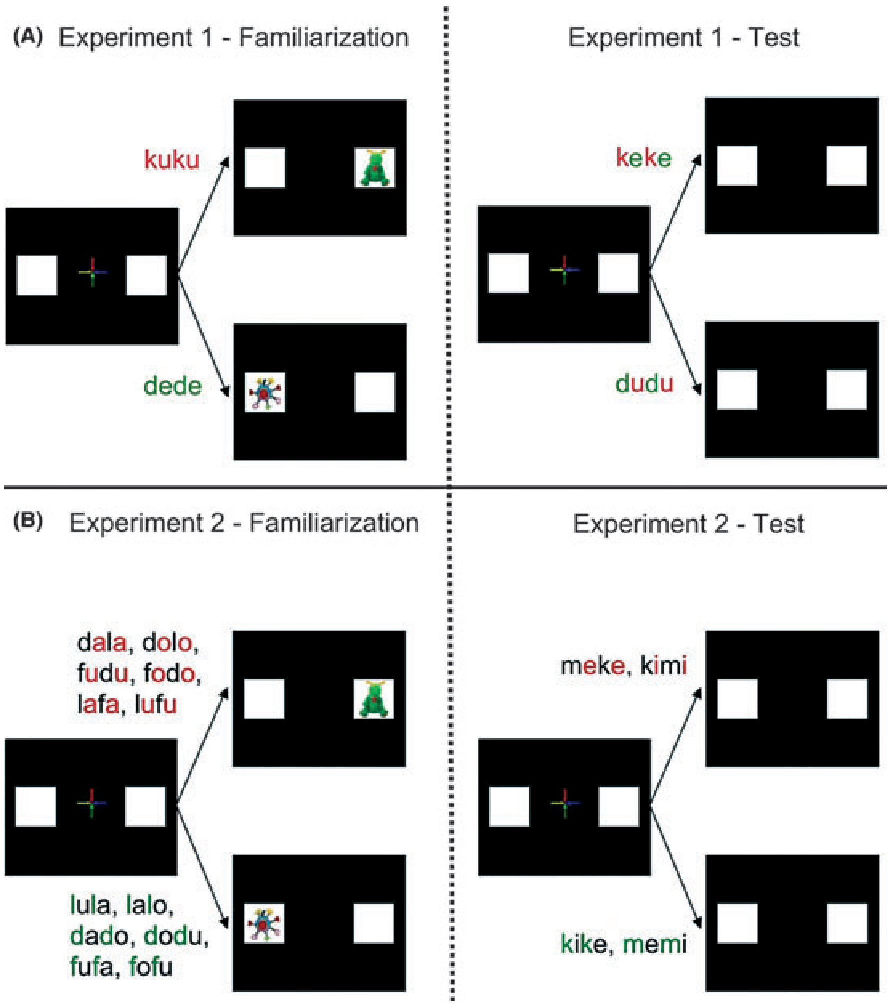

Figure 1 *Experimental paradigm for Experiments 1 (A) and 2 (B).*

dicted by the first word and reliance on vowels would lead infants to search for the toy on the location predicted by the second word. When forming memory representations of a novel word, the CV hypothesis predicts that infants should rely more on consonants than on vowels.

Experiment 2 tests the second part of the CV hypothesis (see Figure 1B), asking whether infants find it easier to learn and generalize regularities defined over vowels or over consonants. We again use the paradigm developed by Kovács (2008) and Kovács and Mehler (2009b). Monolingual 12-month-olds have trouble learning simultaneously two regularities in that paradigm, probably due to limitations in executive function (Kovács & Mehler, 2009b). Indeed, they just learn the simpler one (e.g. adjacent rather than non-adjacent syllable repetition, AAB vs. ABA; repetition rather than absence of repetition, AA vs. AB, ABA vs. ABC; Kovács, 2008). Thus, our paradigm allows us to test in a within-subject design which one of two structural relations is easier for infants to detect and generalize. In our experiment, we teach infants two regularities, consonant repetition and vowel repetition, in order to determine which is easier for 12-month-olds to learn. Each of six words containing a consonant repetition (i.e. *lula*, *lalo*, *dado*, *dodu*, *fufa* and *fofu*) was followed by a toy which appeared on one side of the screen, whereas each of six words containing a vowel repetition (i.e. *dala*, *dolo*, *fodo*, *fudu*, *lafa* and *lufu*) was followed by a toy which appeared on the other side of the screen. We then tested for generalization, asking where infants would search for the toy when hearing novel words respecting either the consonant repetition regularity (i.e. *kike* and *memi*) or the vowel repetition regularity (i.e. *meke* and *kimi*), using novel vowels and consonants that did not appear during the familiarization. These two regularities are strictly equivalent in terms of complexity (a simple repetition), varying only in the category that carries the repetition. The CV hypothesis predicts that the generalization of a repetition regularity should be easier if it is implemented over vowels than if it is implemented over consonants. Therefore, when learning only one regularity, infants should learn and generalize the vowel repetition rather than the consonant repetition.

Experiment 1

Participants

Twenty-six infants were included in the analysis; mean age: 12 months 6 days; range: 11 months 18 days–12 months 26 days. Five other infants participated in the study but were excluded due to fussiness (three) or experimental failure (two). Parents of the infants participating in the two experiments signed the informed consent explanation form before the experiments. The Ethics Committee of SISSA, where the experiments were conducted, approved the study design.

Stimuli

The words in Experiment 1 consisted of one syllable that duplicates. We used two consonants and two vowels to construct four nonsense words: *kuku*, *dede*, *keke* and *dudu*. The two consonants differ in three features, i.e. *k* is back, high and unvoiced, whereas *d* is anterior, coronal and voiced. The two vowels differ in three features, i.e. *e* is unrounded, mid and front, whereas *u* is rounded, high and back.

Two words sharing neither consonants nor vowels were used in the Familiarization (i.e. *kuku* and *dede*). The two remaining words were used in the Test. Words were synthesized with MBROLA (fr4) with a phoneme duration of 120 ms and a monotonous pitch of 200 Hz. There was no silent pause between two syllables within a word.

The visual stimuli were two pictures of colorful toys. Each appeared inside one of two white squares either on the left or on the right side of the screen. The toys loomed from 4 cm to 7 cm inside the squares for 2 s. The squares had a side-length of 8 cm, positioned at a distance of 13.5 cm. In the Familiarization, each toy was paired with one Familiarization word and one side.

Procedure

The procedure was adapted from Kovács and Mehler (2009b) and is presented in Figure 1. Stimuli were presented via an Apple Dual G5 computer running Psyscope X (http://psy.ck.sissa.it). Infants' gaze was recorded with a TOBII 1750 Eye-Tracker (Hofsten, Dahlstrom & Fredrikson, 2005).

The Familiarization phase consisted of 32 Familiarization trials. Familiarization trials started with a display of two white squares on the sides and a central attention-grabber. When the infant looked at the attention-grabber, either one of the two familiarization words was played in a pseudo-random order. We ensured that no word was repeated more than three times in a row. The animated attention-grabber was displayed until the offset of the word, in order to keep the infant's gaze in the middle of the screen. One second after the word offset, a toy appeared in one of the squares, contingent on the word: one word predicted the toy's appearance in one of the squares, while the other word predicted the toy's appearance in the other square. The pairing of the words with toy locations was counterbalanced across participants.

During test, infants were exposed to eight trials in a pseudo-random order. Test trials were similar to the Familiarization trials, except that infants heard words constituted by the consonants of one of the Familiarization words, and the vowels of the other. Thus if the Familiarization words were *dudu* and *keke*, the Test words were *dede* and *kuku*. If the Familiarization words were *dede* and *kuku*, the Test words were *dudu* and *keke*. No toy ever appeared in the test trials. Two seconds after the word onset, the next trial started. Test lists were pseudo-randomized so that the four first trials consisted in two trials for each test word.

Analysis

For the analysis, we divided the screen into three equal parts, left, middle and right. In each trial of the familiarization, we measured the proportion of infants anticipating the toy's appearance to the correct side. In the test, we measured each participant's first fixation, after hearing the new word and before the beginning of the next trial. Infants were coded as targeting either the consonant side or the vowel side. The vowel side was the one where the toy appeared after hearing the familiarization word that had the same vowels as the test word. For example, the vowel side for the test word *keke* was the side where during familiarization they learned to turn to after hearing the word *dede*, whereas the consonant side was where during familiarization they learned to turn to after hearing the word *kuku*. The consonant side for one of the two test words corresponded to the vowel side for the other test word. We also measured the infants' overall accuracy, based on the time spent fixating the consonant side and vowel side of the screen (Kovács & Mehler, 2009b; McMurray & Aslin, 2004). That is, for each trial, infants were scored as searching to the consonant side if the infant looked longer to the consonant side within the 2 s after hearing a new item and before the start of the next trial. Infants were scored as searching to the vowel side otherwise.

We computed difference scores: (#consonant looks − #vowel looks)/(#consonant looks + #vowel looks) for first and for overall accuracy, and computed a *t*-test to compare them to the chance level of 0. Positive differences in scores indicate that infants searched for the toys on the consonant side, while negative difference scores indicate infants searched for the toys on the vowel side.

The eight test trials might display extinction effects due to the absence of the puppets during this phase of the study. Thus, we also ran our analyses considering only the results of the first four test trials.

Results

The first fixation data for familiarization trials are presented in Figure 2. Infants anticipated one or the other side in 55% of the trials. We computed the proportion of infants showing a correct anticipatory look for each trial. A linear regression analysis showed no tendency for more correct anticipations in later familiarization trials, ß = .0011, R^2 = .0109, $t(30)$ = .57, $p > .57$.

The test results are presented in Figure 3. In the test phase, infants looked to the left or the right in 65% of the trials. Considering first fixations, infants' mean difference score was .23, which was significantly greater

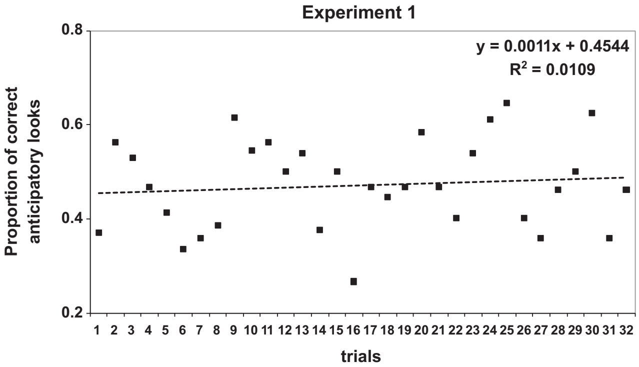

Figure 2 *Proportion of correct anticipatory looks for each familiarization trial in Experiment 1. The dotted line depicts the corresponding linear regression.*

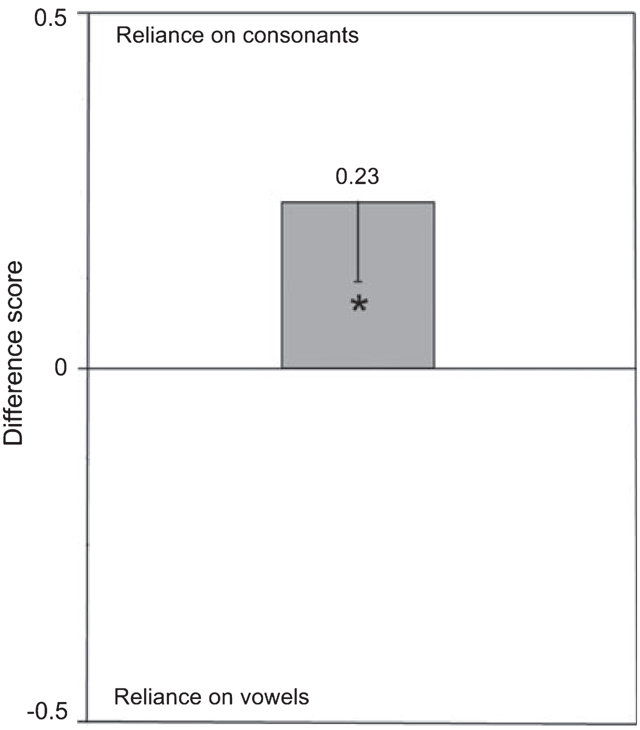

Figure 3 *Mean difference score for Experiment 1 considering the first fixations. Infants looked more at the side predicted by consonants. Error bars represent standard error.*

than 0, $t(25)$ = 2.1077; p = 045; d′ = .41. Seventeen infants displayed a positive difference score, six infants showed a negative difference score, and three infants a null difference score. A binomial test showed that significantly more infants displayed a positive difference score than a negative difference score, p = .035. Considering only the four first test trials, infants' mean difference score was .25, which was significantly greater than 0, $t(23)$ = 2.18; p = 039; d′ = .45. Fourteen infants displayed a positive difference score, six infants a negative difference score, and six infants a null difference score. The number of infants who displayed a positive difference score and the number of infants who displayed a negative difference score did not differ significantly, as evaluated by a binomial test, p = 0.12.

Considering the overall accuracy, infants' mean difference score was .13, which was not significantly different from 0; $t(25)$ = 1.33; p = .20; d′ = .26. Fifteen infants displayed a positive difference score, eight infants a negative difference score, and three infants a null difference score. A comparison of the number of infants who displayed a positive difference score and the number of infants who displayed a negative difference score was not significant, as evaluated by a binomial test, p = 0.21. Considering only the four first test trials, infants' mean difference score was .22, a non-significant trend in the predicted direction, $t(24)$ = 1.81; p = .072; d′ = .36. Fifteen infants showed a positive difference score, five infants a negative difference score, and five infants a null difference score. A binomial test showed that significantly more infants displayed a positive difference score than a negative difference score, p = .041.

Altogether, infants privileged the prediction made by consonants rather than that made by vowels.

Discussion

In this experiment, infants needed to learn that one word predicts a toy's appearance in one location, while another word predicts a different toy's appearance in another location. We further asked what prediction infants would make when presented with ambiguous words formed with the consonants of one of the previous words and the vowels of the other.

The observation of anticipatory looks in the familiarization did not show evidence of an increase in the number of correct anticipations. However, this absence of evidence should not be interpreted as infants' failure to learn the associations, as is shown by the test results. In fact, our paradigm is not designed to evaluate learning in the familiarization phase; rather, it is designed to assess the participants' performance in the test phase. In previous studies (Kovács, 2008; Kovács & Mehler, 2009b), the authors argued that due to limitations in executive function, most infants learned to predict the puppet's appearance for only one of the two screen locations. Similarly, in Experiment 1, most infants probably learned only to predict the puppet's appearance for only one of the two familiarization words. Thus, in all likelihood, infants' anticipations to the unlearned word, namely for 50% of the trials, should be assigned randomly to one or the other location. This predicts that data ought to be noisy during the familiarization phase.

Nevertheless, the observation of first fixations and the overall accuracy in the test phase suggests that infants consider two words sharing consonants as more similar than two words sharing vowels. Twelve-month-olds found *kuku* more similar to *keke* than to *dudu*.[2] A second interpretation is that infants relied solely on consonants when associating the familiarization words to either the side or the specific toy that appeared there. Our paradigm does not allow us to understand whether infants associated words to toys or to locations. In both cases, nevertheless, they needed to store at least one specific word in memory. When encoding specific words in memory, 12-month-old infants appear to give a higher weight to consonants than to vowels.

Experiment 2

In Experiment 2, we address the second part of the CV hypothesis, asking whether vowels are privileged for detecting and generalizing repetition structures. As in Experiment 1, infants heard speech sequences that predicted where toys would next appear. In Experiment 1, though, one word was associated with each location. In Experiment 2, six items were associated with each location. All six items associated with each location exemplified a common regularity, i.e. consonant repetition or vowel repetition.

Experiment 2, unlike Experiment 1, was not designed as a word learning experiment but as an experiment to elicit the generalization of a repetition structure. In that type of experiment, a variety of generalizations are possible from the limited set of examples that are provided to the participants. To avoid the position of specific phonemes providing a cue for generalizations, the same consonants and vowels were used to create the six consonant repetition and the six vowel repetition items used in the familiarization. Moreover, three puppets could appear in each location and were randomly paired with the items, so that the attention of infants could not be attracted by a particular puppet and an item associated with it. In the test phase, we asked whether infants could generalize these associations to novel items formed with novel consonants and vowels, instantiating the repetition structures. Predictions about the location of a toy's appearance in the test trials could be done only if one focused on the identity relation between consonants or vowels. Due to limitations in executive function, monolingual 12-month-olds usually learn to generalize only the easier structure when tested with this paradigm. We thus ask whether vowel repetition or consonant repetition is easier for 12-month-olds to generalize.

Participants

Twenty-four infants were included in the analysis; mean age: 12 months 4 days; range 11 months 26 days–12 months 28 days. Six other infants participated in the study but were excluded due to fussiness (three) or experimental failure (three).

Stimuli

The stimuli in Experiment 2 were bisyllabic items. These items could have either repeated consonants or repeated vowels. For the familiarization, six items containing a consonant repetition (*lula*, *lalo*, *dado*, *dodu*, *fufa* and *fofu*) and six items containing a vowel repetition were created (*dala*, *dolo*, *fodo*, *fudu*, *lafa* and *lufu*). The same three consonants and three vowels were used to generate both sets of items. For the test, four novel items were generated with novel consonants and novel vowels. Two test items had a consonant repetition (*kike* and *memi*) and two test items had a vowel repetition (*meke* and *kimi*). Items were synthesized with MBROLA (fr4) with phoneme durations of 120 ms and a monotonous pitch of 200 Hz. There was no silent pause between two syllables within an item.

Visual stimuli were three pictures of colorful toys. These appeared inside one of two white squares on the left or right side of the screen, as for Experiment 1. All three toys could appear after any familiarization item in the window predicted by that item.

Procedure

The procedure was identical to that of Experiment 1, except for the speech items infants heard (see Figure 1) and the toys they saw. In Experiment 2, a consonant repetition predicted the toys' appearance in one of the squares, while a vowel repetition predicted the toys' appearance in the other square. The pairing of structures with toy locations was counterbalanced across participants. Test lists were pseudo-randomized so that the four first and four last trials consisted of two trials for each structure. The structure of the first test item was vowel repetition for half the participants and consonant repetition for the other half.

[2] The Nazzi group has reported that, in older infants and children, the C-bias may vary depending on the actual contrasts used, even though it applies to most consonantal contrasts (Havy & Nazzi, 2009; Nazzi, 2005; Nazzi & Bertoncini, 2009; Nazzi *et al.*, 2009). In our experimental paradigm, infants were more likely to learn to predict the locations of the toys' appearances if the two novel words were easy to discriminate. For this reason, we chose words that differ by several rather than only one feature. Thus, we chose to use three-feature contrasts, both for consonants and for vowels. Further studies may ask whether one-feature contrasts yield similar results, particularly contrasts involving manner of articulation which did not vary in our experiment.

Analysis

The analysis was similar to that of Experiment 1, except that each fixation was coded as correct or incorrect, according to the structure that preceded infants' response, i.e. consonant repetition or vowel repetition. Independently for each structure, we computed difference scores: (#correct looks − #incorrect looks)/(#correct looks + #incorrect looks) for first fixations and for overall accuracy, and computed a *t*-test to compare them to the chance level of 0. Significantly positive difference scores would indicate that infants learned and generalized the structure.

Because the first four test trials contained only two test trials for each structure, unlike Experiment 1, we did not analyze separately these trials in Experiment 2.

Results

Familiarization results are presented in Figure 4. Infants anticipated to one or the other side in 56% of the trials. We computed the proportion of infants showing a correct anticipatory look for each trial. Qualitatively, the proportion of correct anticipatory looks increased during familiarization for the vowel repetition and decreased for the consonant repetition. A linear regression analysis yielded marginally significant results for the consonant repetition, ß = −.0159, R^2 = .21, $t(14) = -1.92$, $p = .075$, and a non-significant trend for the vowel repetition, ß = .0046, R^2 = .15, $t(14) = 1.60$, $p = .133$.

The test results are presented in Figure 5. Infants looked to the left or the right in 62% of vowel repetition test trials, and in 61% of the consonant repetition test trials. Two infants did not provide data in the vowel repetition tests, so that 24 infants were included in the analysis of the consonant repetition tests and only 22 in the analysis of the vowel repetition tests.

Paired *t*-test showed that infants obtained significantly higher difference scores for the vowel repetition than for the consonant repetition, considering the first fixations, $t(21) = 4.56$; $p < .0002$; d′ = .97; and the overall accuracy, $t(21) = 4.29$; $p < .0004$; d′ = .91.

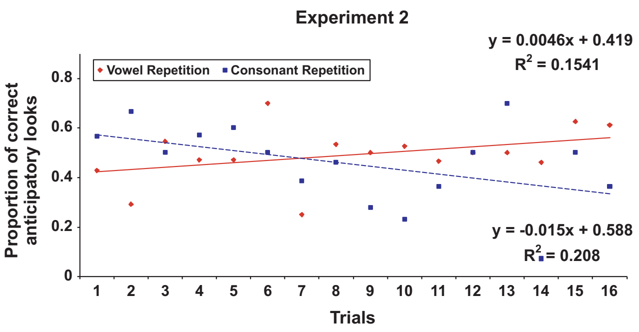

Figure 4 *Proportion of correct anticipatory looks for each familiarization trial in Experiment 2. The dotted lines depict the linear regression for vowel repetition (red) and for consonant repetitions (blue), respectively.*

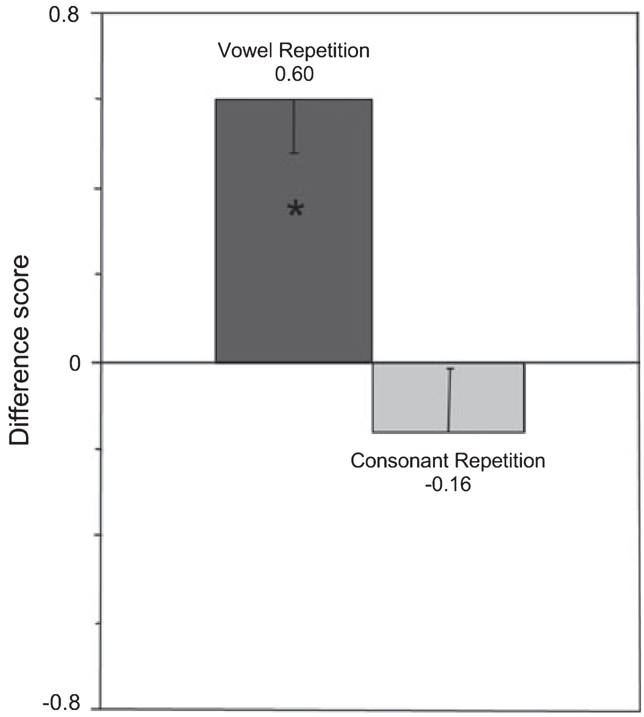

Figure 5 *Mean difference scores for Experiment 2 considering the first fixations. Infants looked more at the correct side predicted by the regularity for the vowel repetition, but not for the consonant repetition. Error bars represent standard errors.*

Considering first fixations in vowel repetition test trials, infants' mean difference score was .60 for the vowel repetition, which was significantly greater than 0, $t(21) = 4.92$; $p < .0001$; d′ = 1.05. Seventeen infants displayed a positive difference score, two infants a negative difference score, and three a null difference score. A binomial test showed that significantly more infants displayed a positive difference score than a negative difference score, $p < .001$. For the consonant repetition, infants' mean difference score was −0.16, which did not differ from chance, $t(23) = -1.06$; $p = .30$; d′ = .22. Eight infants displayed a positive difference score, 11 infants a negative difference score, and five a null difference score. This distribution did not differ from chance, $p = .144$.

Congruently, considering overall accuracy, infants' mean difference score was .53 for the vowel repetition, which was significantly greater than 0; $t(21) = 5.23$; $p < .0001$; d′ = 1.11. Fifteen infants displayed a positive difference score, one infant a negative difference score, and six a null difference score. A binomial test showed that significantly more infants displayed a positive difference score than a negative difference score, $p < .001$. Infants' mean difference score was −.07 for the consonant repetition, which did not significantly differ from chance, $t(23) = -.52$; $p = .61$; d′ = .11. Six infants displayed a positive difference score, nine infants a negative difference score, and nine a null difference score. This distribution did not differ from chance, $p = .61$.

Discussion

The CV hypothesis predicts that generalizing repetition structures should be easier over vowels than over consonants. Experiment 2 directly tested this claim by confronting infants with two competing regularities. Given the limited cognitive capacities of 12-month-olds (Kovács & Mehler, 2009a, 2009b), participants were expected to learn and generalize only one of the patterns. The results suggest that infants learned the association between the vowel repetition regularity and the predicted location of a toy's appearance. Moreover, they could extend this association to new, never heard, items with vowel repetition formed with novel consonants and vowels, thus showing that infants extracted an abstract property from the familiarization and represent an abstract vowel repetition structure. In contrast, they showed no evidence of learning and generalizing the consonant repetition structure. Thus, the same structure (i.e. a repetition[3]) is easier for 12-month-old infants to learn over vowels than over consonants. Vowels rather than consonants appear to be a privileged category for extracting and generalizing abstract structures.

General discussion

The CV hypothesis claims that infants rely more on consonants for word learning and recognition, but privilege vowels for extracting structural syntactic information. In two experiments, infants saw toys appearing in two distinct locations after hearing specific speech items. In Experiment 1, infants had to learn two different words, each associated with a specific location. Thereafter, in the test, infants relied more on consonants than on vowels in their search for the location of the toy. In Experiment 2, infants could generalize a repetition structure implemented over vowels, but not that implemented over consonants. Notably, Experiments 1 and 2 were designed in a similar way, and varied only in the details that should differentiate a word learning experiment from an experiment that evaluates the generalization of a structure. In Experiment 1, in addition to its location, each of the two toys was associated with a specific word, so that these could be interpreted as the names of the respective toys. In Experiment 2, the same consonants and vowels were used to implement both structures, the vowel repetitions and the consonant repetitions, in order to get infants to focus on the relation between segments rather than on the specific phonemes to predict the toy's location. The results allow us to conclude that consonants are privileged for lexical processes, whereas vowels better support the generalization of repetition structures. Thus, we show that functional differences between consonants and vowels are already playing a significant role by the end of the first year of life.

CV hypothesis, Part I: Consonants to build the lexicon

The end of the first year of life coincides with the time when infants begin to develop their vocabulary. According to the MacArthur-Bates Communicative Development Inventory (CDI; Dale & Fenson, 1996) questionnaire studies, infants of that age can understand about 80 words. Fenson, Dale, Reznick, Bates, Thal and Pethick (1994) evaluated that 12-month-olds learn about two words a week, a rate that improves in the following months yielding a 6-year-old vocabulary of about 10,000 words (Bloom & Markson, 1998; Miller, 1996). Thus, given that the consonantal bias for lexical acquisition appears to be in place at 12 months of age, it is likely to play a role in the acquisition of more than 99% of a child's vocabulary.

In the Supplementary Material, we show that the words in French and Italian infants' vocabulary can be better discriminated on the basis of the information carried by consonants than of that carried by vowels. In fact, Keidel *et al*. (2007) attributed the origin of the lexical role of consonants to the comparative distribution of information carried by vowels and consonants in the lexicon, which, in their view, would result from the larger number of consonants attested in the great majority of languages. However, the consonantal superiority for lexical processes is found in languages in which consonants largely outnumber vowels, as well as in languages in which the numbers of consonants and vowels are balanced (see Mehler *et al*., 2006; Toro *et al*., 2008a; and Bonatti, Peña, Nespor & Mehler, 2007, for Italian and French). For instance, Cutler, Sebastián-Gallés, Soler-Vilageliu and van Ooijen (2000) showed that, when asked to change one phoneme to turn a non-word (e.g. *kebra*) into a known word, participants more often altered the vowel (thus generating *cobra*) than the consonant (generating *zebra*). Their results were as significant for speakers of Spanish as for speakers of Dutch. Spanish has many more consonants than vowels, whereas Dutch has a similar number of consonants and vowels. Furthermore, in reading tasks, consonants appear to be privileged for lexical access in various languages such as French (New, Araujo & Nazzi, 2008), English (Berent & Perfetti, 1995; Lee, Rayner & Pollastek, 2001) and Spanish (Carreiras, Gillon-Dowens, Vergara & Perea, 2009). If the lexical statistics hypothesis fully explained the specialization of consonants for lexical access, a larger effect would be expected in languages that have many

[3] Note that, in our stimuli, an alternative description of the vowel repetition structure would be a *non*-repetition of consonants. Indeed, rather than generalizing the repetition of the vowel, generalizing the pattern in which the first and second consonants should differ would yield similar results. However, we consider this possibility unlikely, as Kovács (2008) showed how both 7- and 12-month-olds found it easier to learn and generalize a syllable repetition regularity rather than a non-repetition regularity.

more consonants than vowels. In contrast, we interpret the distribution of information in consonants and vowels as a consequence of the consonantal bias for lexical acquisition (see Bonatti *et al*., 2007).

The functional differences between consonants and vowels may be innate aspects of humans' ability to acquire language, or originate from the fact that different processes have different requirements, and may consequently rely on different categories. Specifically, lexical memory may require more stable, thus reliable, categories that allow the learner to identify words and distinguish each word from other lexical entries. The lexical role of consonants may therefore be due to the categorical mode in which consonants are perceived. In fact, speech perception abilities change in the course of the first year of life. Infants are initially sensitive to all phonemic contrasts that can be found in the languages of the world (Jusczyk, 1997; Mehler & Dupoux, 1994). By 6 months of age, however, infants display a perceptual magnet effect for vowel perception, which is due to the formation of a prototype for each of the vowels of the language of exposure (Kuhl, 1991; Kuhl, Williams, Lacerda, Stevens & Lindblom, 1992). By the end of the first year of life, infants converge to the consonantal categories of the language of exposure, and have lost some sensitivity to non-native consonantal contrasts (Werker & Tees, 1984). Even in adulthood, consonants are perceived more categorically than vowels. Indeed, in specific experimental conditions, when adult participants would fail to detect a within-category consonantal variation, they still detect a within-category vocalic variation (Fry, Abramson, Eimas & Liberman, 1962; Pisoni, 1973). As a consequence, vowel variations due to sentence prosody as well as speaker and dialectal accents may hinder word recognition to a greater extent than consonant variations. Therefore, lexical distinctions should be better instantiated by consonantal contrasts than by vocalic contrasts. This view predicts that the formation of consonantal phonological categories may be necessary for the lexical role of consonants to emerge.

CV hypothesis, Part II: Generalization of structural regularities over vowels

Inspired by the prosodic bootstrapping accounts of syntax acquisition (Morgan & Demuth, 1996), the CV hypothesis predicted that abstract structures should be easier to generalize over vowels than over consonants. In Experiment 2, we show that – in a within-subject design – 12-month-old infants are better at extracting a repetition-based structure over vowels than over consonants. Moreover, ours is the first experiment to show that infants can generalize the vowel repetition structure to completely novel words, formed with vowels and consonants that did not appear in the familiarization. This result indicates a special role of vowels in supporting the extraction and generalization of structures that cannot be reduced to statistical dependencies between constituents (e.g. syllables).

The origin of the specialization of vowels remains unknown, as does the origin of the consonantal bias in word learning. However, in the case of vowels, it is even harder to come up with a scenario where infants would have learned that vowels are more informative for specifying identity relations. Alternatively, the notion that vowels carry structural information may constitute an innate aspect of humans' ability to acquire language.

Consequently, 12-month-old and possibly younger infants ought to be capable of extracting the structural information carried by vowels, including prosodic information that informs syntax (Nespor & Vogel, 1986; Selkirk, 1984). For example, they may quickly learn and generalize prominence alternations signaled either by pitch (Bion, Benavides Varela & Nespor, in press) or by duration, which may allow them to learn the order of heads and complements in their language (Nespor *et al*., 2008).

Obviously, we are not claiming that all linguistic regularities are acquired through the detection and generalization of repetition patterns. However, the acquisition of natural syntactic constituent structures, similarly to the identification of repetition structures, requires the ability to generalize structural relations that cannot be reduced to a statistical regularity and thus require computations that go beyond the reach of memory and statistical computations (Chomsky, 1957; Marcus, 1998). Such computations may involve the formation of algebraic rules (Marcus *et al*., 1999), or the trigger of language-specific learning mechanisms. The precise characterization of the mechanisms involved will require further studies. However, here we have shown that these mechanisms are constrained in the speech domain, and particularly rely more on the vocalic tier than on the consonantal tier, suggesting that these mechanisms have evolved to suit the needs of human speech processing and language acquisition.

Conclusion

Here we document that the functional differences between consonants and vowels are already available to 12-month-old infants. The functional specialization of consonants is present early enough to shape the acquisition of most of the words a child will learn (possibly with the exception of a few very salient words learned very early on[4]). Furthermore, even though infants as young as 6 months are able to detect prosodic cues

[4] Note, however, that cross-linguistically, the words that indicate the mother and the father, two of the first words infants acquire, tend to contrast in their consonants but not in their vowels (e.g. *mamma/papa* in Italian; *anya/apa* in Hungarian; *aai/baba* in Marathi; *māma/bàba* in Mandarin Chinese; Jakobson, 1960; Murdoch, 1959). Thus the lexical role of consonants may already be set when these words are learned.

marking syntactic phrase boundaries (Gout *et al.*, 2004; Soderstrom *et al.*, 2003), it is less clear how they learn about the relation among them. Our results suggest that the structural information carried by vowels is available at least to 12-month-olds and may guide them in this task.

Infants may know or have a mechanism to identify which category better serves the lexicon and which the extraction of structural regularities. We proposed that the distributional and physical properties of consonants make them more reliable for word recognition and lexical distinctions, thus favoring them in word learning processes. Vowels, instead, carry prosody that marks more abstract constituents and provides the learner with information about structural relations. Future studies will explore the link between these specific properties of consonants and vowels and their specializations for language acquisition. These complementary constraints on the processing of consonants and vowels may allow the language acquisition device to work simultaneously on the extraction of regularities and on the acquisition of lexical items.

Acknowledgements

This research has been supported by McDonnell Foundation Grant No. 21002089 and by a grant to the second author from MICIT (Ministerio de Ciencia y Tecnología), CONICIT (Consejo Nacional para Investigaciones Científicas y Tecnológicas) of Costa Rica. We thank Francesca Gandolfo and Marijana Sjekloća for help recruiting participants, and Liuba Papeo and Alan Langus for valuable comments on a previous version of the manuscript.

References

Berent, I., & Perfetti, C.A. (1995). A rose is a reez: the two cycles model of phonology assembly in reading English. *Psychological Review*, **102**, 146–184.

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

Bion, R.H, Benavides Varela, S., & Nespor, M. (in press). Acoustic markers of prominence influence infants' and adults' memory for speech sequences. *Language and Speech*.

Bloom, P., & Markson, L. (1998). Capacities underlying word learning. *Trends in Cognitive Sciences*, **2**, 67–73.

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**, 451–459.

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

Caramazza, A., Chialant, D., Capasso, R., & Miceli, G. (2000). Separable processing of consonants and vowels. *Nature*, **403**, 428–430.

Carey, S., & Bartlett, E. (1978). Acquiring a single new word. *Papers and Reports on Child Language Development*, **15**, 17–29.

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

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

Christophe, A., Mehler, J., & Sebastián-Gallés, N. (2001). Infants' sensitivity for prosodic boundary cues. *Infancy*, **2**, 385–394.

Christophe, A., Nespor, M., Dupoux, E., Guasti, M.-T., & van Ooyen, B. (2003). Reflections on prosodic bootstrapping: its role for lexical and syntactic acquisition. *Developmental Science*, **6**, 213–222.

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

Dale, P.S., & Fenson, L. (1996). Lexical development norms for young children. *Behavioral Research Methods, Instruments, & Computers*, **28**, 125–127.

Donegan, P., & Stampe, D. (1983). Rhythm and the holistic organization of language structure. *Papers from the Annual Regional Meeting, Chicago Linguistic Society: Parasession on the Interplay of Phonology, Morphology and Syntax* (pp. 337–353).

Endress, A.D., Dehaene-Lambertz, G., & Mehler, J. (2007). Perceptual constraints and the learnability of simple grammars. *Cognition*, **105**, 577–614.

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

Endress, A.D., Scholl, B.J., & Mehler, J. (2005). The role of salience in the extraction of algebraic rules. *Journal of Experimental Psychology, General*, **134**, 406–419.

Fenk-Oczlon, G., & Fenk, A. (2005). Crosslinguistic correlations between size of syllables, number of cases, and adposition order. In Gertraud Fenk-Oczlon & Christian Winkler (Eds.), *Sprache und Natürlichkeit: Gedenkband für Willi Mayerthaler* (pp. 75–86). Tübingen: Gunther Narr.

Fenson, L., Dale, P., Reznick, J., Bates, E., Thal, D., & Pethick, J. (1994). Variability in early communication development. *Monographs of the Society for Research in Child Development*, **59** (5, Serial No. 242).

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

Gerken, L.A., Jusczyk, P.W., & Mandel, D.R. (1994). When prosody fails to cue syntactic structure: nine-month-olds' sensitivity to phonological vs. syntactic phrases. *Cognition*, **51**, 237–265.

Gil, D. (1986). A prosodic typology of language. *Folia Linguistica*, **20**, 165–231.

Gleitman, L.R., & Wanner, E. (1982). The state of the state of the art. In E. Wanner & L.R. Gleitman (Eds.), *Language acquisition: The state of the art* (pp. 3–48). Cambridge: Cambridge University Press.

Golinkoff, R.M., Church Jacquet, R., Hirsh-Pasek, K., & Nandakumar, R. (1996). Lexical principles may underlie the learning of verbs. *Child Development*, **67**, 3101–3119.

Golinkoff, R.M., Hirsh-Pasek, K., Bailey, L.M., & Wenger, N.R. (1992). Young children and adults use lexical

principles to learn new nouns. *Developmental Psychology*, **28**, 99–108.
Gout, A., Christophe, A., & Morgan, J. (2004). Phonological phrase boundaries constrain lexical access: II. Infant data. *Journal of Memory and Language*, **51**, 547–567.
Halberda, J. (2003). The development of a word-learning strategy. *Cognition*, **87**, B23–B34.
Havy, M. (2009). Spécificité phonétique et acquisition lexicale en situation de développement normal et atypique. Thèse de Doctorat en Psychologie Cognitive, Université Paris Descartes, soutenance publique: 21 décembre.
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.
Heibeck, T.H., & Markman, E.M. (1987). Word learning in children: an examination of fast mapping. *Child Development*, **58**, 1021–1034.
Hochmann, J.-R., Endress, A.D., & Mehler, J. (2010). Word frequency as a cue for identifying function words in infancy. *Cognition*, **115**, 444–457.
Hofsten, C., Dahlstrom, E., & Fredrikson, Y. (2005). 12-month-old infants' perception of attention direction in static video images. *Infancy*, **8**, 217–231.
Jakobson, R. (1960). Why mama and papa? In B. Kaplan & S. Wapner (Eds.), *Perspectives in psychological theory: Essays in honor of Heinz Werner* (pp. 124–134). New York: International Universities Press.
Johnson, S.P., Fernandes, K.J., Frank, M.C., Kirkham, N.Z., Marcus, G.F., Rabagliati, H., & Slemmer, J.A. (2009). Abstract rule learning for visual sequences in 8- and 11-month-olds. *Infancy*, **14**, 2–18.
Jusczyk, P.W. (1997). *The discovery of spoken language*. Cambridge, MA: MIT Press.
Keidel, J.S., Jenison, R.L., Kluender, K.R., & Seidenberg, M.S. (2007). Does grammar constrain statistical learning? Commentary on Bonatti et al. (2005). *Psychological Science*, **18**, 922–923.
Kemler Nelson, D.G., Hirsh-Pasek, K., Jusczyk, P.W., & Wright-Cassidy, K. (1989). How prosodic cues in motherese might assist language learning. *Journal of Child Language*, **16**, 55–68.
Knobel, M., & Caramazza, A. (2007). Evaluating computational models in cognitive neuropsychology: the case from the consonant/vowel distinction. *Brain and Language*, **100**, 95–100.
Kovács, Á.M. (2008). Learning two languages simultaneously: mechanisms recruited for dealing with a mixed linguistic input. SISSA, Trieste. PhD dissertation.
Kovács, Á.M., & Mehler, J. (2009a). Cognitive gains in 7-month-old bilingual infants. *Proceedings of the National Academy of Sciences, USA*, **106** (2), 6556–6560.
Kovács, Á.M., & Mehler, J. (2009b). Flexible learning of multiple speech structures in bilingual infants. *Science*, **325**, 611–612.
Kuhl, P.K. (1991). Human adults and human infants show a 'perceptual magnet effect' for the prototypes of speech categories, monkeys do not. *Perception and Psychophysics*, **50**, 93–107.
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.
Lee, H.W., Rayner, K., & Pollastek, A. (2001). The relative contribution of consonants and vowels to word identification during reading. *Journal of Memory and Language*, **44**, 189–205.
Lehiste, I. (1970). *Suprasegmentals*. Cambridge, MA: MIT Press.
Lock, A. (1980). *The guided reinvention of language*. London: Academic Press.
McCarthy, J.J. (1985). *Formal problems in semitic phonology and morphology*. New York and Oxford: Garland.
McMurray, B., & Aslin, R.N. (2004). Anticipatory eye movements reveal infants' auditory and visual categories. *Infancy*, **6**, 203–229.
McShane, J. (1979). The development of naming. *Linguistics*, **17**, 879–905.
Maddieson, I. (2008). Consonant–vowel ratio. In M. Haspelmath, M.S. Dryer, D. Gil, & B. Comrie (Eds.), *The world atlas of language structures online*. Munich: Max Planck Digital Library, ch. 3. Available online at: http://wals.info/feature/3
Marcus, G.F. (1998). Rethinking eliminative connectionism. *Cognitive Psychology*, **37**, 243–282.
Marcus, G.F., Vijayan, S., Bandi Rao, S., & Vishton, P.M. (1999). Rule-learning in seven-month-old infants. *Science*, **283**, 77–80.
Markman, E.M. (1990). Constraints children place on word meanings. *Cognitive Science*, **14**, 57–77.
Markman, E.M., & Hutchinson, J.E. (1984). Children's sensitivity to constraints on word meaning: taxonomic vs. thematic relations. *Cognitive Psychology*, **16**, 1–27.
Markman, E.M., Wasow, J.L., & Hansen, M.B. (2003). Use of the mutual exclusivity assumption by young word learners. *Cognitive Psychology*, **47**, 241–275.
Mehler, J., & Dupoux, E. (1994). *What infants know*. Cambridge, MA: Blackwell.
Mehler, J., & Nespor, M. (2004). Linguistic rhythm and the development of language. In A. Belletti & L. Rizzi (Eds.), *Structures and beyond: The cartography of syntactic structures* (pp. 213–221). Oxford: Oxford University Press.
Mehler, J., Peña, M., Nespor, M., & Bonatti, L.L. (2006). The 'soul' of language does not use statistics: reflections on vowels and consonants. *Cortex*, **42**, 846–854.
Mervis, G.B. (1987). Child-basic object categories and early lexical development. In U. Neisser (Ed.), *Concepts and conceptual development: Ecological and intellectual factors in categorization* (pp. 201–233). Cambridge: Cambridge University Press.
Miller, G.A. (1996). *The science of words* (2nd edn.). New York: W.H. Freeman.
Morgan, J.L., & Demuth, K. (Eds.) (1996). *Signal to syntax: Bootstrapping from speech to grammar in early acquisition*. Mahwah, NJ: Lawrence Erlbaum Associates.
Murdock, G.P. (1959). Cross language parallels in parental kin terms. *Anthropological Linguistics*, **1**, 1–5.
Nazzi, T. (2005). Use of phonetic specificity during the acquisition of new words: differences between consonants and vowels. *Cognition*, **98**, 13–30.
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., Bertoncini, J., & Mehler, J. (1998). Language discrimination by newborns: towards an understanding of the

role of rhythm. *Journal of Experimental Psychology: Human Perception and Performance*, **24**, 756–766.

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

Nazzi, T., & New, B. (2007). Beyond stop consonants: consonantal specificity in early lexical acquisition. *Cognitive Development*, **22**, 271–279.

Nespor, M., Peña, M., & Mehler, J. (2003). On the different roles of vowels and consonants in speech processing and language acquisition. *Lingue & Linguaggio*, **2**, 201–227.

Nespor, M., Shukla, M., & Mehler, J. (2011). Stress-timed vs. syllable-timed languages. In M. van Oostendorp, C.J. Ewen, E. Hume, & K. Rice (Eds.), *The Blackwell companion to phonology*. 5 vols. Malden, MA and Oxford: Wiley-Blackwell.

Nespor, M., Shukla, M., van de Vijver, R., Avesani, C., Schraudolf, H., & Donati, C. (2008). Different phrasal prominence realization in VO and OV languages. *Lingue e Linguaggio*, **7**, 1–28.

Nespor, M., & Vogel, I. (1986). *Prosodic phonology*. Berlin: Mouton de Gruyter. (2nd edn., 2008. Dordrecht: Foris.)

New, B., Araujo, V., & Nazzi, T. (2008). Differential processing of consonants and vowels in lexical access through reading. *Psychological Science*, **19**, 1223–1227.

Newport, E.L., & Aslin, R.N. (2004). Learning at a distance I. Statistical learning of non-adjacent dependencies. *Cognitive Psychology*, **48**, 127–162.

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

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

Prunet, J.F., Beland, R., & Adrissi, A. (2000). The mental representation of Semitic words. *Linguistic Inquiry*, **31**, 609–648.

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

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

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

Saffran, J.R., Pollak, S.D., Seibel, R.L., & Shkolnik, A. (2007). Dog is a dog is a dog: infant rule learning is not specific to language. *Cognition*, **105**, 669–680.

Selkirk, E.O. (1984). *Phonology and syntax: The relation between sound and structure*. Cambridge, MA: MIT Press.

Shi, R., Morgan, J., & Allopenna, P. (1998). Phonological and acoustic bases for earliest grammatical category assignment: a cross-linguistic perspective. *Journal of Child Language*, **25**, 169–201.

Shukla, M., & Nespor, M. (2010). Rhythmic patterns cue word order. In N. Erteschik-Shir (Ed.), *The sound patterns of syntax* (pp. 174–188). Oxford: Oxford University Press.

Soderstrom, M., Seidl, A., Kemler Nelson, D., & Jusczyk, P. (2003). The prosodic bootstrapping of phrases: evidence from prelinguistic infants. *Journal of Memory and Language*, **49**, 249–267.

Soja, N., Carey, S., & Spelke, E. (1985). Constraints on word learning. Paper presented at the 1985 Biennial Convention of the Society for Research in Child Development, Toronto, Canada.

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.

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

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

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.

Received: 30 August 2010
Accepted: 27 June 2011

Supporting Information

Additional supporting information may be found in the online version of this article.

Figure S1 Percentage of different consonant sequences (in blue) and different vowel sequences (in red) for vocabulary sizes from 10 to 400 words, in Italian.

Figure S2 Percentage of different consonant sequences (in blue) and different vowel sequences (in red) for vocabulary sizes from 10 to 400 words, in French.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.