

Verbal Positional Memory in 7-months-olds

Silvia Benavides-Varela^{1,2*}, Jacques Mehler¹

¹International School for Advanced Studies (SISSA, ISAS), 34136 Trieste, Italy.

² IRCCS Fondazione Ospedale San Camillo, 30126 Lido-Venice, Italy.

*Correspondence to:

Silvia Benavides-Varela

IRCCS Fondazione Ospedale San Camillo

Via Alberoni 70, 30126 Lido di Venezia, Italia

Phone: +39 0412207111

Email: silviabenavides@gmail.com

Abstract:

Verbal memory is a fundamental pre-requisite for language learning. This study investigated 7-month-olds' (N=62) ability to remember the identity and order of elements in a multisyllabic word. The results indicate that infants detect changes in the order of edge syllables, or the identity of the middle syllables, but fail to encode the order of middle syllables. This suggests that the representational format of multisyllabic words is determined by core mnemonic biases, which favor accurate encoding of edges and limits the encoding of temporal order for internal segments. The studies support accounts proposing that content and order are encoded separately; in addition the data show that this dissociation occurs early in development.

Infants are constantly exposed to new sequences of the sounds that make up words. They extract words from the continuous speech signal using transitional probabilities between adjacent syllables (Saffran, Aslin, & Newport, 1996), word-level stress patterns (Bion, Benavides-Varela, & Nespor, 2011), phonotactic regularities (Mattys & Jusczyk, 2000), allophonic variation (Jusczyk, Hohne, & Bauman, 1999), prosody (Shukla, White, & Aslin, 2011), and cues from other familiar sounds (Bortfeld, Morgan, Golinkoff, & Rathbun, 2005).

In addition to identifying word boundaries, “knowing a word” requires memory for the sound of the word and its semantic reference. A number of behavioral (Swain, Zelazo, & Clifton, 1993; Valiante, Barr, Zelazo, Papageorgiou, & Young, 2006; Horne, Barr, Valiante, Zelazo & Young, 2006) and neuroimaging studies with newborns (Benavides-Varela, Gómez, Macagno, Bion, Peretz, & Mehler, 2011; Benavides-Varela, Hochmann, Macagno, Nespor, & Mehler, 2012; Partanen, Kujala, Naatanen, Liitola, Sambeth, & Huutilainen, 2013) suggest that the basic mechanisms for recognizing familiar word-sounds are already functional at birth. Moreover, by the second half of their first year of life, infants begin recognizing words commonly used in their environment (Hallé & Boysson-Bardies, 1994; 1996; Vihman, Nakai, DePaolis, & Hallé, 2004), words heard in isolation and embedded in short sentences (Kooijman, Junge, Johnson, Hagoort, & Cutler, 2013), or repeatedly heard from storybooks (Jusczyk & Hohne, 1997). They also display some ability to learn new word-object associations. For example, they recognize their own name (Mandel, Jusczyk, & Pisoni, 1995; Parise & Csibra, 2012), associate words like *mommy* and *daddy* with their parents’ faces (Tincoff & Jusczyk, 1999), establish new object-word mappings (Friedrich & Friederici, 2011), and relate common words used in their environment (body parts or food names) to pictures representing these nouns (Bergelson & Swingley, 2012).

The literature to which we refer constitutes compelling evidence that basic auditory memory mechanisms for identifying previously heard word-forms and learning new word-object associations are active in the second half of the first year of life. Most of these studies, however, largely concentrated on the content of infants' memories, by asking whether or not participants remembered a word-sound at a given age and requesting a yes-no answer (i.e. "presence" or "absence" of recognition).

Despite the importance of knowing what infants remember, there are cases in which this information does not suffice to determine the functioning of memory, or how this cognitive function affects language acquisition. For example, one aspect of memory that should influence speech representations during development concerns the configuration of the representation itself, namely when or in what order speech information gets encoded. In fact, many aspects of language acquisition, including recognition of multisyllabic words and word order at the phrasal and sentence levels, depend on encoding the serial order of sounds. For example, a young learner of Italian must at some point remember the order of syllables to distinguish between *mare* and *rema* (sea and paddles), *rape* and *pera* (beet and pear) or *mora* and *ramo* (blackberry and branch). Similarly a learner of English must recognize the difference between "the girl touches the boy" vs. "the boy touches the girl". It would be impossible for a learner incapable of encoding the order of sounds to grasp the differences between such lexical and syntactical distinctions and thus to fully acquire the properties of a given language. Still this particular aspect of early memory development has rarely been integrated in the context of language acquisition (see Wojcik, 2013 for a recent review on early memory mechanisms supporting word learning).

A few studies have addressed the question of whether young infants remember sequential order in the visual and motor domain. In these studies, infants were tested for their ability to track order of events (Leslie & Keeble, 1987) and imitation (Bauer & Mandler, 1992). Participants seemed to remember the order of sequences that have internal structure i.e. events that were causally related, but not those that were arbitrarily presented (see also Gulya, Rovee-Collier, Galluccio & Wilk, 1998; for results on long-term retention of events).

In the auditory domain, Thiessen and Saffran found that 6-8 months-old infants successfully detected changes in number word sequences (e.g., 13589 vs 18539) only when the numbers were sung but not when the numbers were spoken in adult-directed fashion (Thiessen & Saffran, 2009). Likewise, Mandel and colleagues showed that 2 months-old infants detected changes in the order of phonetically similar words (i.e. “Cats would jump benches” vs “Cats jump wood benches”) only when the sentences constituted a single prosodic unit, but failed to do so when the sequence was formed of unrelated sentential fragments (Mandel, Kemler-Nelson, & Jusczyk, 1996). Both studies showed that young infants require additional cues (i.e. melody in Thiessen & Saffran’s study, prosodic sentential units in Mandel *et al.*) to accurately remember the internal order of speech sequences. In fact, in the absence of these cues infants failed to detect changes in the order of words. It is worth mentioning that both these studies investigated the infants’ ability to detect the position of internal elements. Whether participants would have succeeded at encoding the order of items located in other positions of the sequence (i.e. the beginnings and endings of utterances) is an open question. The answer to this question may find its grounds in the adult memory literature, which we briefly describe below.

A long tradition in adult memory studies has shown that when adults are asked to retain new verbal sequences, they remember the first and last items better than the items in the middle

of a sequence. These phenomena –called serial position effects– arise in verbal learning contexts regardless of whether adults are requested to learn a list of separate words or a single sequence of syllables (i.e. a multisyllabic non-word) (Gupta, Lipinski, Bradon, & Po-Han, 2005). Thus, if core properties of sequential memory similar to the ones observed in adults influence infants' word representations, one should expect infants to represent sounds as a function of the position of these sounds within a word.

Additionally, a more profound analysis of the mechanisms that underlie serial learning in adults supports the idea that memory for serial order has two separate components: one for encoding content and another for position of items (e.g., Henson 1998, 1999; Marshuetz, 2005; Ng & Maybery 2002). Crucially, the models in adults also postulate that only edges have proper positional codes, explaining why the position of edges is encoded far more reliably than non-edge positions (Henson, 1998; 1999; Ng and Maybery 2002; Page & Norris, 1998). These properties of positional memory in adults derive from a number of studies in adults showing, for example, that participants often recall a series of items correctly but in an order different from that in which the items were originally presented; they also recall items in the right position across trials but in a sequence different from the one in which they originally appeared (Bjork & Healy, 1974; Henson, Hartley, Burgess, Hitch, & Flude, 2003). Furthermore, it has been shown that order information is lost faster than item information (Bjork & Healy 1974); and that certain concurrent tasks such as finger tapping, interfere with item memory but not with order memory (Henson *et al.*, 2003).

On the basis of this literature and a number of studies in adults, Endress and colleagues have proposed that information located at the edges serves as an anchor for speech representation, and that encoding the edges may be a primitive memory mechanism used for

language learning (Endress, Nespor & Mehler, 2009). In fact, in artificial grammar learning experiments, these authors observed that adult participants extracted structural regularities (e.g. found that a syllable was repeated in a sequence) only when this regularity was located at the edges of sequences (e.g. ABCDEFF, where each letter stands for a syllable) but had greater difficulty when the same regularities were located in the middle of sequences (e.g. ABCDDEF; Endress, Scholl, & Mehler, 2005). Similarly, additional studies showed that participants learned ‘phonotactic constraints’ (e.g. that certain consonants must occur in certain positions within a word) when the critical consonants occupied onset or coda positions but not word-internal positions (Chambers, Onishi & Fisher, 2003; Endress & Mehler, 2010; Onishi, Chambers & Fisher, 2002). Likewise, adults learned that words in an artificial language contained syllables that belonged to distinct classes only when they were in first and last positions but not when they were in medial positions (Endress & Bonatti, 2007; Endress & Mehler 2009). Endress *et al.*, thus suggested that the structural regularities of language are learned through mechanisms similar to those used more generally to track positions in a sequence. In other words they propose the existence of a specialized mechanism devoted to learning grammatical (and presumably other) structures, which is determined by the properties of a more general mechanism for positional memory (Endress *et al.*, 2009).

This view supports the idea that memory predisposes language by facilitating the acquisition of information located at the edges. The question remains however, as to whether memory mechanisms “predispose” language acquisition toward certain developmental patterns from early infancy. Are the items located at the edge also easy to learn in infancy? Are the properties of positional memory deployed, for instance, when infants begin encoding words? If infants experience sequential learning effects similar to those observed in adults, or if their

representations are anchored mainly at the edges -as proposed by Endress *et al.*- we expect infants to remember syllables located in the first and last positions of multisyllabic words more accurately than those in the middle. The following experiments were designed to investigate these questions.

Our aim was to explore whether 7-months-olds are able to remember the content and order of syllables in non-stressed multisyllabic words. In the first experiment, participants were familiarized with a pentasyllabic word that was associated with a puppet appearing on one side of a screen. In the test phase we recorded the infants' anticipatory looking behavior after the presentation of a test word. We presented test words that differed from the familiarization word in either the edge syllables or the middle syllables, in order to assess the degree to which the position of the syllables influences early word representations. In the second experiment, we tested infants with words that differed from the familiarization word in either the position or the content of the syllables. Such contrast enabled a direct assessment of memory for content and order of syllables in word-learning contexts. Finally, as a means to exclude alternative interpretations to the first experiments, in the third experiment we also tested infants in their ability to recognize the exact pentasyllabic word with which they were previously familiarized.

Experiment 1

Methods

Participants: Twenty-two infants participated in the study (12 females; aged from 6 months 15 days to 7 months 27 days; mean age 7 months 5 days). All participants were full term, with no birth complications, and no hearing or visual problems. Eleven infants were excluded from the analysis because of crying or fussiness ($N = 5$), equipment failure ($N = 5$), or because they looked at the screen in only one test trial ($N = 1$). Infants' parents signed the informed

consent after they had understood the procedure and after all their questions had been answered. The Ethics Committee of Scuola Internazionale Superiore di Studi Avanzati approved the study.

Stimuli: Linguistic stimuli were pentasyllabic CVCVCVCV words. Thirty-one CV syllables were constructed by pairing one vowel ([a], [e], [i], [o], [u]) and one consonant ([t], [m], [n], [d], [s], [v], [k], [g], [r], [l], [p], [b], [z], [f]). The C-V pairs served to generate different words that were used during the familiarization phase as non-targets. The syllables “so”, “tu”, “ma”, “ve”, “fi” were used exclusively to generate the two target words (i.e. *sotumavefi* presented to half the infants; and *tusomafive* presented to the other half). Each word –either target or non-target– contained 5 different syllables constructed with five different consonants and five different vowels (see appendix 1). All the words were nonsense in Italian. The stimuli were synthesized with the female voice from the MBROLA Italian database IT4 (Dutoit, Pagel, Pierret, Bataille, & Van de Vrecken, 1996). There was coarticulation within syllables because the speech synthesizer was diphone based. The duration of each syllable was 200ms and a monotonous pitch of 240Hz was used. There were no pauses between syllables.

The visual stimuli consisted of one central attractor and puppets presented inside two white squares (one presented at the time). The white squares had a side-length of 8cm, and the distance between them was 13.5 cm. The colorful puppets were used as visual reinforcement. In the Familiarization, one of the puppets was paired with the target word, whereas another two puppets were randomly paired with the non-target words. The pairing of the target word with the puppet was counterbalanced across participants.

Procedure: We adapted the procedure of (Kovács & Mehler, 2009), implemented previously in word-learning tasks with young infants (Hochmann, Benavides-Varela, Nespor, & Mehler, 2011). The experiment consisted of a familiarization phase of 32 trials and a test phase of 8 trials (Figure 1). Trials in the familiarization phase began with the display of a central visual attractor and two white squares, one on the left and one on the right side of the screen. When the infant fixated the central attractor, one word was played. After the offset of the word, the central attractor disappeared, leaving only the two white squares visible for 1s. Then a looming puppet appeared on one of the two white squares for two seconds. The puppet was accompanied by a 300ms tinkling bell presented 800ms after the onset of the visual stimulus. The target word was presented in sixteen familiarization trials. This word predicted the location of the puppet on one side of the screen. The location was counterbalanced across participants. Half of the infants were familiarized with *sotumaveft* and the other half were familiarized with *tusomafive*. In this way, acoustic manipulations on the edge syllables in one group of infants corresponded to the same manipulations on the middle syllables in the other group, and vice versa (see below). Another sixteen words were used as non-targets; these words preceded the puppet's appearance on the other side of the screen. Target and non-target trials were presented in pseudo-random order. We ensured that targets or non-targets were presented at most three times in a row. The test trials were similar to the familiarization trials, except that after each word no puppet was displayed on either side of the screen. Two conditions (see appendix 1) were tested within-subjects. In both conditions test words were constructed with the same five syllables that constituted the target word in the familiarization phase, but the position of two syllables was interchanged. In one condition we switched the position of the first and last syllables (Edge-switch condition), whereas in the other the second and fourth syllables were switched (Internal-switch condition).

For instance, if the target word in the familiarization phase was *sotumavefi*, the test words were *fitumaveso* (Edge-switch condition) and *sovetmatufi* (Internal-switch condition). In this way the only difference between each test word and the target word was the serial order of the syllables. Each test word was presented four times in the test in pseudo-random order. In the first four trials there were two trials per condition. Each trial started two seconds after the word offset.

Data acquisition: Infants' gaze was recorded with an eye-tracker (TOBII 1750). The eye tracker was integrated into a 17-inch TFT screen, where the stimuli were presented via an IMAC 10,1 running PsyScope X software (<http://psy.ck.sissa.it/>). A loudspeaker was placed behind the screen for the presentation of the acoustic stimuli. Infants were seated on the parent's lap at about 50 cm distance from the monitor. A hidden video camera was used to observe the infant's behavior.

Data analysis: We measured the first fixation after infants heard the test word (looking times of 80 ms or more were considered a fixation). We divided the screen into three equal parts: left, middle and right. Infants were coded as looking either to the target side or to the non-target side. We computed normalized difference scores by subtracting the number of trials in which the infant fixated the target and non-target sides, divided by the total number of trials in which the infants fixated either side (71.82% of the total number of trials on average):

$$\frac{(\# \text{trials fixating the target side} - \# \text{trials fixating the non-target side})}{(\# \text{trials fixating the target side} + \# \text{trials fixating the non-target side})}$$

The difference scores were compared to chance level (zero) using two-tailed one-sample *t*-tests. Positive scores indicate that infants searched for the puppet on the target side whereas negative scores indicate that infants searched for the puppet on the non-target side.

Results

The results are presented in Figure 2. Participants looked for the puppet in 68.75% of the total number of trials. The data show that participants' looking, as measured by the first fixation, was significantly above chance on the target side in the Internal-switch condition (mean difference score = .33; $t(21) = 2.123$; $p < .05$; $d' = .45$), and was at chance for the Edge-switch condition (mean difference score = .19; $t(21) = 1.156$; $p = .262$; $d' = .26$). We obtained similar results using longest fixation and overall accuracy as dependent variables. Longest fixation: Internal-switch condition mean difference score= .36; $t(21) = 2.792$; $p = .01$; $d' = .595$; Edge-switch condition mean difference score = -.075; $t(21) = -.439$; $p = .665$; $d' = -.09$. Overall accuracy: Internal-switch condition mean difference score = .23; $t(21) = 1.81$; $p = .08$; $d' = .397$; Edge-switch condition mean difference score =.15; $t(21) = 1.34$; $p = .19$; $d' = .29$. There was no correlation between infants' age and their performance in the Edge-switch condition ($r=.33$; $p=.148$), nor in the Internal-switch condition ($r =.34$; $p =.125$).

Discussion

The results showed that 7 month-olds' representations of the syllables in a multisyllabic word were influenced by the position of the syllables within the sequence. Participants recognized the familiar sequence when the edge-syllables were preserved even if the middle syllables were swapped. However, infants failed to recognize the familiar sequence (i.e. familiarization word) when the first and last syllables were swapped and the remaining syllables were kept in their original positions.

The results thus suggest that the ability to encode the position of the elements located at the edges is already functional at 7-months of age. These results are compatible with a previous study exploring infants' memory for the content of words located at different positions within a

sentence (i.e. utterance-initial, utterance-medial, or utterance-final position). The study reports that 8-month-olds recognized words that appeared either initially or finally in the passages but failed to recognize words from the middle of the utterances (Seidl & Johnson, 2006). Seidl and Johnson's study evidences the prevalence of -the content- of edges in infants' verbal memory. In addition, our study shows that infants are capable of detecting changes in the order of initial and final elements even in the absence of any prosodic cue. Our findings are thus consistent with accounts suggesting that edge positions are more salient and serve as anchors for verbal representations in adults (Endress *et al.*, 2009). Moreover, the present data extend previous work by suggesting that encoding edges may be a primitive mechanism supporting verbal representations also during development.

The results of the present study are also consistent with the traditional literature on sequential learning. In fact, serial position effects (i.e. primacy and recency) predict a better memorization of the first or the last syllables in a multisyllabic sequence. However, a full attribution of our results to traditional theories of sequential learning would require a demonstration of a u-shaped curve, whereby not only the first and last items are better remembered, but middle syllables are also remembered as a function of their position in the sequence; namely, the closer an item is to the middle the more difficult it is to represent it.

At this point in our research, it is thus difficult to decide whether our results can best be explained by serial learning effects or by an edge primacy primitive. It seems clear, however, that core properties of memory influence verbal representations in young infants, despite the availability of other types of information in the signal that could have facilitated the encoding of middle syllables. For instance, the participants may have used the statistical information to track the sequence of syllables that constitute the familiarization word. Pioneer studies on artificial

speech indicate that 8-month-old infants successfully use transitional probabilities (i.e., the probability with which one syllable predicts the next one) to find words in a continuous speech stream (Saffran *et al.*, 1996). Thus, our participants may have also used transitional probabilities to “connect” the sequence of syllables and encode them in memory -assuming that each syllable serves as a probabilistic retrieval cue for the next one-. However, in the current study infants exhibited the trained response (looking for the puppet on the target side) in the Internal-Switch condition, in spite of the fact that all transitional probabilities between syllables had been altered with respect to the familiarization word. Furthermore, infants did not manifest the trained response in the Edges-Switch condition although two transitional probabilities between syllables were identical to those in the familiarization word. This suggests that participants did not rely on the statistical information to identify the familiar sequence, but were influenced by primitive mnemonic biases. One could argue that in the present experiment each word was presented in isolation, thus no segmentation task was required to identify the sound of a word (as it is generally the case in artificial language learning studies assessing infants’ statistical analysis of the input). This would imply, however, that statistical learning is active during segmentation but not during encoding or recognition of the speech information required for the full word learning process.

The present results clearly indicate the importance of edges in infants’ representations of multisyllabic words; still, the material and manipulations we used prevent additional conclusions. First, it is difficult to decide whether the representation of both edges occurred in the same way, or rather one of the two syllables was better represented than the other. In fact our results could be attributed to the infants noticing that the first, last, or both the first and the last syllables were different from the familiarization word. Previous studies on children’s

phonological knowledge of word-forms and early lexical access may provide some indications on how to evaluate the representation of each edge. For instance, a previous study with 10 months-old Dutch-learning infants showed discrimination of consonantal contrasts in onset positions, but no sensitivity to the same contrasts presented in coda positions (Zamuner, 2006). Similarly, another study on the time course of word recognition in 2 year-olds showed that participants interpreted familiar words incrementally, namely using the information available from the onset to guide their looking towards the picture corresponding to that word (Swingley, Pinto & Fernald, 1999). These studies suggest an asymmetry between onset and coda positions to the advantage of onsets, at least as far as the phonetic specification of known words is concerned. However, more recent studies that have directly contrasted the representation of initial and final segments of a word, could lead to a different conclusion. For instance, Nazzi and Bertoncini (2009) evaluated 20-month-olds' ability to learn two words that differed by only one consonant in either onset or coda position. Infants succeeded for both positions, ruling out the possibility that only syllable-onset positions are specified. Likewise, Swingley (2009) examined 1.5 year-old children's knowledge of the phonological forms of familiar words by measuring their comprehension of correctly and incorrectly pronounced onset or coda consonants. The study showed that participants spotted mispronunciations in both the word's initial or final consonant.

In view of the mixed evidence coming from studies on phonetic specificity in early lexical acquisition, a separate evaluation of infants' positional memory for the initial and final syllables in a multisyllabic word would be justified. Such question, however, goes beyond the scope of the present study, which focuses on the significance of edges in general by contrasting the representation of internal versus external parts of the words.

Another limitation of the present study is the reduced number of phonemes used to create the stimuli (/s/, /v/, /f/, /t/, /o/, /u/, /e/, /i/). Even though the specific features of these phonemes cannot explain our results (because the same manipulations to the stimuli were implemented in the Edge-switch and Internal-switch condition; see methods), the extent to which 7-month-olds display similar effects with other consonants and vowels is yet to be determined. Moreover, we cannot confirm the level at which participants computed similarities or differences between the familiarization and the test words. In the present study we manipulated the syllables because they are fundamental units of speech perception (Bertoni, Floccia, Nazzi, & Mehler, 1995; Goyet, de Schonen, & Nazzi, 2010). However, phonemes are also crucial for learning the lexicon because they are the minimal units for distinguishing meaning (together with cues such as lexical stress or tones in some languages). Judging from previous investigations on early word learning and infants' retention of phonological details (e.g., Swingley, 2005; Nazzi, Floccia, Moquet, & Butler, 2009; Nazzi & Bertoni, 2009), it seems possible that infants represent words by focusing on some of the syllable constituents. In fact a number of studies have shown that French (Havy & Nazzi, 2009; Nazzi, 2005) and Italian infants (Hochmann *et al.*, 2011) represent consonants better than vowels in word learning situations (but see also Floccia, Nazzi, Delle Luche, Poltrack, & Goslin, 2013; and Mani & Plunkett, 2007; 2008 for different results in English). The distinction between consonants and vowels in lexical representations is observed for word-initial, word-final, and word-internal consonants, at least in monosyllabic and bisyllabic words. However, previous studies only tested infants from 12 months of age on. It is thus unclear whether 7-month-olds also focus on consonants to represent multisyllabic words (see also Benavides-Varela *et al.*, 2012 showing that newborns, as opposed to older infants, encode better the sound of vowels). At any rate, the critical finding in this study is that 7 month-old infants

appear to rely mostly on the boundaries of a sequence to represent new multisyllabic words. Which of vowels or consonants is privileged in this context must be the object of further studies.

At this point it is worth asking: if edges play such an important role for word representations, what is the part played by middle syllables? Do infants encode the identity of the middle syllables but not their exact sequential order? Or do infants disregard middle syllables due to memory limitations? These two processes –remembering the occurrence of an item, and remembering where in a sequence the item occurred – could be independent (Henson, 1998). In fact, whereas the brain structures that mediate the representation of temporal order may not be functionally mature until the end of the first year of life (Nelson, 1995), recognition of simple syllables and CVCV sounds, seems to be available since birth (Benavides-Varela *et al.*, 2011; Benavides-Varela *et al.*, 2012; Partanen *et al.*, 2013).

In the following experiment we ask whether the results from experiment 1 can be attributed to the fact that infants cannot encode any information about internal syllables or merely information related to the sequential order of those syllables. We tested infants with words that differed from the familiarization word in either the order or the identity of the middle syllables. If participants cannot encode the middle syllables, any change in those syllables – either item change or sequence change— should go unnoticed. If infants are able to encode the sound of all syllables but not their sequential order, participants should differentiate the target word from a word that contains a completely new set of internal syllables.

Experiment 2

Methods

Participants: Twenty infants (12 males; mean age = 7 months 15 days; age range = 6 months 20 days to 8 months 6 days) were included in the analysis. Additional infants were excluded because of crying or fussiness ($N = 3$), because they looked to the screen only in one test trial ($N = 1$), or because there were technical problems ($N = 3$).

Stimuli: The stimuli were the same as in the previous experiment. Moreover, in this experiment we made sure that the syllables “zi”, “bu”, “ne”, “lo” never appeared during familiarization. They were used only in the test phase to replace the second and fourth syllables in the Internal-change condition (see appendix 1).

Procedure: The procedure was identical to the one used in experiment 1 except that the test words were different. In one condition, the position of the second and fourth syllables in the test word differed from the position of the syllables in the familiarization word (Internal-switch condition). In the other condition it was not the position of the second and fourth syllables that were different but rather their identity (Internal-change condition). For instance, if the familiarization word was *sotumavefi*, the test words were *sovematuſi* (Internal-switch condition) and *sozimaloſi* (Internal-change condition).

Results

The results are presented in figure 3. Infants looked for the puppet in 71.05% of the total number of trials. Participants’ looking was significantly above chance on the target side in the Internal-switch condition (mean difference score = .38; $t(19) = 2.364$; $p < .05$; $d' = .54$) but they

did not perform significantly above chance in the Internal-change condition (mean difference score = .075; $t(19) = .462$; $p = .649$; $d' = .103$). There was no correlation between infants' age and their performance in the Internal-switch condition ($r = .08$; $p = .709$), nor in the Internal-change condition ($r = -.03$; $p = .885$).

Discussion

The results indicate that infants only searched for the puppet on the side of the screen previously associated with the target word when the middle syllables swapped positions but not when they were replaced with novel ones. This suggests that participants encoded the general identity of the internal syllables. They were not able, however, to track their precise sequential order.

The present results are in agreement with a number of adult studies showing that order and item information are two separate components in memory (e.g., Henson 1998, 1999; Marshuetz, 2005; Ng & Maybery 2002). In addition, our data extends the findings in adults by suggesting that differences in the ability to retain order and identity also influence the representation of verbal information in development.

Uncovering these learning constraints may have implications for theories of speech processing and early memory. For instance, as noted earlier, 8-month-olds are able to track the transitional probabilities between syllables to extract words. This appears incompatible with our present results, which demonstrate that memory for the temporal order of internal sounds is not reliable. As a result, it is legitimate to ask whether sound order representation depends on the nature of the task (i.e. segmenting vs remembering), the length of the sequences, or even the

specific characteristics of the paradigm, which may pose additional cognitive demands on the infants.

In fact, before additional conclusions can be derived from the present data, it should be acknowledged that the results obtained in experiments 1 and 2 could be attributed to the paradigm used in our studies. Infants may have encoded the order and content of all syllables of the familiarization word accurately but had trouble associating two different test sequences with the same target puppet or side of the screen. Indeed, previous studies on interleaving learning of regularities using a similar paradigm (Kovács & Mehler, 2009) have shown that most monolingual infants tend to learn and associate only one regularity with the corresponding side of the screen, even though two regularities are provided.

In the preceding experiments, infants were exposed to only one consistent word-referent association during familiarization but during the test they heard two very similar sounds. As a result, recognition of only one word may be due to the infants' inability to cope with two sounds simultaneously rather than the actual format of representation of the sequences in memory. We explore this possibility in the following experiment in which infants are presented with two test words: the Internal-switch word (in which the second and fourth syllables are interchanged), and the Target-word i.e. the word presented during the familiarization phase. Based on the results of the previous experiments, we expected infants to recognize the Internal-switch word. We also expected them to recognize the Target-word in the test phase, given that they had learned this word and its association with the puppet during the familiarization phase. If infants are able to associate two similar pentasyllabic words with the corresponding object or side of the screen, they should perform above chance in the two conditions. Alternatively, if the paradigm

constrains performance in the test, participants should perform above chance in only one of the conditions.

Experiment 3

Methods

Participants: Twenty infants (11 males) aged from 7 months 0 days to 8 months 6 days (mean age 7 months 19 days) participated in the study. All were full term infants, with no birth complications, and no reported hearing or visual problems. Additional infants were excluded from the analysis because of crying or fussiness ($N = 5$), because they looked to the screen only in one test trial ($N = 2$), or because there were technical problems ($N = 1$).

Stimuli: The auditory and visual stimuli were identical to those used in Experiment 1.

Procedure: The procedure was identical to the one used in experiment 1 except that the test words were different. In one condition, the position of the second and fourth syllables in the test word differed from the position of the syllables occupied in the familiarization word (Internal-switch condition). In the other condition the test word was identical to the target word learned during the familiarization phase (Target condition). If the target word in the familiarization phase was *sotumavefi*, the test words were *sotumavefi* (Target condition) and *sovematufi* (Internal-switch condition) (see appendix 1).

Results

Participants looked for the puppet in 75.66% of the total number of trials. As in Experiments 1 and 2, we found that infants looked significantly above chance to the target side in the Internal-switch condition (mean difference score = .33; $t(19) = 2.297$; $p < .05$; $d' = .51$). In addition, their looking was also significantly above chance to the target side in the Target

condition (mean difference score = .36; $t(19) = 2.462$; $p < .05$; $d' = .55$) (see figure 4). There was no correlation between infants' age and their performance in the Internal-switch condition ($r = .25$; $p = .28$), nor in the Target condition ($r = .13$; $p = .595$).

Discussion

The current experiment replicates the main results of Experiments 1 and 2: participants searched for the puppet on the target side in the Internal-switch condition. Furthermore, the infants successfully recognized the familiarization word and associated it with the corresponding side of the screen in the test phase. Given that the infants displayed similar looking behavior in both the Internal-switch condition and the Target condition, the results ruled out the possibility that the data from experiments 1 and 2 had to do with the specific features of the paradigm. Instead, the data confirmed our initial interpretation: speech representation in early word-learning tasks is modulated by the position of the syllables within a word.

General Discussion

The present study explored 7-month-olds' abilities to remember the identity and the order of syllables that constitute a non-stressed pentasyllabic word. Infants were first familiarized with a word associated to a puppet appearing on one side of the screen. Thereafter, in the test, infants were presented with either the target item or modifications of a) the order or b) identity of the syllables that formed the familiarization word. The results showed that infants successfully searched for the puppet on the correct side of the screen when they heard the target pentasyllabic sound in the test (Experiment 3). They also looked to the target side when presented with sequences that differed from the familiarization word in the order of middle syllables

(Experiments 1, 2 and 3). In contrast, participants did not look to the target side when there were alterations in the order of edge-syllables (Experiment 1), or changes in the identity of the middle syllables (Experiment 2). The data thus indicate that 7-month-olds are able to encode precise information concerning the order and identity of edge syllables but have trouble encoding the order of syllables located in the middle of words.

These results add significantly to the existing literature on sequential learning during development, and provide evidence that infants like adults (Endress *et al.*, 2009), are influenced by primitive mnemonic biases, which favor the recognition of elements located at the edges. Moreover, this is the first study to show that information about order and identity are encoded differently in infants' memory representations.

One might argue that these results can be attributed to infants' incapacity to encode or even process a long multisyllabic word rather than to primitive cognitive biases favoring edge information. In fact, information regarding the actual size of infants' verbal memory span is so sparse that one could assume 7-month-olds remembering only the first phoneme or syllable of the non-stressed sequence. This possibility would have been enough to explain why the replacement of the first syllable in Experiment 1, prevented participants from recognizing the familiarization word. The results of Experiment 2, however, call for a different explanation. The data showed clearly that the identity of the internal syllables was also relevant for the recognition of the familiarization word. Infants appeared to identify the specific set of syllables that form the pentasyllabic word but failed to recognize the order of internal syllables. Critically, in the two manipulations of Experiment 2 the edge syllables remained in their original positions, but the infants only recognized the familiar sequence in one of these conditions. It thus seems plausible that the edge syllables were necessary, yet not sufficient for the recognition of the familiar

sequence.

Paralleling the adult studies on positional memory, the present findings show that infants remember the content of the syllables that form a pentasyllabic word but have trouble remembering the position of internal syllables. Our results thus indicate that the division between the two components of sequential memory (order and item information) appears early in development, possibly reflecting a fundamental structure of memory representations. In particular, it seems that recognition of verbal content is already functional in young infants whereas detail recognition of serial order appears later in development, possibly with the maturation of the brain structures that mediate the representation of temporal order (Nelson, 1995).

In the present study, we found a relative prominence of edges when the verbal material contained basic phonological information without prosody. However, we do not exclude the possibility that in everyday life infants use other information contained in the speech signal to encode internal elements. In fact previous studies showed that infants learn the order of internal elements more easily in the presence of melody (Thiessen & Saffran, 2009) or coherent prosodic sentential units (Mandel *et al.*, 1996). This suggests that memory processes and cues from the speech signal interact bidirectionally during language acquisition. While fundamental mechanisms of memory predispose the system to encode certain information (i.e. content and position of the edges), extra cues in speech modulate the ease with which other portions (i.e. order of middle elements) are encoded. One could hypothesize that lexical stress at the word level –like prosody or melody at the utterance level– facilitates learning the internal portions of a word, either by increasing the saliency of certain syllables or by segmenting longer speech sequences in chunks (see Moher, Tuerk & Feigenson, 2012 for an example of how infants use

chunking to increase the visual memory). Previous studies have shown that some implementation of lexical stress (Bion *et al.*, 2010; Johnson & Jucszyk, 2001) as well as coarticulation (Johnson & Jucszyk, 2001) may facilitate the segmentation and representation of speech streams in infancy. Exploring whether these or similar implementations influence the way infants encode sequential information may be an important avenue for future research.

The basic features of verbal memory in infancy, and in particular the proneness to pay attention to the beginnings and endings of utterances may also relate to the emergence of edge-based positional regularities in the language (Endress *et al.*, 2005; 2009) and may facilitate its acquisition. For example morphemes (e.g. suffixes or prefixes) are generally appended either to the final edge of a word (as to signal the past-tense in “regard-ed”) or to the first edge of a word (as in “dis-regard”), but not to middle; in fact prefixes and suffixes are cross-linguistically much more frequent than infixes (Greenberg, 1957). Reduplication provides another example of the importance of edges in morphology. Across the languages of the world, reduplications are widespread either in initial or in final position. Middle reduplications, by contrast, are rarely attested (Broselow & McCarthy, 1983). Extracting and learning the particular stress pattern in a given language may be also facilitated by this fundamental edge-based mechanism. Generalizing the position of stress in a given language requires both the ability to perceive stress and the ability to remember where stress is placed in most words of that language. Young infants easily perceive and process stress (Nazzi, Floccia, & Bertoni, 1998), however they may have difficulty in retaining positional information (see above). Still, the task of remembering the position of stress might be facilitated across languages, possibly because stress is most frequently found in word-initial or word-final positions (Hayes, 1995).

Although the findings of the present experiment resemble the properties of sequential

memory in adults, it is difficult to claim that the mechanisms supporting phonological memories during development and adulthood are the same. Extant theories of adult short-term memory propose that verbal memory traces – necessary for learning new words – fade away within a few seconds unless an articulatory rehearsal process prevents them from decaying (Baddeley, 2003). These models however, do not make predictions regarding the way preverbal infants maintain new phonological representations in short-term memory, for instance when learning multisyllabic words. Thus, an interesting area for speculation constitutes the degree to which articulatory rehearsal processes are active in young infants. One possibility is that, internal speech-motor programs favor a rehearsal-like mechanism that facilitates the storage of verbal information in preverbal infants. In fact, it has been shown that the links between auditory and motor representations exist in infancy. Preverbal infants are able to match simple speech sounds (i.e. syllables or vowels) and mouth movements before they can utter their first words (Bristow, *et al.*, 2009; Yeung & Werker, 2013). Moreover, De Paolis and colleagues have shown that production patterns in prelinguistic infants (babbling specific features) influence infants' preferential looking behavior to non-words containing these features (DePaolis, Vihman, & Keren-Portnoy, 2011; DePaolis, Vihman, & Nakai, 2013; Majorano, Vihman, & DePaolis, 2014). These early links between auditory and motor representations indicate that basic articulatory mechanisms may support at least some verbal memory representations in preverbal infants. The alternative would be to hypothesize that, different from adults, 7-month-olds do not rely on articulation to maintain phonological information in memory. This alternative finds support on the fact that, even though some infants begin babbling at 7-months of age, their speech motor control is still immature (Steeve, Moore, Green, Reilly, & McMurtrey, 2008) particularly if required for rehearsing multisyllabic words. This would suggest that only later in

development, when infants begin producing entire multisyllabic words, the temporary storage of verbal information rely on an articulatory mechanism similar to that in adults. The acceptance of one or the other alternative might be premature at present. Thus, the exact way preverbal infants maintain the traces of the phonological representation when recruited for working memory tasks (for instance word learning) awaits clarification. At any rate, the critical finding of this study is that the mechanism by which syllables located at the edges are privileged over middle syllables is present before the onset of multisyllable production. These findings provide new insights for understanding the mechanisms of verbal memory in development, and open new vistas for future investigation.

Based on these results and the material used in the present study, the functioning of the positional memory system underlying multisyllabic word recognition at 7-months could be anticipated. The representational system seems to detect previously heard sounds by verifying first the accurate content and position of the edges. Then, the system appears to validate the presence of the middle syllables, disregarding the order of these elements. This implies that, when it comes to word recognition in development, content would be a necessary requisite for word-sound recognition, while order –of middle syllables– might be taken into consideration separately and perhaps later in development. Differently from other syllables within the word, the positional information of the edges would be stored accurately at this age.

To conclude, these data suggest that the first stages of word learning are influenced by mnemonic biases. Infants' representational system encodes edges accurately, whereas memory for middle elements appears fallible for temporal order. We argue that the relative advantage of syllables located at the edges should not be considered a limitation for lexical learning, but rather another feature of human memory that interacts with the language acquisition device. It is

possible that these edge-based regularities appeal to a positional memory mechanism that is not specific to the human species or to the linguistic knowledge, yet it appears to influence humans' linguistic representations early in life. The elements at the edges seem to be learned particularly easily, explaining why – arguably – edge-based regularities are frequently found in human languages.

Acknowledgements:

We thank Alessio Isaja and Marijana Sjekloca for their technical and administrative support, Laurence White for his helpful comments during the preparation of these experiments and Francesca Gandolfo for her valuable assistance on the recruitment of babies. Our gratitude also goes to the parents of the infants for their participation. To Judit Gervain and Susana Frank for helpful comments on previous versions of this paper. The research leading to these results received funding from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013) / ERC grant agreement n° 269502 (PASCAL) to J.M.

References

- Baddeley, A. D. (2003). Working memory: looking back and looking forward. *Nature Reviews Neuroscience*, 4, 829-839. doi:10.1038/nrn1201
- Bauer, P. J., & Mandler, J. M. (1992). Putting the horse before the cart: The use of temporal order in recall of events by one-year-old children. *Developmental Psychology*, 28, 441-452. doi:10.1037/0012-1649.28.3.441
- Benavides-Varela, S., Gómez, D. M., Macagno, F., Bion, R. A., Peretz, I., & Mehler, J. (2011). Memory in the Neonate Brain. *PlosOne*, 6, e27497. doi:10.1371/journal.pone.0027497
- Benavides-Varela, S., Hochmann, J.-R., Macagno, F., Nespor, M., & Mehler, J. (2012). Newborn's brain activity signals the origin of word memories. *Proceedings of the National Academy of Sciences USA*, 109, 17908-17913. doi: 10.1073/pnas.1205413109
- Bergelson, E., & Swingley, D. (2012). At 6–9 months, human infants know the meanings of many common nouns. *Proceedings of the National Academy of Sciences USA*, 109, 3253-3258. doi:10.1073/pnas.1113380109
- Bertoni, J., Floccia, C., Nazzi, T., & Mehler, J. (1995). Morae and Syllables: Rhythmic Basis of Speech Representations in Neonates. *Language and Speech*, 38, 311-329.
- Bion, R. A., Benavides-Varela, S., & Nespor, M. (2011). Acoustic Markers of Prominence Influence Infants' and Adults' Segmentation of Speech Sequences. *Language and Speech*, 54, 123-140. doi:10.1177/0023830910388018
- Bjork, E. L., & Healy, A. F. (1974). Short-term order and item retention. *Journal of Verbal Learning and Behavior*, 13, 80-97. doi: 10.1016/s0022-5371(74)80033-2

- Bortfeld, H., Morgan, J. L., Golinkoff, R. M., & Rathbun, K. (2005). Mommy and Me: familiar names help launch babies into speech-stream segmentation. *Psychological Science*, 16, 298-304. doi: 10.1111/j.0956-7976.2005.01531.x
- Bristow, D., Dehaene-Lambertz, G., Mattout, J., Soares, C., Gliga, T., Baillet, S., et al. (2009). Hearing Faces: How the Infant Brain Matches the Face It Sees with the Speech It Hears. *Journal of Cognitive Neuroscience*, 21, 905-921. doi: 10.1162/jocn.2009.21076
- Broselow, E. & McCarthy, J. (1983) A theory of internal reduplication. *The Linguistic Review*. 3, 25–88. doi: 10.1515/tlir.1983.3.1.25
- Chambers, K. E., Onishi, K. H., & Fisher, C. (2003). Infants learn phonotactic regularities from brief auditory experience. *Cognition*, 87, B69–B77. doi: 10.1016/s0010-0277(02)00233-0
- DePaolis, R. A., Vihman, M. M., & Keren-Portnoy, T. (2011). Do production patterns influence the processing of speech in prelinguistic infants. *Infant Behavior and Development*, 34, 590-601. doi: 10.1016/j.infbeh.2011.06.005
- DePaolis, R. A., Vihman, M. M., & Nakai, S. (2013). The influence of babbling patterns on the processing of speech. *Infant Behavior and Development*, 36, 642–649. doi: 10.1016/j.infbeh.2013.06.007
- Dutoit, T., Pagel, V., Pierret, N., Bataille, F., & Van de Vrecken, O. (1996). The MBROLA project: towards a set of high quality speech synthesizers free of use for non commercial purposes . *Fourth International Conference on Spoken Language, 1996. Proceedings* . 3, pp. 1393-1396. ICSLP. doi: 10.1109/icslp.1996.607874
- Endress, A. D., & Bonatti, L. L. (2007). Rapid learning of syllable classes from a perceptually continuous speech stream. *Cognition*, 105, 247–299. doi: 10.1016/j.cognition.2006.09.010

- Endress, A. D., & Mehler, J. (2010). Perceptual constraints in phonotactic learning. *Journal of Experimental Psychology: Human Perception and Performance*, 36, 235-250. doi: 10.1037/a0017164
- Endress, A. D., & Mehler, J. (2009). Primitive computations in speech processing. *The Quarterly Journal of Experimental Psychology*, 62, 2187-2209. doi: 10.1080/17470210902783646
- Endress, A., Nespor, M., & Mehler, J. (2009). Perceptual and memory constraints on language acquisition. *Trends in Cognitive Sciences*, 13, 348-353. doi: 10.1016/j.tics.2009.05.005
- Endress, A., Scholl, B. J., & Mehler, J. (2005). The Role of Salience in the Extraction of Algebraic Rules. *Journal of Experimental Psychology: General*, 134, 406-419.
- Floccia, C., Nazzi, T., Delle Luche, C., Poltrock, S., & Goslin, J. (2013). English-learning one-to two-year-olds do not show a consonant bias in word learning. *Journal of Child Language*. doi: 10.1017/s0305000913000287
- Friedrich, M., & Friederici, A. D. (2011). Word learning in 6-month-olds: Fast encoding–weak retention. *Journal of Cognitive Neuroscience* (23), 3228-3240. doi: 10.1162/jocn_a_00002
- Goyet, L., de Schonen, S., & Nazzi, T. (2010). Words and syllables in fluent speech segmentation by French-learning infants: An ERP study. *Brain Research*, 1332, 75-89. doi: 10.1016/j.brainres.2010.03.047
- Greenberg, J. H. (1957). *Essays in Linguistics*. Chicago, IL: University of Chicago Press.
- Gulya, M., Rovee-Collier, C., Galluccio, L., & Wilk, A. (1998). Memory Processing of a Serial List by Young Infants. *Psychological Science*, 9, 303-307. doi: 10.1111/1467-9280.00060
- Gupta, P., Lipinski, J., Bradon, A., & Po-Han, L. (2005). Serial position effects in nonword repetition. *Journal of Memory and Language*, 53, 141-162. doi: 10.1016/j.jml.2004.12.002

Hallé, P. A., & Boysson-Bardies, B. (1994). Emergence of an early receptive lexicon: Infants' recognition of words. *Infant Behavior and Development*, 17, 119–129 doi: 10.1016/0163-6383(94)90047-7

Hallé, P. A., & Boysson-Bardies, B. (1996). The format of representation of recognized words in the infants' early receptive lexicon. *Infant Behavior and Development*, 19, 463 – 481 doi: 10.1016/s0163-6383(96)90007-7

Hayv, M., & Nazzi, T. (2009). Better processing of consonantal over vocalic information in word learning at 16 months of age. *Infancy*, 14, 439-456. doi: 10.1080/15250000902996532

Hayes, B. (1995). *Metrical stress theory: principles and case studies*. Chicago: University of Chicago Press.

Henson, R. N. (1998). Short-Term Memory for Serial Order: The Start-End Model. *Cognitive Psychology*, 36, 73-137. doi: 10.1006/cogp.1998.0685

Henson, R. (1999). Positional information in short-term memory: relative or absolute? . *Memory and Cognition*, 27, 915-927. doi: 10.3758/bf03198544

Henson, R., Hartley, T., Burgess, N., Hitch, G., & Flude, B. (2003). Selective interference with verbal short-term memory for serial order information: A new paradigm and tests of a timing-signal hypothesis. *The Quarterly Journal of Experimental Psychology A*, 56, 1307-1334 doi: 10.1080/02724980244000747

Hochmann, J.-R., Benavides-Varela, S., Nespor, M., & Mehler, J. (2011). Consonants and vowels: different roles in early language acquisition. *Developmental Science*, 14 (6), 1445-1458. doi: 10.1111/j.1467-7687.2011.01089.x

Horne, P., Barr, R. G., Valiante, G., Zelazo, P. R. & Young, S. N. (2006). Glucose enhances newborn memory for spoken words. *Developmental Psychobiology*, 48: 574–582.

doi: 10.1002/dev.20172

Johnson, E.K. & Jusczyk, P.W. (2001). Word segmentation by 8-month-olds: when speech cues count more than statistics. *Journal of Memory and Language*. 44, 548–567 doi:

10.1006/jmla.2000.2755

Jusczyk, P. W., Hohne, E. A., & Bauman, A. (1999). Infants' sensitivity to allophonic cues for word segmentation . *Perception & Psychophysics* , 61, 1465-1476. doi: 10.3758/bf03213111

Jusczyk, P., & Hohne, E. A. (1997). Infant's memory for spoken words. *Science* , 277, 1984-1986. doi: 10.1126/science.277.5334.1984

Kooijman, V., Junge, C., Johnson, E. K., Hagoort, P., & Cutler, A. (2013). Predictive brain signals of linguistic development. *Frontiers in Psychology* , 4. doi: 10.3389/fpsyg.2013.00025

Kovács, A. M., & Mehler, J. (2009). Flexible Learning of Multiple Speech Structures in Bilingual Infants. *Science* , 325, 611-612. doi: 10.1126/science.1173947

Leslie, A. M., & Keeble, S. (1987). Do six-month-old infants perceive causality? *Cognition*, 25, 265-288. doi: 10.1016/s0010-0277(87)80006-9

Majorano, M., Vihman, M., & DePaolis, R. A. (2014). The relationship between infant production experience and their processing of speech. *Language Learning and Development*, 10, 179-204 doi: 10.1080/15475441.2013.829740

Mandel, D. R., Jusczyk, P. W., & Pisoni, D. B. (1995). Infants' Recognition of the Sound Patterns of Their Own Names. *Psychological Science* , 6, 314-317. doi : 10.1111/j.1467-9280.1995.tb00517.x

Mandel, D. R., Kemler-Nelson, D. G., & Jusczyk, P. W. (1996). Infants remember the order of words in a spoken sentence. *Cognitive Development* , 11, 181-196 doi: 10.1016/s0885-2014(96)90002-7

- Marshuetz, C. (2005). Order information in working memory: an integrative review of evidence from brain and behavior. *Psychological Bulletin*, 131, 323-339 doi: 10.1037/0033-2909.131.3.323
- Mattys, S. L., & Jusczyk, P. W. (2000). Phonotactic cues for segmentation of fluent speech by infants. *Cognition*, 78, 91-121 doi: 10.1016/s0010-0277(00)00109-8
- Moher, M., Tuerk, A.S. & Feigenson, L. (2012). Seven-month-old infants chunk items in memory. *Journal of Experimental Child Psychology*. 112, 361-377. doi: 10.1016/j.jecp.2012.03.007
- Nazzi, T. (2005). Use of phonetic specificity during the acquisition of new words: differences between consonants and vowels. *Cognition*, 98, 13-30 doi: 10.1016/j.cognition.2004.10.005
- Nazzi, T., & Bertoni, J. (2009). Phonetic specificity in early lexical acquisition: new evidence from consonants in coda positions. *Language and Speech*, 52, 463-480. doi:10.1177/0023830909336584
- Nazzi, T., Floccia, C., & Bertoni, J. (1998). Discrimination of pitch contours by neonates. *Infant Behavior and Development*, 21, 779-784. doi: 10.1016/s0163-6383(98)90044-3
- 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. doi: 10.1016/j.jecp.2008.05.003
- Nelson, C. A. (1995). The ontogeny of human memory: A cognitive neuroscience perspective. *Developmental Psychology*, 5, 723-738. doi: 10.1037//0012-1649.31.5.723

Ng, H. L., & Maybery, M. T. (2002). Grouping in short-term verbal memory: is position coded temporally? *Quarterly Journal of Experimental Psychology A*, 55, 391-424. doi: 10.1080/02724980143000343

Onishi, K. H., Chambers, K. E., & Fisher, C. (2002). Learning phonotactic constraints from brief auditory experience. *Cognition*, 83, B13–B23. doi: 10.1016/s0010-0277(01)00165-2

Page, M. P., & Norris, D. (1998). The primacy model: A new model of immediate serial recall. *Psychological Review*, 105, 761-781. doi: 10.1037/0033-295x.105.4.761-781

Parise, E., & Csibra, G. (2012). Electrophysiological Evidence for the Understanding of Maternal Speech by 9-Month-Old Infants. *Psychological Science*, 23, 728-733 doi : 10.1177/0956797612438734

Partanen, E., Kujala, T., Naatanen, R., Liitola, A., Sambeth, A., & Huotilainen, M. (2013). Learning-induced neural plasticity of speech processing before birth. *Proceedings of the National Academy of Sciences, USA*, 110, 15145-15150 doi: 10.1073/pnas.1302159110

Saffran, J., Aslin, R., & Newport, E. (1996). Statistical Learning by 8- Month-Old Infants. *Science*, 274, 1926-1928. doi: /10.1126/science.274.5294.1926

Seidl, A., & Johnson, E. K. (2006). Infant word segmentation revisited: edge alignment facilitates target extraction. *Developmental Science*, 9, 565-573. doi: 10.1111/j.1467-7687.2006.00534.x

Shukla, M., White, K. S., & Aslin, R. N. (2011). Prosody guides the rapid mapping of auditory word forms onto visual objects in 6-mo-old infants . *Proceedings of the National Academy of Sciences USA*, 108, 6038 -6043. doi: 10.1073/pnas.1017617108

- Steeve, R. W., Moore, C. A., Green, J. R., Reilly, K. I., & McMurtrey, J. R. (2008). Babbling, chewing, and sucking: Oromandibular coordination at 9 months. *Journal of Speech, Language, and Hearing Research*, 51, 1390–1404. doi: 10.1044/1092-4388(2008/07-0046)
- Swain, I. U., Zelazo, P. R., & Clifton, R. K. (1993). Newborn infants' memory for speech sounds retained over 24 hours. *Developmental Psychology*, 29, 312-323. doi: 10.1037//0012-1649.29.2.312
- Swingley, D. (2005). 11-month-olds' knowledge of how familiar words sound. *Developmental Science*, 8, 432-443. doi: 10.1111/j.1467-7687.2005.00432.x
- Swingley, D. (2009). Onsets and codas in 1.5-year-olds' word recognition. *Journal of Memory and Language*, 60, 252-269. doi: 10.1016/j.jml.2008.11.003
- Swingley, D., Pinto, J. P., & Fernald, A. (1999). Continuous processing in word recognition at 24 months. *Cognition*, 71, 73–108 doi: 10.1016/s0010-0277(99)00021-9
- Thiessen, E. D., & Saffran, J. R. (2009). How the Melody Facilitates the Message and Vice Versa in Infant Learning and Memory. *Annals of the New York Academy of Sciences*. 1169, 225–233. doi: 10.1111/j.1749-6632.2009.04547.x
- Tincoff, R., & Jusczyk, P. W. (1999). Some Beginnings of Word Comprehension in 6-Month-Olds. *Psychological Science*, 10, 172-175. doi: 10.1111/1467-9280.00127
- Valiante, A. G., Barr, R. G., Zelazo, P. R., Papageorgiou, A. N., & Young, S. N. (2006). A typical Feeding Enhances Memory for Spoken Words in Healthy 2- to 3-Day-Old Newborns. *Pediatrics*. 117, e476-e486. doi: 10.1542/peds.2004-2859
- Vihman, M. M., Nakai, S., DePaolis, R. A., & Hallé, P. (2004). The role of accentual pattern in early lexical representation. *Journal of Memory and Language*, 50, 336-353 doi: 10.1016/j.jml.2003.11.004

- Wojcik, E. H. (2013). Remembering new words: integrating early memory development into word learning. *Frontiers in Psychology*, 4. doi: 10.3389/fpsyg.2013.00151
- Yeung, H. H., & Werker, J. F. (2013). Lip Movements Affect Infants' Audiovisual Speech Perception. *Psychological Science*, 24, 603-612. doi: 10.1177/0956797612458802
- Zamuner, T. S. (2006). Sensitivity to word-final phonotactics in 9-to-16-month-old infants. *Infancy*, 10, 77–95. doi: 10.1207/s15327078in1001_5

FIGURE CAPTIONS

Figure 1. Schematic diagram of the procedure. The experiments consisted of a familiarization phase (32 trials) and a test phase (8 trials). In the familiarization phase, there were 16 target trials and 16 non-target trials randomly interleaved. Trials started with a fixation display showing a central visual attractor. In each trial, participants listened to one word (either target or non-target). After the offset of the word, the anticipatory period (duration, 1s) began. At the end of the anticipatory period a puppet appeared for 2s on one side of the screen. In the test phase no puppet appeared after each test word.

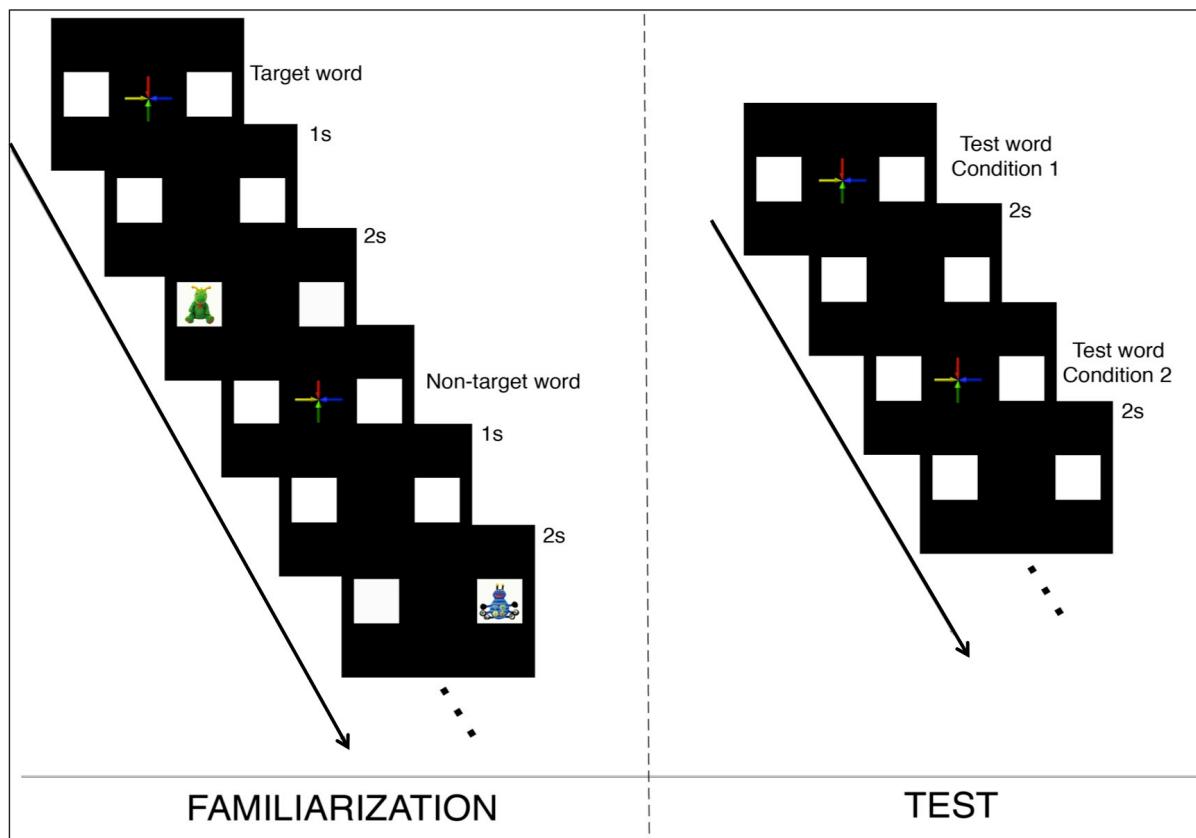


Figure 2. Normalized difference scores considering first fixations in the test phase of experiment

1. The y-axis shows the mean difference scores. Positive difference scores indicate that infants searched for the puppet on the target side while negative difference scores indicate that infants searched for the puppet on the non-target side. Bars indicate standard error of the mean (SEM).

* $p < 0.05$.

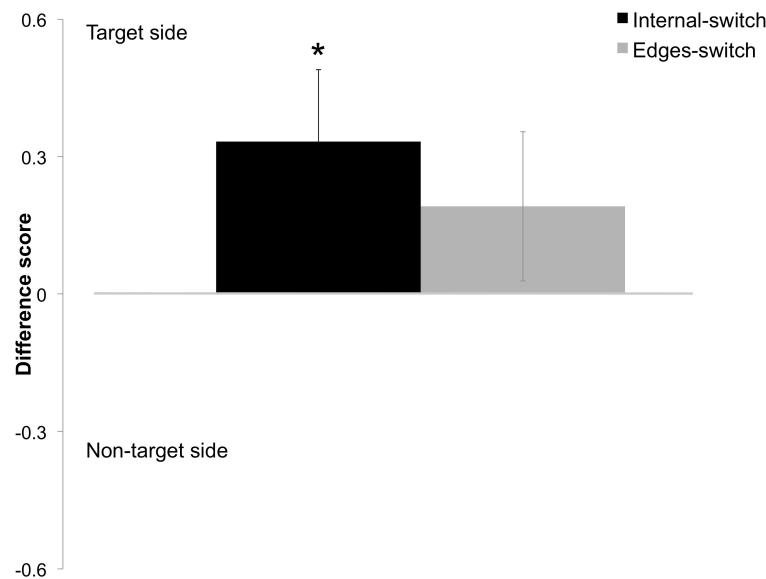


Figure 3. Normalized difference scores considering first fixations in the test phase of experiment

2. Infants searched for the puppet on the target side in the Internal-Switch Condition and not in the Internal-Change Condition. Bars indicate standard error of the mean (SEM). * $p < 0.05$.

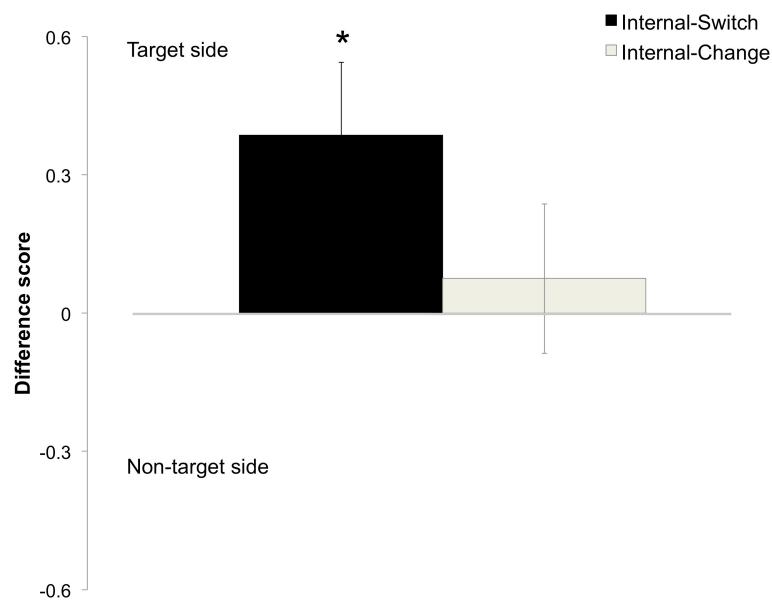
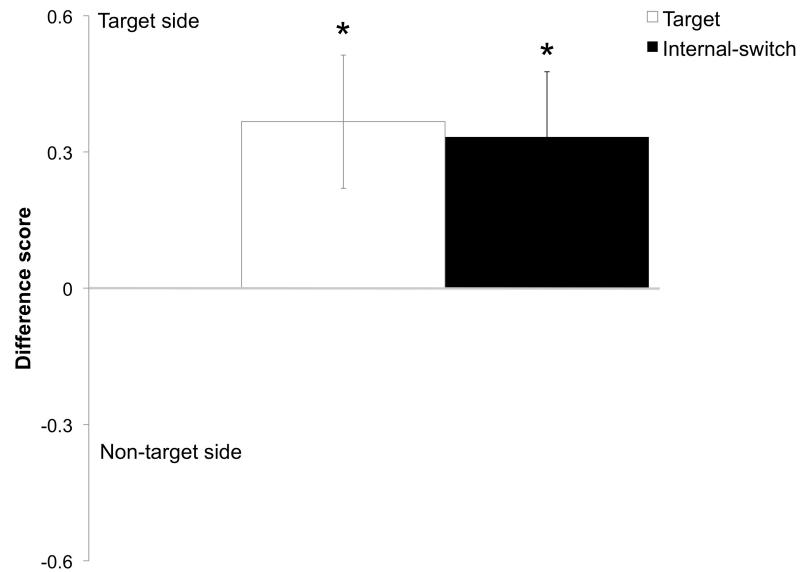


Figure 4. Normalized difference scores considering first fixations in the test phase of experiment

3. The y-axis shows the mean difference scores. Positive difference scores indicate that infants searched for the puppet on the target side while negative difference scores indicate that infants searched for the puppet on the non-target side. Infants searched for the puppet on the target side both in the Target Condition and in the Internal-Switch Condition. Bars indicate standard error of the mean (SEM). * $p < 0.05$.



Appendix 1

Pseudowords used in the familiarization and in the test phase of the three experiments

FAMILIARIZATION PHASE	
Non-targets	
lumorisape	
ganidemovu	
zunikesano	
banigukemo	
puderimola	
rudesakeni	
todegusazi	
vogukerita	
bogusaderi	
fagunimobe	
vorikegusa	
merizanigu	
vusadenibo	
tisamokezu	
lukerideno	
pesaniguto	

FAMILIARIZATION PHASE	TEST PHASE			
Targets		Experiment 1	Experiment 2	Experiment 3
sotumavefi	Condition 1	sovematufi	sovematufi	sovematufi
	Condition 2	fitumaveso	sozimalofi	sotumavefi
or				
tusomafive	Condition 1	tufimasove	tufimasove	tufimasove
	Condition 2	vesomafitu	tunemabuve	tusomafive