



**HAL**  
open science

# Infants' statistical word segmentation in an artificial language is linked to both parental speech input and reported production abilities

Mélanie Hoareau, H. Henny Yeung, Thierry Nazzi

## ► To cite this version:

Mélanie Hoareau, H. Henny Yeung, Thierry Nazzi. Infants' statistical word segmentation in an artificial language is linked to both parental speech input and reported production abilities. *Developmental Science*, 2019, 22 (4), pp.e12803. 10.1111/desc.12803 . hal-02292021

**HAL Id: hal-02292021**

**<https://hal.science/hal-02292021>**

Submitted on 19 Sep 2019

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

**Infants' statistical word segmentation in an artificial language is linked to both parental speech input and reported production abilities**

**Running head** : Input, babbling & statistical word segmentation

*Mélanie Hoareau<sup>1</sup>, H. Henny Yeung<sup>2</sup> & Thierry Nazzi<sup>1,3</sup>*

<sup>1</sup> Integrative Neuroscience and Cognition Center, UMR 8002, Université Paris Descartes, Sorbonne Paris Cité, Paris, France

<sup>2</sup> Department of Linguistics, Simon Fraser University, Burnaby, BC, Canada

<sup>3</sup> CNRS (Integrative Neuroscience and Cognition Center, UMR 8002), Paris, France

melanie.hoareau@parisdescartes.fr, henny\_yeung@sfu.ca, thierry.nazzi@parisdescartes.fr

Please, note that this version of the manuscript may be slightly different from the final accepted version

## **Research highlights**

- Statistical learning mechanisms for word segmentation are related to environmental and infant-specific factors
- The amount of parental input at 4 months is linked to the strength in 8-month-olds' statistical word segmentation
- Production abilities at 8 months are also correlated with 8-month-olds' success in statistical word segmentation
- Implications for individual differences in laboratory-based statistical word segmentation tasks are discussed

## **Abstract**

Individual variability in infant's language processing is partly explained by environmental factors, like the quantity of parental speech input, as well as by infant-specific factors, like speech production. Here we explore how these factors affect infant word segmentation. We used an artificial language to ensure that only statistical regularities (like transitional probabilities between syllables) could cue word boundaries, and then asked how the quantity of parental speech input and infants' babbling repertoire predict infants' abilities to use these statistical cues. We replicated prior reports showing that 8-month-old infants use statistical cues to segment words, with a preference for part-words over words (a novelty effect). Crucially, 8-month-olds with larger novelty effects had received more speech input at 4 months and had greater production abilities at 8 months. These findings establish for the first time that the ability to extract statistical information from speech correlates with individual factors in infancy, like early speech experience and language production. Implications of these findings for understanding individual variability in early language acquisition are discussed.

*Key words:* infant word segmentation, transitional probabilities, individual variability, language development predictors, input, babbling

## Introduction

Infants' ability to recognize word forms in continuous speech, or *word segmentation*, develops during the first year of life. This ability is crucial for the acquisition of words, as words are rarely heard in isolation, even when caregivers are explicitly asked to teach new words to their infants (Aslin, Woodward, LaMendola & Bever, 1996). Moreover, since there is no deterministic cue in running speech that indicates the onsets of words, infants must be sensitive to a constellation of potential word boundary markers. Cross-linguistic studies have revealed that many cues which infants rely upon are language-specific: metrical units in stressed-based languages, like English, (Jusczyk, Houston & Newsome, 1999) versus syllables in syllable-based languages, like French (Nazzi, Iakimova, Bertoni, Frédonie & Alcantara, 2006; Nishibayashi, Goyet & Nazzi, 2015); or the use of familiar words in the native language as anchors for determining word boundaries (Bortfeld, Morgan, Golinkoff, & Rathburn, 2005; Mersad & Nazzi, 2012).

By their very nature, the cues described above need to be learned, which stands in contrast with the use of 'domain-general' cues to word boundaries, like statistical regularities (Newport & Aslin, 2000; Aslin, Saffran & Newport, 1999). These statistical cues are often thought to be less dependent on other aspects of language development (Erickson & Thiessen, 2015), but this issue is relatively unexplored. Here we asked whether statistical cues to word segmentation are linked to well-known developmental factors that are highly variable between individuals, like quantity of early parental input, or individual differences in language production abilities, like babbling. Specifically, we evaluated whether concurrent and prior measures of speech input and babbling ability could predict 8-month-old infants' abilities to use statistical information to segment words, like transitional probabilities (TPs) between syllables.

### Individual variability in infants' speech environment

Understanding links between mechanisms of word segmentation and individual variability is important, because many researchers have shown that performance in speech segmentation tasks is highly predictive of later vocabulary learning (Junge & Cutler, 2014; Newman, Ratner, Jusczyk, Jusczyk & Dow, 2006; Singh, Steven Reznick, & Xuehua, 2012; Von Holzen, Nishibayashi & Nazzi, 2018). However, to our knowledge, only one study has previously investigated links between parental speech input and word segmentation abilities in the first year of life (Newman, Rowe & Ratner, 2016). While results from that longitudinal study showed that vocabulary development at 24 months of age was linked to both early speech input and early segmentation skills at 7.5 months, there was no clear evidence that maternal input and word segmentation abilities were themselves

directly interrelated.

Our study makes two novel contributions to the understanding of the links between early speech input and infant word segmentation. First, infant segmentation abilities emerge from around 4 to 8 months in both English (the language of the infants tested in Newman et al., 2016) and French (the language of the infants tested in the present study; see Berdasco, Nishibayashi, Baud, Biran & Nazzi, 2018), which means that the infants at 7.5 months that Newman et al. (2016) tested had already developed some segmentation abilities. However, speech input might play a greater role *earlier* in development, when these speech segmentation abilities are first emerging, something not assessed in Newman et al. (2016). Indeed, because any word segmentation procedure must operate over linguistically defined units (such as phonemes and syllables), speech input in the earliest developmental stages may contribute to the building of these linguistic categories. Accordingly, it would be important to measure the prospective effects of language input on later segmentation abilities. Here, we tested French-learning infants in a longitudinal design at both 4 to 8 months of age in order to determine how speech input might impact the development of both subsequent and concurrent segmentation abilities. Input was measured at both ages, but segmentation was only tested at 8 months, since this was the only age at which the ability to use the kind of statistical cues explored here (see next paragraph) had been tested and found in French (Mersad & Nazzi, 2012).

A second contribution of our present study is to use an artificial language to isolate a single segmentation cue, which differs from Newman et al. (2016), who used natural speech utterances containing numerous word boundary cues, like phrasal and word-level prosody, and whose stimuli likely contained many familiar word forms that infants could use as word boundary anchors. Their ability to find links between segmentation abilities and input may thus have been complicated by infants' varying sensitivity to these overlapping cues. Hence, in the present work, we focus on the link between speech input and infants' ability to use statistical co-occurrences as cues to segmentation. It is well established that infants are sensitive to statistical regularities in continuous speech, and may exploit these regularities in order to discover word forms. Statistical approaches to word segmentation have mostly focused on transitional probabilities (TPs) between syllables, which could be a cue to segmentation since syllables co-occur more consistently within words than between words (Curtin, Mintz & Christiansen, 2005). Moreover, infants seem to track TPs between syllables in both synthesized and naturally produced speech (e.g. Aslin, Saffran & Newport, 1998; Pelucchi, Hay & Saffran, 2009; Saffran, Aslin & Newport, 1996), suggesting that TP computation is a useful learning mechanism for word segmentation outside the laboratory, which also interacts with other cues to segmentation (Erickson & Thiessen, 2015; Johnson & Jusczyk, 2001; Mersad &

Nazzi, 2012). Although some limitations have been identified (Johnson & Tyler, 2010; Mersad & Nazzi, 2012), several studies now show that infants' learning of word meanings is easier for word forms with high internal TPs (Graf Estes, 2012; Graf Estes, Evans, Alibali & Saffran, 2007), that infants who are better able to identify speech structures with high internal TPs are also faster to recognize familiar words (Lany, Shoaib, Thompson & Graf Estes, 2018), and that infant statistical learning as measured in *visual* tasks are linked to individual differences in concurrent (Shafto, Conway, Field & Houston, 2012) and later language abilities (Ellis, Gonzalez & Deák, 2014).

Crucially though, no study has explored where variation in statistical learning comes from in the first place. This gap in knowledge about sources of variation in infants' statistical segmentation abilities is surprising, given that any word segmentation procedure depends on the existence of linguistically defined units. Thus, it is likely that linguistic variation among individuals, like the amount of speech input they receive, may affect the robustness of linguistic representations upon which statistical segmentation operates. Here we investigate whether variation in earlier and concurrent amount of speech input is related to the ability of infants to use statistical procedures for segmentation.

#### Individual variability in infants' speech production

Factors intrinsic to infants may also affect word segmentation abilities. For example, canonical babbling, or the rhythmic production of adult-like syllables, emerges between 6 and 8 months of age (Oller, 2000), around the same time as segmentation abilities. Yet little is known about how individual variation in speech production might be linked to segmentation. One approach in the literature has suggested that infants' mastery of certain speech forms in an individual's production inventory, particularly consonants, predicts their ability to detect word forms containing these consonants in fluent speech (DePaolis, Vihman, & Keren-Portnoy, 2011). The ability to produce consistent (consonant) patterns in production may help in constructing a stable phonological memory for target word forms (Vihman, 1993), which could then aid speech segmentation by lessening the processing load for related word forms.

In addition to increases in infants' productive speech sound inventory, a second type of production ability that may predict speech segmentation abilities is the complexity of the syllable structure in babbling itself. Specifically, word segmentation in French is highly dependent on the recognition of a syllable as an important rhythmic unit (Nazzi et al., 2006), and so infants' emerging ability to produce cohesive speech syllables might also be correlated with the stability of

phonological memories for syllables, which could then impact infants' ability to segment words using a syllable-based procedure in French.

To capture variability in these intrinsic speech production factors, parents completed a detailed questionnaire about segmental and syllable structure aspects of their infants' speech production abilities, evaluating whether the sophistication of segmental and syllable structure production at 4 and 8 months is related to infants' ability to use TP information to segment words at 8 months.

## Methods

Our two main objectives for this study asked how the strength of individual statistical segmentation abilities (as attested by magnitude of individual novelty preferences) was modulated by two measures of infants' skill in their native language. First, we investigated links with the amount of parental speech input, both earlier in infancy (four months prior), and at the same age when infants were tested on word segmentation. Second, we investigated how word segmentation was associated with speech production abilities at both these earlier and concurrent time points.

### Participants

Thirty-three monolingual French-learning infants (19 females) participated in the segmentation study when they were 8 months of age (*mean age*: 8 months and 12 days; *range*: 8 months to 9 months and 9 days). The sample size was specified a priori on the basis of previous research that used a similar design and procedure (Mersad & Nazzi, 2012: Experiment 1, which had an effect size of  $d = .57$ ). A power analysis suggested that in order to have a power of .80-.90, final sample size should be about 27 to 35 participants, which was the target range of our final sample. It should also be noted that the infants tested here were a subset of a sample from a larger longitudinal study on a different research question (about audiovisual speech processing, for which infants completed a short 2-minute task that involved watching a woman on a screen talking). Here, we report only and all factors related to the present research question.

All infants were born full-term and were exposed to French at least 90% of the time in their environment. Ten additional infants were tested but not included due to fussiness (5) or crying (5). All parents gave informed consent before participation. Infants' concurrent babbling abilities was assessed on the day the segmentation task was performed, and concurrent speech input was assessed about a week after the lab visit. Infants' input and babbling were also assessed about 4 months prior



to the segmentation task (*mean age*: 4 months and 14 days; *range*: 4 months–5 months and 4 days).

## Stimuli

The stimuli used in the experiment were taken from Experiment 1 of Mersad & Nazzi (2012). Two language variants (A and B) were constructed such that words and part-word test items were counterbalanced (see Table 1). Specifically, the familiarization streams of each language were constructed by concatenating four trisyllabic nonce words (hereafter words), with two frequent words occurring 90 times each, and the two target words occurring 45 times each. TPs within words were thus equal to 1, while concatenation was pseudo-random with the following constraints: The same word never occurred twice in a row; Each frequent word was followed by the other frequent word half of the time (resulting in TPs between the two frequent words equal to .5). The resulting languages were structurally identical to those used in Aslin et al. (1998).

The words used at test were the two target words and two part-words. The part-words were constructed by concatenating the last syllable of a frequent word and the first two syllables of the other frequent word. Words and part-words were thus all trisyllabic (see Table 1). TPs within words were equal to 1 whereas TPs within part-words were equal to .75 (average of .5 for the two syllables crossing the frequent word boundary and of 1 for the two syllables taken from the second frequent word). Hence, TPs within words were higher than TPs within part-words. However, words and part-words had appeared with the same frequency in the familiarization strings due to the way part-words were made from the frequent words. Test stimuli were presented in lists each consisting of 15 occurrences of one of the words or part-words separated by 500 ms silences. All lists were 13.65 s long.

In order to control the acoustic properties of the stimuli, familiarization and test stimuli were all synthesized with MBROLA (Dutoit, Pagel, Pierret, Bataille, & Van der Vrecken, 1996) using the French female diphone database (fr2). All phonemes had the same duration (111 ms) and F0 (200 Hz). There were no pauses or acoustic cues to word boundaries in the familiarization sequences. The familiarization streams lasted 3.03 minutes with a 4.45 syllable/s speech rate. An increasing and decreasing amplitude ramp was applied, respectively, to the first and last 5 seconds of the streams to ensure that words corresponding to the fade-in or the fade-out of the familiarization were not clearly audible.

## Procedure

**Home recordings.** To measure quantity of parental speech input at 4 and 8 months of age, we used

audio recorders from the LENA™ system to collect home recordings of a typical weekend day in which both parents were present, independently of the mode of daycare during the week (see more on this issue in Discussion section). We used raw LENA™ Adult Word Counts (AWCs) to estimate the average number of adult words directed at or uttered close to the infant per hour of recording, thus controlling for differences in recording duration. Under quiet conditions, AWCs for European French-learning infants are comparable to measures based on human coders, while background noise can introduce error in the calculation of AWCs (Canault, Le Normand, Foudil, Loundon, & Thai-Van, 2016). To maximize reliability of our measures, we instructed parents to make the recording on a day in which no atypically noisy activities were planned (e.g., birthday parties, large outdoor gatherings, etc.).

Due to constraints from our ethics board, the recordings could not be listened to assess measures regarding the quality of input, like when infants were being addressed. These AWC measures were obtained for 33 infants at 4 months of age, and 32 infants at 8 months of age (due to one family being unavailable to record at the later age).

**Language production outcomes.** All parents filled out a babbling questionnaire at both ages (adapted from Gonzalez-Gomez & Nazzi, 2012a). The choice of using a caregiver-based repertoire assessment was based on several studies showing the reliability of parental reports for word production from 8 to 30 months (Fenson et al., 1991) and the onset of babbling (Oller, Eilers, & Basinger, 2001). In particular, one study showed that parents' reporting of the type of syllables their infants produce between 8 to 12 months of age strongly correlates with laboratory-based observations (Ramsdell, Oller, Buder, Ethington & Chorna, 2012).

Our questions were posed as a checklist, and parents indicated whether their infant produced certain vowels, consonants, or syllable structures. The questionnaire had a total of 16 items. It asked whether infants produced 7 vowels common in French-learning infants' early productions (/a/, /e/, /ø/, /i/, /o/, /u/, /y/), 6 different consonant categories (labials, coronals, liquids, velars, nasals), and what types of syllable sequences (syllable repetitions, like /jaja/ or /baba/; vowel repetitions with different consonants, like /yaba/ or /baya/; or concatenations of syllables with different consonants and vowels, like /badi/). The production outputs of the LENA™ system (Child Vocalizations Counts) were not considered as relevant for our project, as they give only a broad approximation of vocalization quantity, and are not measures of the sophistication of speech production (i.e., cries, coughs, and speech productions are all counted).

**Statistical segmentation task.** The experiment was conducted inside a sound-attenuated room, which contained a three-sided test booth. The test booth had a red light and a loudspeaker (SONY xs-F1722) mounted at eye level on each of the side panels and a green light mounted on the center panel. A video camera used to monitor infants' behavior was placed below the center light. A Dell Optiplex computer, a TV screen connected to the camera, and a response box connected to the computer were located outside the sound-attenuated room. The box was controlled by the observer, who was looking at the video of the infant on the TV screen. The response box allowed the observer to send to the computer the information about the infant's head direction and hence to control the flashing of the lights and the presentation of the sounds. The observer, and also the infant's caregiver, wore earplugs and listened to masking music over tight-fitting headphones, which prevented them from hearing the stimuli presented.

The procedure used in the present study was the version of the Headturn Preference Procedure (HPP) adapted from Saffran et al. (1996) by Mersad & Nazzi (2012). Each infant was held on a caregiver's lap seated in a chair in the center of the test booth. Each session began with a familiarization phase during which infants heard the stimuli, which were delivered by the loudspeakers via an audio amplifier (Marantz PM4000). During the 3.03-minute familiarization, the central and lateral lights were alternatively made to blink to keep infant's attention, but orientation times were not measured. Note that in this phase of the experiment, the blinking of the lights was not contingent on the infant's head-turns. The test phase, immediately following the end of the familiarization phase, consisted of two blocks, each corresponding to the presentation of the four lists (one for each word and part-word of the language) in a unique pseudo-random order within each block. Each test trial began with the green light on the center panel blinking until the infant had oriented in that direction. Then, the center light was extinguished and the red light above the loudspeaker on one of the side panels began to flash. When the infant made a turn of at least 30° in the direction of the loudspeaker, the stimulus of the trial began to play. Stimulus was played to completion or stopped immediately after the infant failed to maintain the 30° head-turn for 2 consecutive seconds (with a 200 ms fade-out). If the infant turned away from the target by 30° in any direction for less than 2 s and then turned back again, the trial continued, but the time spent looking away was not included in the orientation time. Thus, the maximum orientation time for a given trial was the duration of the entire speech sample. If a trial was less than 1.5 s, the trial was repeated and the original orientation time was discarded. The flashing red light remained on for the entire duration of the trial. Half of the infants were assigned to Language A, and the other half to Language B (see Table 1).

## Results

Mean orientation times (OTs) were calculated for the lists of words versus part-words. A repeated measures ANOVA with the within-subjects factors of Familiarity (words vs. part-words) and Block (1 vs. 2) was conducted. The effect of Familiarity (see Figure 1) was significant ( $F(1,32) = 4.91, p = .034, \eta_p^2 = .133$ ) indicating that infants oriented longer to the part-words ( $M = 7.2$  s, 95% CI = 6.39 to 8.02) than the words ( $M = 6.5$  s, 95% CI = 5.69 to 7.41). The effect of Block also reached significance ( $F(1,32) = 2.28, p = .003, \eta_p^2 = .243$ ) indicating that overall infants oriented significantly less in Block 2 ( $M = 6.1$  s, 95% CI = 5.37 to 6.84) than in Block 1 ( $M = 7.6$  s, 95% CI = 6.57 to 8.73). The Familiarity x Block interaction did not reach significance ( $F(1,32) = .17, p = .681, \eta_p^2 = .005$ ), indicating that infants kept their novelty preference throughout the test phase.

Our primary question was whether individual differences in statistical learning are related to quantity of speech input and language outcomes. To address this question, we computed number of adult words (AWCs) and production scores for each infant. Means and SDs at both ages are presented in Table 2. Paired t-test revealed that while AWCs were stable over the two ages ( $t = 1.65, p = .109$ ), production scores increased with age ( $t = -13.38, p < .001$ ). Then, we calculated the OT difference scores ( $OT_{words} - OT_{part-words}$ ; with negative scores indicating novelty preferences) to conduct correlational (Table 3) and regression (Table 4) analyses. First, we found that the size of 8-month-olds' novelty effect increased with amount of adult words at 4 months (*Pearson's*  $r = -.37, p = .035$ , see Figure 2), though not at 8 months ( $r = -.16, p = .372$ ). Second, we found that the size of 8-month-olds' novelty effect increased with production abilities at the same age ( $r = -.42, p = .018$ , see Figure 3) but not at 4 months ( $r = -.08; p = .657$ ).

We last asked whether amount of parental speech input and babbling repertoire, when considered simultaneously, could explain unique variation in infant segmentation ability at 8 months. We conducted multiple regression analyses using parental input and babbling scores at 4 and 8 months of age as predictors of infant novelty preference at 8 months (see Table 4). All variables were mean centered to allow for meaningful interpretation. To control for multicollinearity (see also Table 2), we determined variance inflation factors (VIF) from a model including all four variables using the function VIF in SPSS. This revealed collinearity not to be an issue (mean VIF = 1.8 so not substantially greater than 1; mean tolerance = .655, with the minimum tolerance = .5 for the babbling at 4 months, so substantially greater than .2). The model significantly explained 38% of the variance in TP-based segmentation ability. However, only parental input at 4 months and babbling abilities at 8 months were independent and significant predictors of the

novelty effect in TP-based segmentation at 8 months.

## Discussion

As a group, we replicated the expected preference for part-words over words (i.e., a novelty effect) found in previous reports (Mersad & Nazzi, 2012; Saffran et al., 1996), showing again that French-learning infants are able to use statistical cues to segment words at 8 months of age. Importantly, we also found that statistical learning efficiency for speech segmentation at this age is related to both infants' earlier experience with linguistic input (as evaluated using automatic analyses provided by LENA) 4 months prior, and to babbling ability (as evaluated through parental questionnaire) at 8 months of age.

A conservative way of interpreting these findings is that they reveal an interaction of overall environmental stimulation with global cognitive skill. On this account, infants receiving richer environmental input might develop better perceptual and production abilities, and infants with higher perceptive and production abilities would elicit richer input from their parents, creating a virtuous circle that would lead to links between speech input, production, and performance in this word segmentation task (see Johnson, Lahey, Ernestus, & Cutler, 2013, for related ideas on this issue). However, two features of our results suggest more complex patterns of causality between our measured factors. First, the correlation analyses failed to reveal links between input measures and production abilities. While production milestones have sometimes been linked to environmental factors (Franklin et al., 2014; Oller et al., 2013), and social feedback loops between parent interactions and babbling are well established (Albert, Schwade, & Goldstein, 2018; Gros-Louis & Miller, 2018; Warlaumont, Rochards, Gilkerson, & Oller, 2014), it is unclear whether more sophisticated productions necessarily lead to increased talk to infants: We did not find such links in our data. Second, the regression analysis revealed that both input and production abilities contribute uniquely to explaining part of the variance in statistical segmentation performance. Thus, more specific links than simply positive feedback loops between general cognitive abilities and speech input may be at play. In the following, we discuss how input and production ability might individually be implicated in TP-based segmentation.

We first consider the role of input on statistical word segmentation. From a methodological point of view, our findings with AWC measures indirectly show that automated LENA measures are accurate enough to be used in research studies, in spite of known technical limitations (particularly the negative effects of noise, see Canault et al., 2016). However, it is also crucial to moderate our comments by considering how well the LENA™ measurements (recorded during a weekend day) reflect infants' global speech environment at different ages. In Figure 4, for example,

we observed that the modes of child-care during the week change from mostly home-care with the mother at 4 months to mostly shared nanny (small-group care in a private residence with ~2-5 other infants) or day-care (in a dedicated child-care space with >5 other infants and children) at 8 months. Our measure of parental input might thus be less representative of infants' linguistic environment at 8 than at 4 months, which might explain why we found a link between segmentation abilities and earlier, but not concurrent, speech input. Future studies will have to continue exploring the relative strengths and weaknesses of such automated analyses (compared to manual coding) and their optimal conditions of use.

From a theoretical point of view, our finding is compatible with the idea that statistical learning, although often considered a domain general ability observed from birth (Teinonen, Fellman, Näätänen, Alku, & Huotilainen, 2009; Kudo, Nonaka, Mizuno, Mizuno, & Okanoya, 2011; Bulf, Johnson, & Valenza, 2011), is modulated by domain-specific experience. One way this effect may occur is that linguistic experience primes infants' ability to segment words by attuning their perceptual sensitivities to language-specific sounds, building more mature linguistic categories (i.e., for phonemes and syllables). This enhanced development of linguistic category types may in turn help infants subsequently apply their innate statistical learning abilities to the problem of finding words in running speech by improving the robustness of these linguistic units over which statistical calculations can be made.

Given that we found that segmentation ability is related to input at 4 but not 8 months, and if this is not solely an artefact of less reliable measurements at 8 months (as discussed above), it may be that once infants have sufficiently mature phonological knowledge, and have learned how to apply statistical mechanisms to segment the speech stream, concurrent amount of parental speech no longer predicts their use of TPs. That is, it may be that after a certain point in development, infants no longer benefit from additional parental speech input to improve their statistical ability to find word boundaries. This, added to the fact that the stimuli in Newman et al. (2016) contained multiple segmentation cues, might also explain the lack of direct relationship between parental input and segmentation abilities at the age of 7.5 months reported in that previous study. One way to evaluate this hypothesis in greater detail would be to test the same infants on both TP-based segmentation abilities and phonological processing abilities. For example, one might examine categorical perception, like Hoonhorst et al. (2009), or the acquisition of the language-specific consonant bias in lexical processing, as done by Nishibayashi & Nazzi (2016), since both of these phonological abilities are likely language-specific by 8 months of age, and could then be used to test the prediction that variability in performance on phonological and statistical segmentation tasks should be linked.

This aspect of our results is of high interest because it contrasts with the dominant view about the development of infants' word segmentation. Instead of supporting the view that statistical information is used before language-specific information (Erickson & Thiessen, 2015), our results suggest instead that there are a number of back-and-forth interactions between language-dependent factors (e.g., amount of parental speech input) and innate, domain-general learning mechanisms (e.g., tracking of TPs). At a minimum, we provide support for the idea that infant TP-based ability to detect words is related to their language environment early in life. In future work it will be important to continue investigating the link between early experience and statistical learning by replicating and extending our work to the use of non-linguistic visual stimuli in a statistical learning task. If no relationship were found between parental speech input and TPs between visual forms, this would support the idea that robust speech input early in life helps develop domain-specific perceptual categories or units that are necessary for statistical computations. If a positive relationship were found, it would imply that early experience with speech is related to a more general cognitive mechanism that may contribute to the detection of structures and regularities in infants' environment.

A second major finding is that 8-month-old infants' ability to use TPs to detect novel sequential structures in auditory perception is also related to more mature language outcomes in production at the same age (but not at the earlier age of 4 months). From a methodological point of view, it suggests that at least by 8 months of age, parents were able to evaluate their infants' productions sufficiently well to be able to observe such links (adding to the validity of parent-based measures in some defined contexts; see also Ramsdell et al., 2012).

From a theoretical point of view, the regression results showed that our babbling measure taken at 8 months of age explains variance in individual statistical segmentation ability above and beyond that of measures of speech input is that speech perception and production develop in tandem. This interpretation is in line with recent research on infants supporting the notion that the ability to produce certain sounds is linked to various measures of speech perception (e.g., Altvater-Mackensen, Mani, & Grossman, 2016; DePaolis et al., 2011, DePaolis, Vihman, & Nakai, 2013; Majorano, Vihman, & DePaolis, 2014; Streri, Coulon, Marie, & Yeung, 2016). In the current study, better segmentation abilities might have led to better specified phonological representations and syllable productions. Conversely, a more diverse babbling inventory might have facilitated infants' ability to recognize familiar sound sequences while listening to speech in the familiarization phase (or even to recognize individual word forms in the test phase; DePaolis et al., 2011). Moreover, infants' production of more complex syllable structures might have helped them to better process the syllables in the continuous speech stream, and to detect their statistical structure. One way to

further probe the link between word segmentation and production abilities would be to create two artificial streams with syllables that are produced either early or late in babbling and to see whether detecting the statistical structures for the earlier versus the later produced forms in babbling would differ.

## Conclusions

This study explored potential environmental and intrinsic factors that could contribute to individual variability in statistical segmentation skills. Results show a link between the amount of speech input received at 4 months, the level of speech production at 8 months, and the effective use of statistical procedures for word segmentation at 8 months of age. While correlations do not specify causal links between the correlated factors, and we have discussed the possibility that the observed links were in part driven by underlying factors common to all of these factors, we have also argued, based on the lack of link between input and production abilities, and their unique contribution to explain variance in TP-based performance, that our study nevertheless suggests more specific links between the strength of TP-based word segmentation abilities and the amount of speech input, as well as speech production abilities. Future studies will be needed to further support and better specify these links. Similar to other developmental phenomena, it is likely the case that there is a high degree of interdependence between domain-general competencies, like statistical learning, and domain-specific ones, like phonological encoding. Overall, however, our finding is notable because it suggests that variation in experience with language can impact even perceptual-cognitive processes that are widely assumed to be domain-general.



## Acknowledgements

This study was supported by the EU Horizon 2020 research and innovation program (H2020-MSCA-ITN-2014, N°641858) to TN and MH, by ANR blanche grant (ANR-13-BSH2-0004) to TN and Labex-EFL (ANR-10-LABX-0083) grant to TN and HY.

**Conflict of interest:** The authors report no conflict of interest.

## References

- Altvater-Mackensen, N., Mani, N., & Grossmann, T. (2016). Audiovisual speech perception in infancy: The influence of vowel identity and infants' productive abilities on sensitivity to (mis)matches between auditory and visual speech cues. *Developmental Psychology, 52*(2), 191–204. <http://dx.doi.org/10.1037/a0039964>
- Albert, R. R., Schwade, J. A., & Goldstein, M. H. (2017). The social functions of babbling: acoustic and contextual characteristics that facilitate maternal responsiveness. *Developmental science, e12641*.
- Aslin, R. N., Saffran, J. R., & Newport, E. L. (1998). Computation of conditional probability statistics by 8-month-old infants. *Psychological Science, 9*(4), 321–324. <https://doi.org/10.1111/1467-9280.00063>
- Aslin, R. N., Saffran, J. R., & Newport, E. L. (1999). Statistical learning in linguistic and nonlinguistic domains. In MacWhinney B. (Ed.), *The emergence of language* (pp. 359–380). Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.
- Aslin, R. N., Woodward, J. Z., LaMendola, N. P., & Bever, T. G. (1996). Models of word segmentation in fluent maternal speech to infants. In Morgan J. L. & Demuth K. (Eds.), *Signal to syntax: Bootstrapping from speech to grammar in early acquisition* (pp. 117–134). Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.
- Berdasco-Muñoz, E., Nishibayashi, L. L., Baud, O., Biran, V., & Nazzi, T. (2018). Early segmentation abilities in preterm infants. *Infancy, 23*(2), 268–287. <http://doi.org/10.1111/infa.12217>
- 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*(4), 298–304. <https://doi.org/10.1111/j.0956-7976.2005.01531.x>
- Bulf, H., Johnson, S. P., & Valenza, E. (2011). Visual statistical learning in the newborn infant. *Cognition, 121*(1), 127–132. <https://doi.org/10.1016/j.cognition.2011.06.010>
- Canault, M., Le Normand, M.-T., Foudil, S., Loundon, N., & Thai-Van, H. (2016). Reliability of the Language ENvironment Analysis system (LENATM) in European French. *Behavior Research Methods, 48*(3), 1109–1124. <http://doi.org/10.3758/s13428-015-0634-8>
- Curtin, S., Mintz, T. H., & Christiansen, M. H. (2005). Stress changes the representational landscape: Evidence from word segmentation. *Cognition, 96*(3), 233–262. <https://doi.org/10.1016/j.cognition.2004.08.005>
- DePaolis, R. A., Vihman, M. M., & Keren-Portnoy, T. (2011). Do production patterns influence the processing of speech in prelinguistic infants? *Infant Behavior & Development, 34*(4), 590–601. <https://doi.org/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(4), 642–649. <https://doi.org/10.1016/j.infbeh.2013.06.007>

- Dutoit, T., Pagel, V., Pierret, N., Bataille, F., & Van der Vrecken, O. (1996). The MBROLA project: Towards a set of high quality speech synthesizers free of use for noncommercial purposes. In H. T. Bunnell & W. Idsardi (Eds.), *Proceedings of the Fourth International Conference on Spoken Language Processing* (pp. 1393–1396). Wilmington, DE: Applied Science and Engineering Laboratories.
- Ellis, E. M., Gonzalez, M. R., & Deák, G. O. (2014). Visual prediction in infancy: what is the association with later vocabulary?. *Language Learning and Development*, 10(1), 36-50.
- Erickson, L. C., & Thiessen, E. D. (2015). Statistical learning of language: Theory, validity, and predictions of a statistical learning account of language acquisition. *Developmental Review*, 37, 66–108. <https://doi.org/10.1016/j.dr.2015.05.002>
- Franklin, B., Warlaumont, A. S., Messinger, D., Bene, E., Nathani Iyer, S., Lee, C. C., ... Oller, D. K. (2014). Effects of parental interaction on infant vocalization rate, variability and vocal type. *Language Learning and Development*, 10(3), 279–296. <https://doi.org/10.1080/15475441.2013.849176>
- Fenson, L., Dale, P. S., Reznick, J. S., Thal, D., Bates, E., Hartung, J. P., ... & Reilly, J. S. (1991). Technical manual for the MacArthur communicative development inventories. *San Diego, CA: San Diego State University*.
- Gonzalez-Gomez, N., & Nazzi, T. (2012). Phonotactic acquisition in healthy preterm infants. *Developmental Science*, 15(6), 885–894. <https://doi.org/10.1111/j.1467-7687.2012.01186.x>
- Graf Estes, K. (2012). Infants generalize representations of statistically segmented words. *Frontiers in Psychology*, 3(447). <https://doi.org/10.3389/fpsyg.2012.00447>
- Graf Estes, K., Evans, J. L., Alibali, M. W., & Saffran, J. R. (2007). Can infants map meaning to newly segmented words?: Statistical segmentation and word learning. *Psychological Science*, 18(3), 254–260. <https://doi.org/10.1111/j.1467-9280.2007.01885.x>
- Hoonhorst, I., Colin, C., Markessis, E., Radeau, M., Deltenre, P., & Serniclaes, W. (2009). French native speakers in the making: From language-general to language-specific voicing boundaries. *Journal of Experimental Child Psychology*, 104(4), 353-366.
- 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(4), 548–567. <https://doi.org/10.1006/jmla.2000.2755>
- Johnson, E. K., Lahey, M., Ernestus, M., & Cutler, A. (2013). A multimodal corpus of speech to infant and adult listeners. *The Journal of the Acoustical Society of America*, 134(6), EL534-EL540.
- Johnson, E. K., & Tyler, M. D. (2010). Testing the limits of statistical learning for word segmentation. *Developmental Science*, 13(2), 339–345. <https://doi.org/10.1111/j.1467-7687.2009.00886.x>
- Julie, G. L., & Miller, J. L. (2018). From ‘ah’ to ‘bah’: social feedback loops for speech sounds at key points of developmental transition. *Journal of child language*, 45(3), 807-825.
- Junge, C., & Cutler, A. (2014). Early word recognition and later language skills. *Brain Sciences*, 4(4), 532–559. <https://doi.org/10.3390/brainsci4040532>
- Jusczyk, P. W., Houston, D. M., & Newsome, M. (1999). The beginnings of word segmentation in english-learning infants. *Cognitive Psychology*, 39(3–4), 159–207.

<https://doi.org/10.1006/cogp.1999.0716>

- Kudo, N., Nonaka, Y., Mizuno, N., Mizuno, K., & Okanoya, K. (2011). On-line statistical segmentation of a non-speech auditory stream in neonates as demonstrated by event-related brain potentials. *Developmental Science*, *14*(5), 1100–1106. <https://doi.org/10.1111/j.1467-7687.2011.01056.x>
- Lany, J., Shoaib, A., Thompson, A., & Graf Estes, K. (2018). Infant statistical-learning ability is related to real-time language processing. *Journal of Child Language*, *45*(2), 368–391. <https://doi.org/10.1017/S0305000917000253>
- Majorano, M., Vihman, M. M., & DePaolis, R. A. (2014). The relationship between infants' production experience and their processing of speech. *Language Learning and Development*, *10*(2), 179–204. <https://doi.org/10.1080/15475441.2013.829740>
- Mersad, K., & Nazzi, T. (2012). When mommy comes to the rescue of statistics: Infants combine top-down and bottom-up cues to segment speech. *Language Learning and Development*, *8*(3), 303–315. <https://doi.org/10.1080/15475441.2011.609106>
- Nazzi, T., Iakimova, G., Bertoni, J., Frédonie, S., & Alcantara, C. (2006). Early segmentation of fluent speech by infants acquiring French: Emerging evidence for crosslinguistic differences. *Journal of Memory and Language*, *54*(3), 283–299. <https://doi.org/10.1016/j.jml.2005.10.004>
- Newman, R., Ratner, N. B., Jusczyk, A. M., Jusczyk, P. W., & Dow, K. A. (2006). Infants' early ability to segment the conversational speech signal predicts later language development: A retrospective analysis. *Developmental Psychology*, *42*(4), 643–655. <https://doi.org/10.1037/0012-1649.42.4.643>
- Newman, R. S., Rowe, M. L., & Ratner, N. B. (2016). Input and uptake at 7 months predicts toddler vocabulary: the role of child-directed speech and infant processing skills in language development. *Journal of Child Language*, *43*(5), 1158–1173. <https://doi.org/10.1017/S0305000915000446>
- Newport, E. L., & Aslin, R. N. (2000). Innately constrained learning: Blending old and new approaches to language acquisition. *Proceedings of the 24th Annual Boston University Conference on Language Development*, 1–21.
- Nishibayashi, L.-L., Goyet, L., & Nazzi, T. (2015). Early speech segmentation in French-learning infants: Monosyllabic words versus embedded syllables. *Language and Speech*, *58*(3), 334–350. <https://doi.org/10.1177/0023830914551375>
- Nishibayashi, L. L., & Nazzi, T. (2016). Vowels, then consonants: Early bias switch in recognizing segmented word forms. *Cognition*, *155*, 188–203.
- Oller, D. K. (2000). *The emergence of the speech capacity*. New York, NY: Psychology Press.
- Oller, D. K., Buder, E. H., Ramsdell, H. L., Warlaumont, A. S., Chorna, L., & Bakeman, R. (2013). Functional flexibility of infant vocalization and the emergence of language. *Proceedings of the National Academy of Sciences*, *110*(16), 6318–6323. <https://doi.org/10.1073/pnas.1300337110>
- Oller, D. K., Eilers, R. E., & Basinger, D. (2001). Intuitive identification of infant vocal sounds by parents. *Developmental Science*, *4*(1), 49–60.
- Pelucchi, B., Hay, J. F., & Saffran, J. R. (2009). Statistical learning in a natural language by 8-month-old infants. *Child Development*, *80*(3), 674–685. Retrieved from <http://www.jstor.org/stable/29738646>
- Ramsdell, H. L., Oller, D. K., Buder, E. H., Ethington, C. A., & Chorna, L. (2012). Identification of prelinguistic phonological categories. *Journal of Speech, Language, and Hearing Research*,

55(6), 1626-1639.

- Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, 274(5294), 1926–1928. <http://doi.org/10.1126/science.274.5294.1926>
- Shafto, C. L., Conway, C. M., Field, S. L., & Houston, D. M. (2012). Visual sequence learning in infancy: Domain-general and domain-specific associations with language. *Infancy*, 17(3), 247–271.
- Singh, L., Steven Reznick, J., & Xuehua, L. (2012). Infant word segmentation and childhood vocabulary development: A longitudinal analysis. *Developmental Science*, 15(4), 482–495. <https://doi.org/10.1111/j.1467-7687.2012.01141.x>
- Streri, A., Coulon, M., Marie, J., & Yeung, H. H. (2016). Developmental change in infants' detection of visual faces that match auditory vowels. *Infancy*, 21(2), 177–198. <https://doi.org/10.1111/infa.12104>
- Teinonen, T., Fellman, V., Näätänen, R., Alku, P., & Huotilainen, M. (2009). Statistical language learning in neonates revealed by event-related brain potentials. *BMC Neuroscience*, 10(21). <https://doi.org/10.1186/1471-2202-10-21>
- Vihman, M. M. (1993). Vocal motor schemes, variation and the production-perception link. *Journal of Phonetics*, 21(1–2), 163–169.
- Von Holzen, K., Nishibayashi, L. L., & Nazzi, T. (2018). Consonant and vowel processing in word form segmentation: An infant ERP study. *Brain Sciences*, 8(2), 24. <https://doi.org/10.3390/brainsci8020024>
- Warlaumont, A. S., Richards, J. A., Gilkerson, J., & Oller, D. K. (2014). A social feedback loop for speech development and its reduction in autism. *Psychological science*, 25(7), 1314–1324.

#### Figure and Table Legends

**Table 1.** All trisyllabic words and part-words used in languages A & B (f = frequent words)

**Table 2.** Descriptive statistics of AWCs and babbling scores

**Table 3.** Correlational analysis between all variables

**Table 4.** Regression models predicting infant segmentation ability at 8 months based on parental input and production abilities at 4 and 8 months (n= 32)

**Figure 1.** Mean OT differences between words and partwords (each dot represent the OT difference score of a given infant, and the bold line the group average)

**Figure 2.** Link between the segmentation effect at 8 months and input at 4 months

**Figure 3.** Link between the segmentation effect at 8 months and babbling at 8 months

**Figure 4.** Types of child care at 4 and 8 months during the weekdays

## Tables and figures

**Table 1.** All trisyllabic words and part-words used in languages A & B (f = frequent words)

<b>Language A</b>		<b>Language B</b>	
<b>Words</b>	<b>Part-words</b>	<b>Words</b>	<b>Part-words</b>
pabiku	tudaro	tudaro	pabiku
tibudo	pigola	pigola	tibudo
golatu <sup>f</sup>		budopa <sup>f</sup>	
daropi <sup>f</sup>		bikuti <sup>f</sup>	

**Table 2.** AWCs and production scores

		<b>Mean</b>	<b>SD</b>	<b>Range (min. - max.)</b>
<b>AWCs</b>	4m	1420	567	(393 - 2548)
(per hour)	8m	1262	516	(340 - 2323)
<b>Babbling</b>	4m	3	2	(0 - 10)
(raw scores)	8m	8	3	(2 - 16)

**Table 3.** Correlational analysis between all variables

		1	2	3	4	5
1	Novelty Effect	–	-.37*	-.16	-.08	-.42*
2	Input at 4m		–	.40*	-.13	-.06
3	Input at 8m			–	-.26	-.08
4	Babbling at 4m				–	.69***
5	Babbling at 8m					–

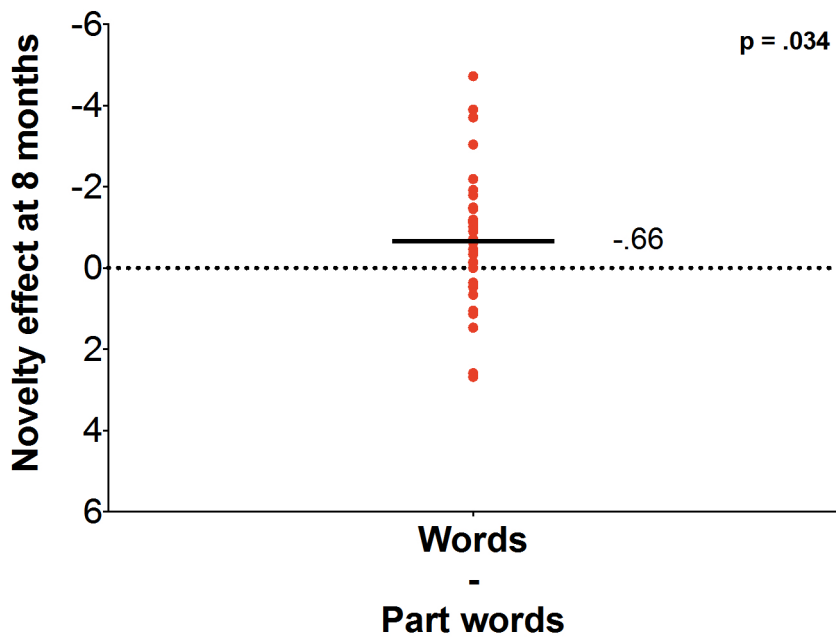
*Note:* \* is  $p < .05$ ; \*\* is  $p < .01$ ; \*\*\* is  $p < .001$

**Table 4.** Regression models predicting infant segmentation ability at 8 months based on parental input and production abilities at 4 and 8 months (n= 32)

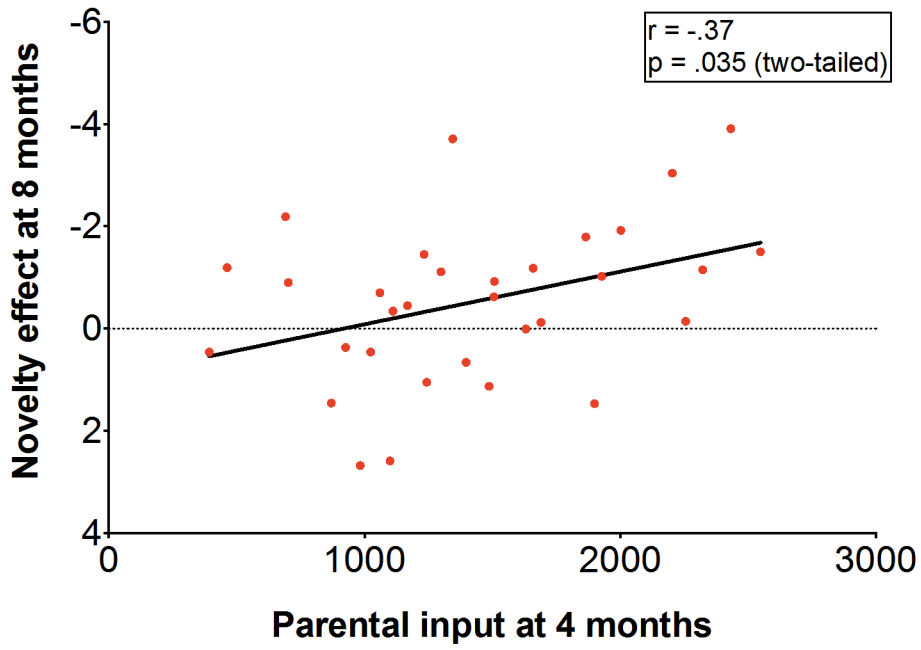
<i>Predictors</i>	<i>Novelty Effect at 8 months</i>		
	<i>B coefficients</i>	<i>95% CI for B</i>	<i>P value</i>
Intercept	<.001	-.305 – .305	1.000
*Input at 4 months	-.373	-.712 – -.035	.032
Input at 8 months	.017	-.333 – .368	.921
Babbling at 4 months	.327	-.119 – .772	.144
**Babbling at 8 months	-.660	-1.091 – -.229	.004
**R-squared stat		.384 (p = .009)	

*Note:* \* is  $p < .05$ ; \*\* is  $p < .01$





**Figure 1.** Mean OT differences between words and partwords (each dot represent the OT difference score of a given infant, and the bold line the group average)



**Figure 2.** *Link between the segmentation effect at 8 months and input at 4 months.*

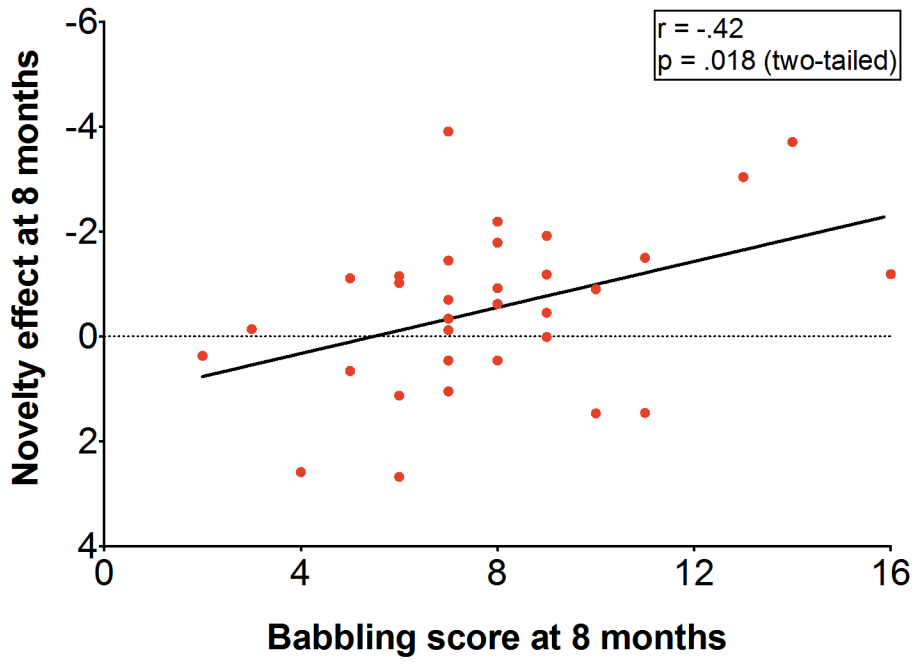
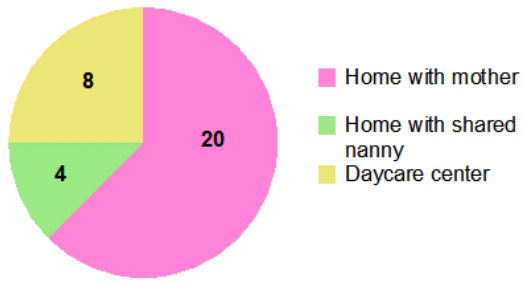


Figure 3. Link between the segmentation effect at 8 months and babbling at 8 months

Type of child care at 4m (weekdays)



Type of child care at 8m (weekdays)

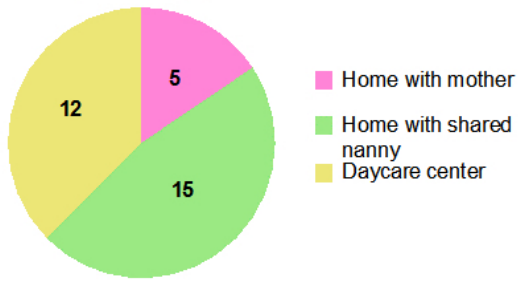


Figure 4. Types of child care at 4 and 8 months during the weekdays