IMPACT OF LATE-TO-MODERATE PRETERM BIRTH ON MINIMAL PAIR WORD-LEARNING Running head: Minimal pair word-learning in preterms

Clément François, Antoni Rodriguez-Fornells, Xim Cerda-Company, Thaïs Agut, Laura Bosch

IMPACT OF LATE-TO-MODERATE PRETERM BIRTH ON MINIMAL PAIR WORD-LEARNING

Running head: Minimal pair word-learning in preterms

Clément François[1*], Antoni Rodriguez-Fornells[2,3,4], Xim Cerda-Company[2,3], Thaïs Agut[5,6], Laura Bosch[3,6,7*]

[1]Aix-Marseille University, CNRS, LPL, Aix-en-Provence, France

[2]Cognition and Brain Plasticity Group [Bellvitge Biomedical Research Institute-] IDIBELL, L'Hospitalet de Llobregat, Barcelona, Spain

[3]Department of Cognition, Development and Educational Psychology, University of Barcelona, Barcelona, Spain

[4]Catalan Institution for Research and Advanced Studies, ICREA, Barcelona, Spain

[5]Department of Neonatology, Hospital Sant Joan de Déu, Barcelona, Spain

[6]Institut de Recerca Sant Joan de Déu (IRSJD), Barcelona, Spain

[7]Institute of Neurosciences (UBNeuro), University of Barcelona, Barcelona, Spain

*Corresponding authors:

Laura Bosch, Department of Cognition, Development and Educational Psychology, University of Barcelona, Faculty of Psychology, Institute of Neurosciences (UBNeuro), University of Barcelona, Barcelona, Spain. Email: laurabosch@ub.edu

Clément François, Laboratoire Parole et Langage, Aix-Marseille University, CNRS, UMR7309, Aix-en-Provence, France. Email: clement.francois@univ-amu.fr

ABSTRACT

Little is known about language development after late-to-moderate premature birth, the most significant part of prematurity worldwide. We examined minimal-pair word-learning skills in eighteen 18-months-old healthy full-term (mean gestational age at birth = 39.6 weeks; 7 males; 100% Caucasian) and eighteen healthy late-to-moderate preterm infants (mean gestational age at birth 33.7 weeks; 11 males; 100% Caucasian). Data were collected in the local urban area of Barcelona city from May 2015 to August 2016. Toddlers first associated two pseudo-words, forming a minimal pair based on a voice onset time (VOT) distinction of the initial consonant, with two unfamiliar objects during a habituation phase. A visual choice test assessed their recognition of the two novel word-object associations and some familiar word-object pairs. While full-terms successfully mapped the similar sounding pair of novel words (d = 1.57), preterms could not (d = 0.17). These results suggest that late to moderate preterm birth can hinder basic associative learning mechanisms relying on fine temporal speech features.

KEYWORDS: Late to moderate preterm birth; minimal pair word learning; looking while listening procedure; voice onset time (VOT) contrast

In 2010, the World Health Organization reported that almost 15 million babies were born preterm worldwide (Blencowe et al., 2012). Preterm birth can have a dramatic impact on early development, with around 30% of very preterm infants (i.e., born between 28- and 32-week gestational age [wGA]) showing delays in language acquisition at age two (Sansavini et al., 2010). Due to immature brain networks and prenatal limited experience with low-pass filtered speech, premature birth may alter developmental trajectories by increasing the risk for cognitive and linguistic delays (Gervain, 2015). Importantly, a meta-analysis involving 15 studies showed that immaturity at birth is linked to negative cognitive outcomes at school age (r^2 = 0.49; Bhutta et al., 2002). Another meta-analysis focusing on lexical and grammatical processing of very preterm children, assessed with standardized tests and parental questionnaires at preschool and school age, revealed that preterm scores were falling between 0.38 and 0.77 SD below those of control samples (Barre et al., 2011). Moreover, neurodevelopmental models of language processing propose that the interplay between early experience and brain maturation is crucial for an optimal language developmental trajectory (Werker & Hensch, 2015; Sanchez-Alonzo & Aslin, 2022). Therefore, identifying early markers for later language deficits or delays is crucial to better understand the long-term impact of preterm birth on core cognitive and language functions. One way to apprehend the likelihood of later cognitive and language deficits in preterm infants is to explore their early word-learning abilities as they can be linked to later language, cognitive and academic outcomes (Martinez-Nadal & Bosch, 2021; Marchman et al., 2022).

Regarding early speech perception abilities, electrophysiological data collected in newborns have shown that preterm birth can negatively impact the encoding of various speech features, such as vowel quality or voice identity (François et al., 2021; Mahmoudzadeh et al., 2017). Preterm infants before 10 months of postnatal age (i.e. listening experience have also shown poor higher-order linguistic abilities, such as perception and differentiation of

phonotactic properties and segmentation of continuous speech (Bosch, 2011; Gonzalez-Gomez & Nazzi, 2012; but see Berdasco-Muñoz et al., 2018 for positive evidence of monosyllabic segmentation already at 6 months postnatal age). At 24 months, studies have reported a lag on multiple measures of linguistic competencies, showing smaller expressive and receptive vocabularies or reduced use of complex grammatical structures (Kern & Gayraud, 2007; Sansavini et al., 2011; Vohr et al., 1988). Comprehension and production difficulties have often been observed until school age and even adolescence (Ortiz-Mantilla et al., 2008; see van Noort-van der Spek et al., 2010 for a meta-analysis). Crucially, the poor language abilities in childhood reported in preterm infants predict later academic under-achievement (Pritchard et al., 2009; Wolke et al., 2008). However, most of these results were obtained from infants born very preterm (i.e., born before 32 weeks of gestation). Despite these recent data showing a negative impact of moderate premature birth on cognition at age two (de Jong et al., 2015; Natarajan & Shankaran, 2016), we still know too little about language development after moderate-to-late prematurity (i.e., infants born between 32-34 and 34-36 completed gestation weeks, respectively), which corresponds to the greater number of preterm births worldwide (84 %; Shapiro-Mendoza & Lackritz, 2012; Gouyon et al., 2012). For instance, lower scores in cognition, receptive and productive communication can also be observed at age two in moderate preterm children compared to controls, although with small effect sizes (de Jong et al., 2015). A recent study focused on the long-term outcomes of late-to-moderate preterms (L-MPT) at school age. Of their L-MPT population, 24.3% had a borderline intelligence function and memory quotient, 32.4% had borderline executive functions, 13.5% had internalizing problems, and 66.7% had an abnormal score for at least one variable for attention deficit hyperactivity disorder (Jin et al., 2020). Further, previous studies on the long term neurodevelopmental outcomes of L-MPT infants evaluated at school age revealed an increased risk of developmental, language, and cognitive delays as well as poorer academic performance

(Hornman et al., 2017; Putnick et al., 2017; Vohr, 2013; Natarajan & Shankaran, 2016). Interestingly, a recent meta-analysis based on 16 studies involving late preterm children and control group assessments at preschool and school age explored cognitive and language functions as well as academic achievement (Martinez-Nadal & Bosch, 2021). Results revealed the presence of mild difficulties, pointing to suboptimal outcomes in areas such as executive function, short term verbal memory, literacy skills, attention and processing speed. Furthermore, this meta-analysis revealed the influence of maternal education and socio economic status on the scores obtained by late preterm participants rather than the gestational age factor per se. The results of this meta-analysis suggested that a combination of factors can negatively affect the catch up process in late preterm participants. Therefore, the close monitoring of L-MPT children's neurodevelopmental status and early referral of those demonstrating a neurodevelopmental delay is important (Vohr, 2013).

The ability to construct meaning from experience with the environment is a core aspect of early language acquisition that has been a hot topic for modern cognitive neuroscience, both from a developmental and a neurocognitive perspective (Gervain, 2015; Werker & Hensch, 2015). The emergence of robust word-referent associations and their storage in memory are crucial steps that infants must accomplish when acquiring their first receptive vocabulary. The fast association between new word-forms and their referents is a viable learning mechanism that may explain the rapid increase in word-learning rate during the first two years of life (Bion, Borowsky & Fernall, 2019; Waxman & Gelman, 2009). For example, 6- to 9-month-old infants already know the meaning of a few highly frequent words, suggesting that semantic knowledge can begin to emerge from these early word-object associations favored by daily experience in home contexts (Bergelson & Swingley, 2012; Bergelson & Aslin, 2017). The interplay between refined phonological processing and lexical knowledge has been studied in young infants using minimal-pair word-learning paradigms (Stager & Werker, 1997; Yoshida et al., 2009). Indeed,

the ability to distinguish between two similar-sounding words that differ by a single segment, thus forming a minimal pair (e.g., "moon" [mu::n] and "noon" [nu:n]), is fundamental for accurate lexical development. Because infants have a limited behavioral repertoire, tracking of gaze orientation and visual fixation measures can be instrumental in assessing the early development of lexical learning (Fennell & Werker, 2003; Ellis et al., 2015). Different versions of the inter-modal preferential looking paradigm (or "looking-while-listening" procedure) have been developed to study linguistic skills in infants and toddlers (Fernald, Zangl, Portillo & Marchman, 2008; Fernald, Pinto, Swingley, Weinberg & McRoberts, 1998; Fernald & Marchman, 2008; Bailey & Plunkett, 2002; Ballem & Plunkett, 2005). Importantly, this procedure has been successfully used to study the speed of word recognition (Fernald et al., 1998; Marchman & Fernald, 2008; Hurtado et al., 2008) or the development of early phonological representations (Swingley & Aslin, 2000; Fennell & Werker, 2003; Fennell, 2012; Yoshida et al., 2009).

Here, we explored toddlers' sensitivity to a minimal phonological distinction in a word-learning task using a variant of the looking-while-listening procedure originally used in previous studies (Fernald et al., 2008; Yoshida et al., 2009) and that had been successfully applied in our lab (Ramon-Casas et al., 2013; Ramon-Casas, Fennell & Bosch, 2017). Specifically, we collected visual fixation measures to target and distractor elements, representing a minimal pair of disyllabic CVCV pseudo-words contrasting in the voicing of the initial stop consonant (i.e., ['bali] - ['pali]) in healthy full-terms (FT) and healthy L-MPT tested at 18 months of corrected age for gestation. The perception of a voicing distinction in stop consonants relies on a fine-grained temporal speech feature, the voice onset time (VOT), which represents the temporal gap between the noise burst produced at consonant release and the onset of the waveform periodicity associated with vocal cord vibration (Lisker & Abramson, 1967). In Romance languages, such as Catalan and Spanish, the voicing contrast is

signaled by presence or absence of vocal fold vibration during the consonant constriction, that is, long voicing lead vs. lack of voicing lead or short positive VOT values (Solé, 2007). Certainly, the implementation of stop voicing contrasts in different languages can be affected by different factors, related to the position of the segments (onset, prevocalic, intervocalic), and also speaking rate (Solé, 2007). However, studies on Romance languages suggest that negative VOT values for voiced stops increase at slow speech rate, while VOT values for voiceless stops are not affected. Regarding our stimuli it is important to recall that they were produced with a rather similar slow speaking rate, uttered as isolated pseudowords and the contrast was in onset position always followed by the same vowel. These characteristics constrain the range variability that could be attributed to speaking rate.

The perception of a VOT distinction, present in many different languages, has been largely studied in human infants, children, and adults using behavioral and electrophysiological measures (Aslin et al., 1981; Elliot et al., 1986; Sharma & Dorman, 1999; Medina et al., 2010; Chobert et al., 2011, 2012, 2014). However, despite evidence showing that categorical perception is present early after birth for a VOT bilabial consonant contrast (Eimas et al., 1971), the VOT has been poorly studied in the context of minimal-pair word learning. While most of the minimal-pair word-learning studies performed in young infants have used consonantal contrasts involving differences in place of articulation (Stager & Werker, 1999; Yoshida et al., 2009; Fennell, Byers-Heinlein, Werker, 2007), only a small set of experimental and simulation studies have used voicing contrasts (Apfelbaum & McMurray, 2011; Pater et al., 2004; Rost & McMurray, 2009; 2010). For instance, Pater and colleagues (2004) explored associative word learning of minimal pairs differing in place of articulation (i.e., [bIn] - [dIn]), in voicing (i.e., [bIn] - [pIn]) and a combination of place and voicing ([phIn] - [dIn]) in three independent experiments conducted with 14-month-old infants. Results of the three experiments showed that infants failed to learn the word-object associations independently of the phonetic contrast

used. Thus, these three experiments showed that such contrasts could not be used to trigger learning of new object-label associations, suggesting that place of articulation and voicing contrasts may not yet be fully integrated into the phonological system at this age.

Besides, one problem regarding the location of the VOT boundary for different languages relies on the fact that the VOT continuum can be roughly divided into three categories, i.e., lead, short-lag, and long-lag VOTs, with different languages using different distinctions, not all of them being equally discriminated, as research in infancy has revealed. More specifically, the short-lag/long-lag VOT distinction (as found in English) does seem to be more easily discriminated by infants (likely to be more salient from an acoustic perspective), compared to the lead/ short-lag distinction characteristic of Spanish or Dutch, among other languages, representing a more subtle distinction, more perceptually demanding from the young learner's perspective (see Narayan, 2019, pp. 3-54, for a short synthesis of the inconsistent results relative to this type of VOT distinction). In the present minimal-pair word-learning study, the selection of a typical VOT contrast involving /b/ - /p/ segments is relevant for capturing possible differences between the groups under study. This contrast is frequent in the young learner's input language but its perceptual saliency is relatively low. Moreover, preterm birth implies an earlier exposure to the broad spectrum of frequencies, that is, at an earlier stage of brain maturation compared with the same type of exposure in infants born at term. This difference may impact the phonetic discrimination skills and phonological representation of words involved in early word-learning activities. If early exposure to the broad spectrum of frequencies before term age positively impacts phonological processing, such as the processing of fine-grained temporal speech cues, then we could expect L-MPT infants to exhibit a robust distinction of the novel words forming a minimal pair, with a longer proportion of looking times to target elements over distractors, compared to the behavior of the FT peers. Alternatively, suppose this early exposure negatively impacts L-MPT toddlers. In

that case, they should exhibit a shorter proportion of looking time to target elements over the distractor ones, significantly differing from the FT group. Importantly, because both groups were within the normal range for expressive vocabulary, we expected both groups to show no differences in the familiar word recognition measures.

METHODS

Participants

A total of 36 healthy infants completed the experiment with 18 healthy full-term infants (7 males; mean gestational age [GA] at birth = 39.6 weeks [37-42]; mean birth weight = 3195 g [2600-4200]; Apgar-5 > 9; mean chronological age at evaluation = 547 days ± 7; mean corrected age for gestation at evaluation = 547 days ± 7, 100% Caucasian) and 18 healthy late-to-moderate preterm infants (11 males; mean GA= 33.7 weeks [32-35.5]; mean birth weight = 1956 g [1625-2700]; Apgar-5 > 9; mean chronological age at evaluation = 592 days ± 18; mean corrected age for gestation at evaluation = 544 days ± 12, 100% Caucasian). Seven additional participants (4 FT and 3 L-MPT) had to be excluded from the final sample due to incomplete data collection. The mean age at evaluation (corrected age in the L-MPT group) was not significantly different between the two groups (t_{34} = 0.62, p = .54, two-tailed). However, the mean chronological (postnatal) age (t_{34} = -10.11, p < .0001, two-tailed), the mean GA at birth (t_{34} = 13.45, p < .0001, two-tailed) and the mean birth weight (t_{34} = 10.69, p < .0001, two-tailed) were significantly different, as expected. Importantly, none of the preterms fell in the small for gestational age category known to exhibit developmental deficits (Hayes & Sharif, 2009). Participants were recruited from the local urban area of Barcelona city using an existing database of interested families. Data collection took place at the APAL Infant Lab at the Hospital Sant Joan de Déu in Barcelona (Spain) between May 2015 and August 2016. Regarding the language status of the participants, information was gathered by using a

language exposure questionnaire (Bosch & Sebastián-Gallés, 2001), yielding an estimate of the language/s regularly used by caretakers in their environment (i.e., percentage of time infants were exposed to each language). This information was especially necessary to identify the most frequently used language within the family and daycare center for infants growing up in bilingual contexts (Spanish or Catalan) and select the corresponding version of the word learning task. Following Byers-Heinlein (2015) and Bosch (2024), the criterion that was applied to consider participants as belonging to monolingual or bilingual environments was the most usual in infancy research with bilingual groups, neither too inclusive, nor extremely conservative, that is, 100 % - 80 % was the range of exposure to the main language for monolingual infants, and when exposure to the most frequent language in their environment was below 80%, participants were classified as bilinguals. The range of exposure to the two languages in their environment fluctuated between 50% - 50% (rather infrequent) and a more unbalanced distribution (more usual), with 78 % - 22 % as the maximum range difference of exposure to the languages for infants in bilingual contexts. Importantly, the proportion of bilingual participants in each group was quite similar. In the FT group N = 11 participants were being raised in monolingual family/daycare environments (i.e., ≥ 80 % regular and daily exposure to a single language; range 80-99 %), and the remaining N = 7 were growing up in Catalan-Spanish bilingual contexts (between 50% - 50% and 35% - 65% of regular and daily exposure to their two languages). According to the more frequent language of exposure identified from the questionnaire, N = 12 participants were tested in the Catalan version of the task and N = 6 in the Spanish one. In the L-MPT group, similar criteria were applied and the distribution was N = 12 and N = 6 monolingual (range 80-93 %) and bilingual (between 40 % - 60 % and 22 % - 78 %), respectively. Nine of the participants were tested using the Catalan version, and the other 9 were tested in the Spanish one. Parents were informed and signed a consent form at the beginning of the experimental session. The hospital ethics committee

approved our procedures and protocols (CEIC-PIC-69-13). At birth, all infants had normal hearing from the universal screening test (automated auditory brainstem response) and normal examination made by a neonatologist at the delivery ward.

Stimuli

Several tokens of the non-words /bali/ and /pali/, produced by a female bilingual Catalan-Spanish native speaker, were recorded as independent elements in infant-directed speech using four different intonation patterns (flat low, flat high, ascending, and descending). Acoustical measures confirmed that the /b/ - /p/ consonants were contrastive, with no overlap in their VOT values (see **Figure 1**). Even though two versions of the paradigm were created with Spanish and Catalan elements (specifically concerning known words or familiar items, all cognates, and a verbal form in one of the carrier questions), the minimal pair tokens were all selected from the recording of the same female bilingual speaker in which no specific language difference was evident. As previously mentioned, based on descriptive phonetics of Catalan and Spanish languages and the properties of the voicing contrast in bilingual speakers (Amengual, 2012; Fernandez Planas, 2005), one should not expect different VOT production values for the /b/ - /p/ contrast in the novel word tokens for the Spanish and Catalan versions of the experiment. Six tokens of each word were selected and used during habituation trials, and four tokens, different from those used in the habituation phase, were selected for the test phase. The mean VOT values (in ms) for the onset consonants (/b/ and /p/) were -156 ms (SD = 23.5) and +12.3 ms (SD = 3.5) respectively. Nine words to be used as familiar items were recorded in both languages. One of them was to be used only in the initial warm-up trial, and the other eight were to be presented in familiar word recognition trials for the test phase. The criteria for selecting these words were their presence in toddlers' vocabulary checklists and their cognate status in Spanish and Catalan. Words and their corresponding images had been

successfully used in previous studies in our lab. These items corresponded to the following objects: *spoon* (Spanish and Catalan form, respectively: "**cuchara**", "**cullera**"); *ball* ("**pelota**", "**pilota**"); *cow* ("**vaca**", "**vaca**"); *car* ("**coche**", "**cotxe**"); *fish* ("**pez**", "**peix**"); *biscuit* ("**galleta**", "**galeta**"); *cat* ("**gato**", "**gat**"); *pencil* ("**lápiz**", "**llapis**"). Unlike novel words, which were always presented in isolation, familiar words were presented in a sentence context, similar in Spanish and Catalan, alternating between *Ball!..., Look, a ball!* ("***Pelota/Pilota!..., Mira, una pelota/pilota!***") or *Cow!..., do you see a cow? (*"**Vaca/Vaca... ¿Ves/Veus una vaca?"**), following the procedure implemented in Yoshida, Fennell, Swingley & Werker (2009). Visual stimuli were colorful pictures presented in slow motion on a black background.

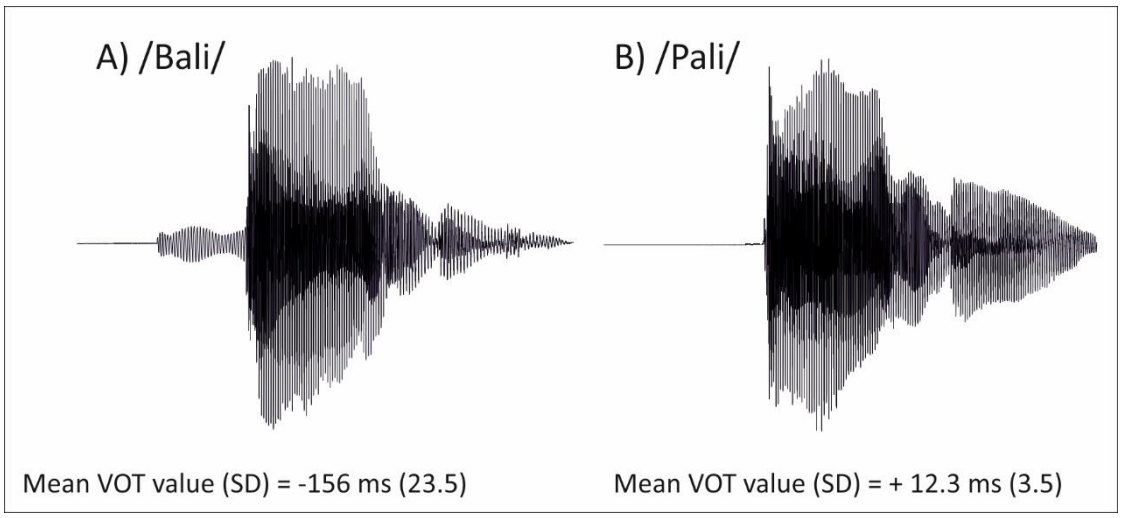

Figure 1: Illustration of one pair of novel-word tokens used in the present experiment showing /bali/ (A) and /pali/ (B). Mean VOT values in ms and SD of the total set of tokens used in the experiment are shown below.

Procedure

Testing took place in a soundproof room of the APAL Infant Lab at the Hospital Sant Joan de Déu in Barcelona. Infants were seated on their parents' laps in front of a 50" television monitor at 130 cm from the screen. The experimental procedure involving a habituation phase followed by a visual preference test phase was adapted from the task initially developed by Yoshida and colleagues (2009), which includes features of the intermodal preferential-looking

paradigm suitable to test minimal pair word learning in young toddlers (Fernald et al., 1998; see **Figure 2** for an illustration of the procedure).

The habituation phase began after a warm-up trial presenting a familiar object being named. Each trial presented one of the two novel word-object pairs with the object slowly moving horizontally on the screen, from side to side, while the corresponding label was being played. Participants were exposed to blocks of 4 habituation trials (containing two trials of each novel word-object pairing in any of these sequences: ABAB, BAAB, BABA, ABBA) until they reached the habituation criterion: either an attention decrease of 60% of that of the longest previous block or after a maximum of 24 trials (**Figure 2**). Each habituation trial lasted approximately 20 seconds and contained six different tokens of one of the target novel words, each token produced by the same voice with a different prosodic contour.

The test phase comprised 16 trials organized in blocks of 4 trials, presenting two objects side-by-side moving vertically in synchrony. Two types of trials were used: newly learned word-object pairs (novel-*word learning trials*) and familiar word-object pairs (familiar-*word recognition trials*) presented in a pseudo-random order, beginning with two familiar-word recognition trials first, followed by two word-learning trials, with the remaining trials randomly ordered. The presentation side for the first familiar-word recognition and novel-word learning trials was counterbalanced. The experimenter controlled the experiment from an adjacent room using Habit 1.0 software (Cohen, Atkinson & Chaput, 2004). The entire session was video recorded. Participants' looking time to target and distractor objects on each trial was coded offline on a frame-by-frame basis by an expert coder familiarized with the coding task but blind to the experimental conditions. A second blind coding by an independent coder was also done, and looking times were compared to check for any gross inconsistencies. After their identification, a third checking was done until an agreement was reached or the trial was discarded. Following this step, a Pearson correlation of test trial looking time measures from

the two independent offline coders was calculated and yielded significant results [$r(34) = .987$, $p < .001$] based on 4 randomly selected trials per participant. We ensured that only known associations (familiar words) were considered in the analyses by asking the parents to fill out an expressive vocabulary checklist (López-Ornat et al., 2005) and confirm known words from a short list of receptive vocabulary items, including toys, animals, vehicles, and food items.

Vocabulary and cognitive measures

Parents filled out the expressive vocabulary checklist from the Spanish version of the MacArthur-Bates Communicative Development Inventory (MCDI; López-Ornat et al., 2005) and a Catalan translation of the checklist used in the Lab. This questionnaire provided a parental estimate of expressive vocabulary size, including words produced in any of the bilingual's two languages and excluding translational equivalents (i.e., conceptual vocabulary). Participants were also tested using the Cognitive and Language subscales of the Bayley Scales of Infant and Toddler Development (BSID-III; Bayley, 2006). The Language subscale composite score resulted from the sum of the receptive and expressive scaled scores, and it represented a standardized measure of toddlers' receptive and expressive language development. The MCDI and the BSID-III Language composite score were positively correlated [$r(34) = 0.65$, $p < .001$], thus limiting the presence of parental biases in how they reported toddlers' spontaneous word productions. A general measure of Cognitive development was obtained to ensure that all infants in the sample fell within the normal range. The Cognitive and Language subscales of the BSID-III and the MCDI questionnaire were administered in a room of the APAL Infant Lab at the Hospital Sant Joan de Déu.

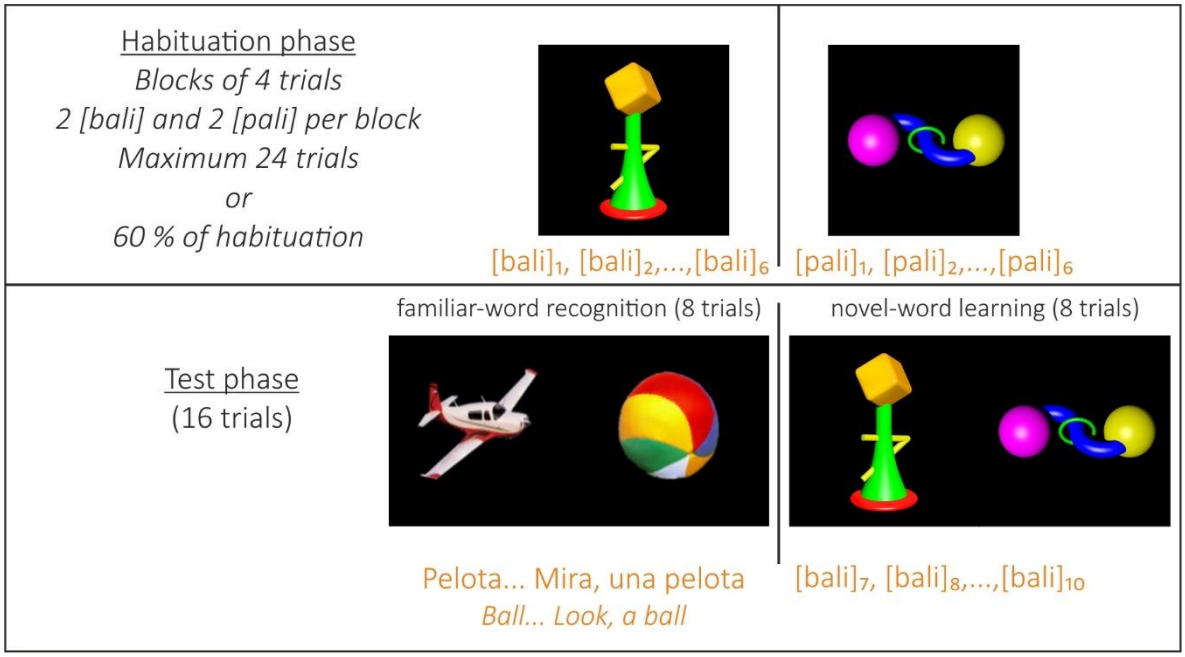

Figure 2. Illustration of the experimental procedure with the habituation and test phases. After the habituation phase, during which infants learned to associate two pseudo-words forming a minimal pair with two unfamiliar objects, infants were tested on their ability to recognize familiar word-object pairs (left panel: *familiar-word recognition trial*) and the newly learned associations (right panel*: novel-word-learning trial*).

Statistical analyses

We used the mean Proportion of Looking time to Target (PLT) as the dependent variable: Looking to Target Object ÷ (Looking to Target object + Looking to Distractor Object). We extracted the PLT within a standard time window, beginning 360 to 2000 ms after the onset of the target word (Swingley & Aslin, 2000, 2002). First, we compared the average PLTs across conditions to chance level (50%) separately for novel-word learning and familiar-word recognition trials. Second, we used generalized mixed-effect models with beta family function to evaluate the experimental measures (mean PLT), comparing FT and L-MPT groups across conditions (novel-word learning and familiar-word recognition). In each model, mean PLT was defined as the dependent variable, and random intercepts for participants were included to allow for differences in performance on individual participants. Fixed effects included Group (FT and L-MPT), condition (novel and familiar words), age, linguistic background, gender, and language of test. Since the PLT measure ranged from 0 (no looking

at target at all) and 1 (looking at the target the whole trial), we built beta-regression models, considering as control variables the above-mentioned fixed effects except for Group, which was defined as the variable of interest. Thus, the null model was defined by all the control variables, and the full one also considered the variable of interested. We assessed model performance using the Akaike Information Criterion (AIC) and the Likelihood Ratio Test (LRT) to determine the most effective model in explaining the data's variability (see **Table 1**). If necessary, post-hoc analyses were performed through z-test comparisons to identify significant differences among levels, with the application of the False Discovery Rate (FDR) for p-value correction. Both z-ratio and corrected p-values (p_{adj}) were reported. Importantly, we considered the interaction condition by language of test variables in the null model, to account for the possibility that some stimuli in one condition (e.g. novel words) could be more difficult in Catalan than in Spanish (or vice-versa). Finally, the link between experimental measures and linguistic outcomes was explored with Pearson correlations.

RESULTS

Language and cognitive measures

For FTs, the mean productive vocabulary size from the MCDI parental questionnaire was 45 words (median = 34; SD = 31.99; range: 13-110). Cognitive and language attainment assessed with the BSID-III yielded mean standardized scores of 112 (SD = 8.26; range: 90-125) and 104 (SD = 9.6; range: 89-121) for the cognitive and language subscales, respectively. For L-MPT, the mean productive vocabulary size from the MCDI was 36 words (median = 26; SD = 29.26; range: 8-114). Cognitive and language attainment yielded mean standardized scores of 108 (SD = 8.04; 90-125) and 99 (SD = 8; range: 83-112) for the Cognitive and Language subscales, respectively. All these measures were within the normal range (100 +/− 15). No between-group differences were found in the BSID-III for the Cognitive subscale [t_{34} = 1.53, p = .135, two-tailed, Cohen's d = 0.51], nor in the mean productive vocabulary size [t_{34} = 0.88, p = .388, two-

tailed, Cohen's d = 0.29]. Nonetheless, we found a slightly significant difference in the BDSID-III Language subscale between the two groups [t_{34} = 2.05, p = .049, two-tailed, Cohen's d = 0.68; see **Figure 3**].

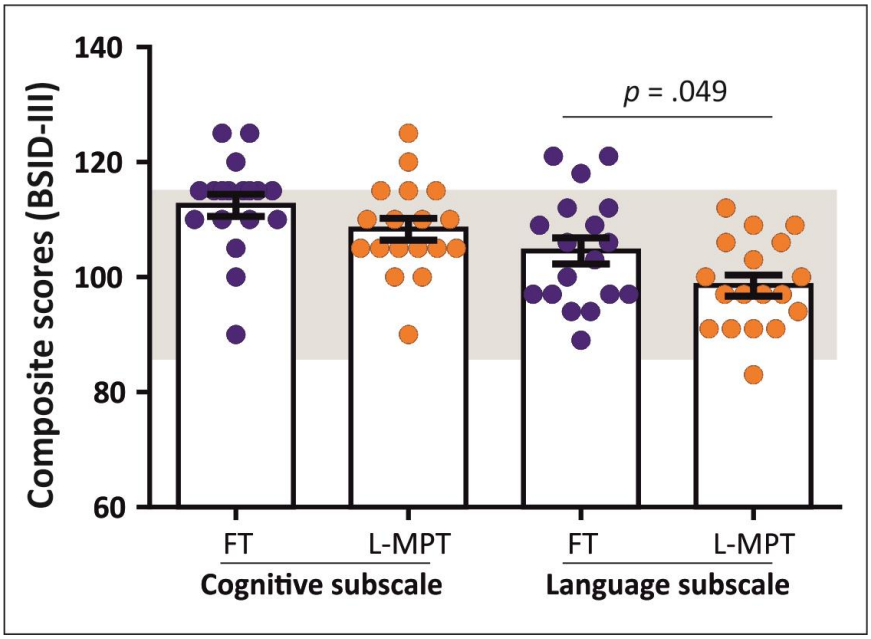

Figure 3. Distribution of the composite scores from the BSID-III Cognitive and Language subscales in full-term (FT) and late to moderate preterm toddlers (L-MPT). The dots represent each individual composite score. The mean value and standard error of the mean for each Group are also shown. The grey area indicates normative scores (85-115). Notice the significant between-groups difference found for the Language subscale only.

Looking time measures

First, Figure 4 shows the number of trials needed to reach the habituation criterion (i.e., either an attention decrease of 60% of the longest previous block or after a maximum of 24 trials) during the habituation phase in the two groups of infants. While FTs needed an average number of 15.4 trials (SD = 5.5; range: 7-24), L-MPTs needed 14.3 trials (SD = 5.6; range: 6-24) to reach the habituation criterion with no difference between the two groups [t_{34} = 0.57, p = .571]. This result suggests that both groups had a similar level of attention during this phase.

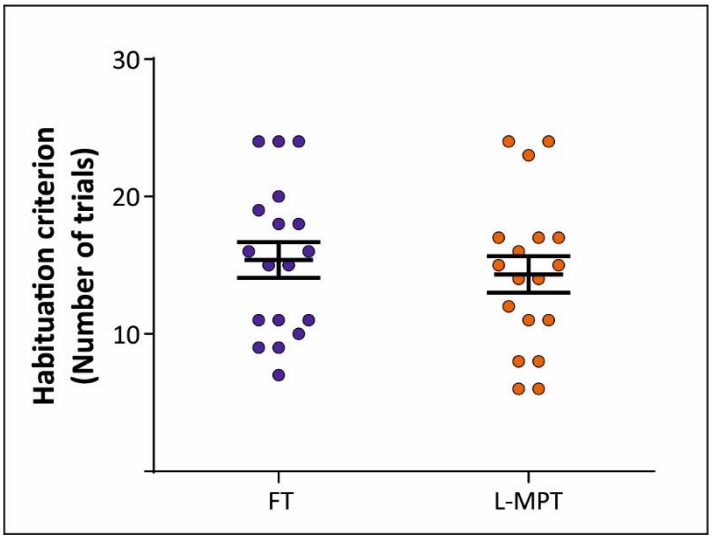

Figure 4. Distribution of the number of trials needed to reach the habituation criterion for full-term (FT) and late-to-moderate preterm infants (L-MPT). The dots represent each infant's number of trials needed to reach the criterion. The mean and standard error of the mean for each Group are also shown. No between-group difference was found.

Figure 5 displays the looking behaviour (i.e., mean PLT) for both novel-word-learning and familiar word-recognition trials in the two groups of participants. As can be seen, infants recognized the familiar word-object associations at the group level. For these familiar-word-recognition trials, the PLT was significantly higher than chance level for FTs [57.7 %; t_{17} = 2.86, p = .01; Cohen's d = 1.38] and L-MPT [59.1 %; t_{17} = 3.28; p = .004; Cohen's d = 1.59].

For novel-word-learning trials, the mean PLT to targets was significantly above chance level in FTs [M = 59.4 %; t_{17} = 3.25, p = .005; Cohen's d = 1.57], with thirteen infants looking more than 50% of the time to the correct object. By contrast, the PLT to targets in L-MPT was not significantly different from chance [M = 50.7 %; t_{17} = 0.36, p = .72; Cohen's d = 0.17], with only eight infants looking more than 50% of the time to the correct object. These results were confirmed by the beta-regression models, with the final model (see **Table 1**) showing a significant Group by Condition interaction [$\chi^2(1) = 5.59$; p = .0181, AIC = -100.6, LL = 61.3, $\eta_p^2 = 0.13$]. Post-hoc *t*-tests comparing the two groups showed that the mean PLT in the novel-word-learning trials was significantly higher for FTs than for L-MPT (z = 2.41, p_{adj} = .0322, Cohen's d = 0.80). This was not the case for familiar-word-recognition trials (z = -0.45, p_{adj} =

.6528, Cohen's d = -0.15). Therefore, these results show that 18-month-old FTs successfully learned the minimal pairs, as indicated by their greater than chance looking time to the targets, while L-MPT, as a group, did not.

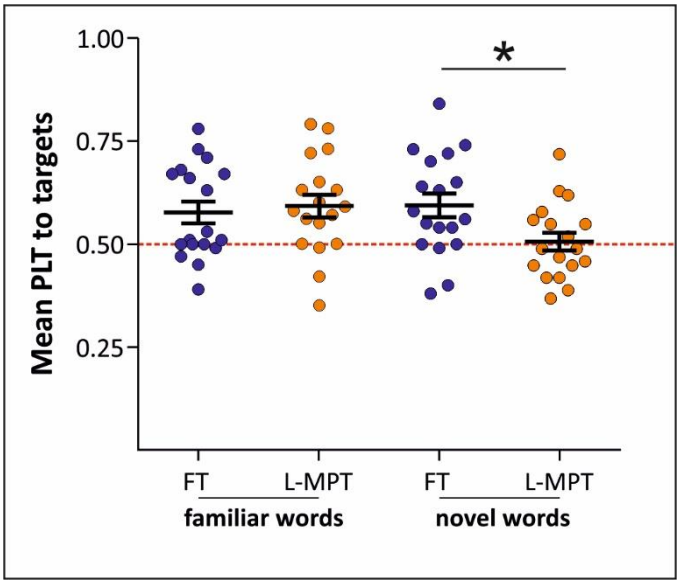

Figure 5. Distribution of the mean proportion of looking-time (PLT) scores to targets in familiar-word recognition and novel-word-learning trials for full-term (FT) and late-to-moderate preterm infants (L-MPT). The dots represent each infant's mean PLT score. The mean PLT score and standard error of the mean for each Group are also shown. The red dotted line indicates chance level. Notice the significant between-group difference found for novel-word-learning trials only.

Exploratory correlation analyses

We explored the relationship between measures of looking time in the experimental task (mean PLT) and MCDI scores using Pearson correlations. We found a significant positive correlation between the mean PLT to targets for novel-word-learning trials and expressive vocabulary measured by the MCDI in the Group of L-MPT [r(16) = 0.48, p = .04; see **Figure 6**] but not in FTs [r(16) = -0.29, p = .24].

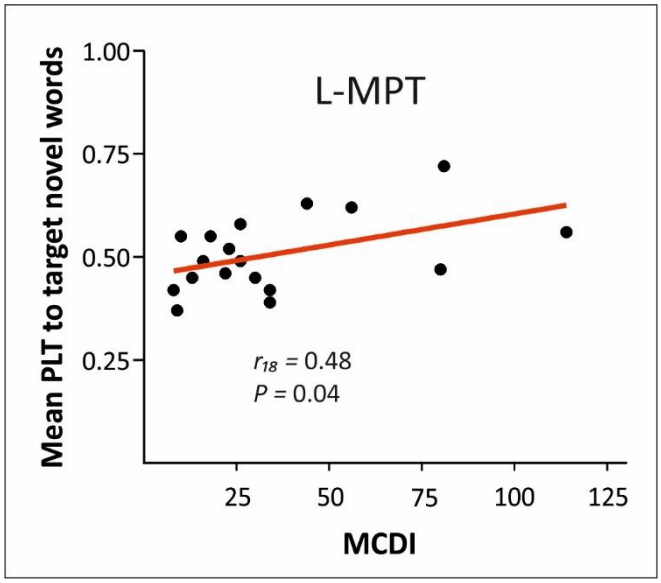

Figure 6. Association between mean PLT scores to novel words and expressive vocabulary measures obtained from the MCDI in the L-MPT Group of 18 mo. (N=18). The scatterplot shows the significant correlation between the raw MCDI scores and the mean PLT scores for novel words in the Group of L-MPT.

DISCUSSION

Here, we provide new evidence on the negative impact that moderate premature birth can have on basic associative word-learning skills involving fine-grained phonological representations. Specifically, we compared healthy late-to-moderate preterm 18-month-old toddlers matched in maturational age with participants born at term. We used a reliable and frequently used behavioral paradigm to measure novel word-object associative learning at an early age to characterize better the impact of late-to-moderate prematurity on the development of a core language learning function. While familiar word recognition was preserved, we observed a negative impact of late-to-moderate premature birth on the ability to learn minimal pairs differing in VOT, a somewhat challenging feature here involving the lead/short lag distinction characteristic of the native language of the participants. Specifically, when we compared the average proportion of looking time to target with chance level, FTs displayed a preference towards the target with a large effect size (d = 1.57), whereas L-MPTs exhibited no preference with a small effect size (d = 0.17). Further, LMM analysis revealed that the mean PLT in the novel-word-learning trials was significantly higher for FTs than for L-MPT and

showed a moderate effect size (η_p^2 = 0.07). Moreover, a significant association was found between looking-time behavior in novel word-learning test trials and expressive vocabulary measures in the group of L-MPT infants. Even considering the small sample of participants in the present study, these results refine our knowledge of the mechanisms that may hinder the deployment of early word-learning skills after premature birth.

While full-term infants showed significantly higher than chance mean looking time to targets in novel-word learning trials, preterm infants could not differentiate them from distractors. This pattern of results suggests that FTs correctly learned and recognized the minimal pairs based on VOT differences, extending prior literature on minimal pair word learning using other versions of the looking-while-listening paradigm and testing different contrasts (Werker et al., 1998; Fennell, 2012; Höhle et al., 2020). The finding that L-MPT infants could not learn and recognize the minimal-pair target words not only fits well with previous studies showing that preterm infants may have suboptimal speech-processing abilities (François et al., 2021; Paquette et al., 2015; Peña, Werker, & Dehaene-Lambertz, 2012), but it further suggests that late-to-moderate preterm birth may negatively impact associative word-learning, especially when the properties of stimuli are too demanding (i.e., coping with similar sounding labels to encode the names of visually distinct objects).

One could argue that the lack of learning in late-to-moderate preterm infants may be related to different factors: a slower speed in information processing, attention deficits that they often exhibit, or even the task itself (Marchman, Adams, Loi, Fernald, & Feldman, 2016; Ramon-Casas, Bosch, Iriondo, & Krauel, 2013). Previous reports in healthy FT infants revealed successful minimal-pair word learning abilities at 17-18 months of age (Werker et al., 2002), while 14 month-old infants may not always be able to learn similar sounding novel words (Stager & Werker, 1997; Pater et al., 2004; Rost & McMurray, 2009; 2010). Here, we used the corrected age for gestation to match the L-MPT group to the FT group, so all infants were older

than 17 months. In other words, both groups were tested at a mean age of 18 months, even though the mean postnatal age of the L-MPT group was actually 20 months if correction for gestation had not been applied. Therefore, both groups should already show a fairly established capacity for minimal-pair associative word learning based on their age at testing, whether chronological (in FT) or corrected for gestation (in L-MPT). Crucially, both groups of infants recognized familiar word-object pairs with a mean proportion of looking time to targets higher than chance in the familiar word recognition trials. Besides, the number of trials needed to reach the habituation criterion was not different between the two groups, suggesting that no clearly distinct pattern of attention between groups was evident during the habituation phase. Therefore, even late-to-moderate preterm infants could correctly perform the task with enough attention directed toward the stimuli, favoring the view that the lack of learning was not due to fatigue during the initial phase of habituation. This result may not be surprising for FTs, considering previous studies showing that highly frequent familiar words are already known as early as by 6-9 months of age (Bergelson & Swingley, 2012). But what the pattern of looking behavior of the group of L-MPT participants reflects (in novel-word learning test trials) is precisely that the late-to-moderate premature birth can negatively impact basic associative mechanisms involving fine-grained phonological representations, that is, dealing with the processing of fine phonetic detail, encoding a contrastive phonological representation while connecting these minimally distinct labels to novel but clearly distinct objects.

A number of factors must be taken into account in order to gain a better understanding of the challenges of the task for L-MPT infants. More specifically, one relevant factor that needs to be considered in interpreting the failure from the L-MPT group of infants in minimal-pair word-learning might be related to the range of variability present in the novel items to be learned. Interestingly, recent data suggest that talker variability during infants' word learning situations may facilitate the discrimination of consonants and early word representation (Rost

& McMurray, 2009; Quam, Clough, Knight, Gerken, 2021; Bulgarelli & Bergelson, 2022; Höhle et al., 2020). Here, we used stimuli recorded from one single speaker but with different prosodic contours to introduce variability. This reduction of variability to just one factor (i.e., prosody, but not voice and prosody), might have precluded a more direct capture of the VOT difference, making the task more demanding for the L-MPT group. However, it is also possible to consider that this more limited range of variability could have been an advantageous property for preterm participants, although it was not. Further studies involving premature infants may consider this factor to understand better the novel word learning deficits observed in our group of late-to-moderate moderate preterms.

From a different perspective, although still exploratory, the positive association between looking behavior in novel word-learning test trials and expressive vocabulary in the L-MPT group deserves some comments. The finding that a larger expressive lexicon was associated with a higher proportion of looking time to target elements in the novel word-learning trials is in line with a pattern of association that has been recently reported, showing that eye-tracking-based measures of familiar-words' lexical processing are positively associated with expressive vocabulary in 18-month-old very preterm infants (Ståhlberg-Forsén et al., 2022). Importantly, language processing at 18 months can predict the number of words produced and vocabulary learning abilities over the second and third years of life (Fernald & Marchman, 2012; Law & Edwards, 2015; Peter et al., 2019). Further, this association is consistent with previous data on lexical development involving larger samples of preterms and showing significant correlations between gestational age at birth and language delays in terms of vocabulary size and grammatical processing (Foster-Cohen et al., 2007; Kern & Gayraud, 2007). Overall, our results suggest that expressive vocabulary is also somehow linked to basic associative word learning skills involving fine-grained phonological representations.

Our study has some methodological limitations that must be considered. First, our sample size was rather modest to solve a group by condition interaction which may decrease the statistical power. Further studies with larger cohorts are needed to understand the impact of preterm birth on associative word learning skills. Second, a bilingual background was inevitably present to a greater or lesser extent in the linguistic environment the participants were growing up. This fact resulted in a certain heterogeneity between the groups regarding the exposure to the home/daycare languages. A recent meta-analysis of 141 associative word learning studies based on the Switch task paradigm (Werker et al., 1998) revealed that, on average, infants raised in bilingual environments at home outperformed monolingual peers, suggesting a bilingual advantage in learning word-object associations for phonetically distant words and when facilitative experimental manipulations are used (Tsui et al., 2019). Results from a minimal-pair word learning study in 18 month-olds based on a vowel difference had also suggested a bilingual advantage in the task (Singh et al., 2015). Here, the proportion of bilinguals was quite similar in the two groups and this factor did not seem to modulate the pattern of looking behavior. However, more work with larger and more linguistically controlled groups is needed to better understand the interaction between preterm birth and linguistic environment with associative word learning abilities.

Third, infants performed the looking-while-listening experiment with a direct presentation of the novel objects without any interactive context, with no additional visual or gestural cues from any visible speaker. Such a situation may be considered artificial and may have lacked ecological validity, even though we did our best to provide a friendly setup. Indeed, recent data have already shown the importance of multisensory cues to trigger word learning in infants and children (Ramos-Escobar et al., 2021; Seidl, Indarjit, Borovsky, 2023). In this sense, more precise and time-resolved measures based on recent devices such as eye-tracking glasses can be more appropriate for studying young infants' learning during "real" interactions

with caregivers or experimenters (Tomalski et al., 2022), thus offering a more ecologically valid methodological approach in this type of study (Thiele et al., 2021). Fourth, we explored looking-time behaviors using offline frame-by-frame coding of videos which may not give a detailed perspective of the looking behaviors during the test phase. In this sense, eye-tracking measurements gathered in "looking-while-listening" paradigms can offer additional information such as the fine-grained time course of infants looking behaviors associated with early lexical processing (Ståhlberg-Forsén et al., 2022). Fifth, visual fixation measurements during the habituation phase required the active involvement of the experimenter. Again, eye-tracking setups can be instrumental in automatizing the experiment by offering a more direct way to compute the attention behavior in this type of task (Shukla, Wen, White, Aslin, 2022). Additional measures from the habituation phase could reveal subtle between-group differences that might shed light on the time-course of the minimal-pair word learning.

Finally, it is important to comment on the difference between novel word-learning trials and familiar word-recognition trials in the test phase. While novel word-learning trials involved the presentation of tokens of the pseudo-words produced in isolation, the familiar word-recognition trials included carrier sentences (e.g., Look, a ball!), which might have contributed to the differences in the novel word-learning trials. But on the other hand, the similarity between naming the novel objects in the habituation phase and the presentation of the labelling in the test phase could also be seen as a feature favoring the correct identification of the target elements in test trials. Overall, what must be emphasized here is that the format difference between familiar word-recognition trials and novel word-learning trials may not be sufficient to explain the between-group significant difference we observed as both groups were assessed with the same experimental procedure.

CONCLUSION

Minimal-pair word learning can be a useful experimental tool to better understand the impact of premature birth on core linguistic skills before the emergence of higher-order cognitive functions during early development. Our data provide additional evidence for the critical role that premature birth may have on the interplay between early phonetic perception and the development of core associative learning mechanisms required for successful language development. Thus, the study of moderate to late preterm infants and toddlers can inform neurodevelopmental models of language processing by providing relevant information focusing on the interplay between early experience and brain maturation required for optimal language development.

ACKNOWLEDGMENTS

We thank all the families and their babies for participating in the study. We also thank D. Cucurell, M. Teixidó, M. Ramon-Casas, J. Sánchez-Galán, and J. Solé for their help in various aspects of the study. Finally, we thank the three anonymous reviewers for their constructive comments on previous versions of the manuscript.

FUNDING INFORMATION

This research was supported by a Spanish MINECO project (PSI2015-69132P) and the French National Agency for Research (ANR), "BabyLang project" (ANR-20-CE28-0017) to CF. It has also been supported by Spanish MICINN projects (PGC 2018-097487-B-100 and PID 2021-128159NB-100) to L.B. This work carried out within the Institut Convergence ILCB (ANR-16-CONV-0002), has benefited from support from the French government, managed by the French National Agency for Research (ANR) and the Excellence Initiative of Aix- Marseille University (A ∗ MIDEX).

CONFLICT OF INTEREST STATEMENT

None declared.

DATA AVAILABILITY STATEMENT

The analyses presented here were exploratory. The anonymized data and analytic R codes necessary to reproduce the analyses presented here are available on the OSF platform (https://osf.io/p3856/). The materials necessary to attempt to replicate the findings presented here are not publicly accessible, but are available from the corresponding author upon request. The analyses presented here were not preregistered.

REFERENCES

Amengual, M. (2012). Interlingual influence in bilingual speech: Cognate status effect in a continuum of bilingualism. *Bilingualism: Language and Cognition*, *15*(3), 517-530.

Apfelbaum, K.S., McMurray, B. (2011). Using variability to guide dimensional weighting: Associative mechanisms in early word learning. *Cognitive Science*, 35(6):1105-1138.

Aslin, R.N., Pisoni, D.B., Hennessy, B.L., Perey, A.J. (1981). Discrimination of voice onset time by human infants: New findings and implications for the effects of early experience. *Child development*, 1135-1145.

Bailey, T.M., Plunkett, K. (2002). Phonological specificity in early words. *Cognitive Development*, 17(2), 1265-1282.

Ballem, K.D., Plunkett, K. (2005). Phonological specificity in children at 1; 2. *Journal of Child Language*, 32(1), 159-173.

Barre, N., Morgan, A., Doyle, L. W., & Anderson, P. J. (2011). Language abilities in children who were very preterm and/or very low birth weight: a meta-analysis. The Journal of pediatrics, 158(5), 766-774.

Bayley, N. Bayley Scales of Infant and Toddler Development. 3rd edn. San Antonio, TX: Harcourt Assessment Inc (2006).

Berdasco-Muñoz, E., Nishibayashi, L. L., Baud, O., Biran, V., & Nazzi, T. (2018). Early segmentation abilities in preterm infants. *Infancy*, 23(2), 268-287.

Bergelson, E., Swingley, D. (2012). At 6–9 months, human infants know the meanings of many common nouns. *Proc. Natl. Acad. Sci. U.S.A*, 109(9), 3253-3258.

Bosch, L. (2011). Precursors to language in preterm infants: speech perception abilities in the first year of life. *Progress in brain research*, 189, 239-257.

Bosch, L. (2024). Acquisition of segmental phonology in child bilingualism. In M. Amengual (Ed.) The Cambridge Handbook of Bilingual Phonetics and Phonology. C.U.P. ISBN 9781009098601

Bosch, L.; Sebastián-Gallés, N. (2001). Evidence of early language discrimination abilities in infants from bilingual environments. *Infancy*, 2(1), 29- 49.

Bhutta, A.T., Cleves, M.A., Casey, P.H., Cradock, M.M., Anand, K.J. (2002). Cognitive and behavioral outcomes of school-aged children who were born preterm: a meta-analysis. *JAMA*, 288(6), 728-737.

Bion, R. A., Borovsky, A., Fernald, A. (2013). Fast mapping, slow learning: Disambiguation of novel word–object mappings in relation to vocabulary learning at 18, 24, and 30 months. *Cognition*, 126(1), 39-53.

Bulgarelli, F., Bergelson, E. (2022). Talker variability shapes early word representations in English-learning 8-month-olds. *Infancy*, 27(2), 341-368.

Byers-Heinlein, K. (2015). Methods for studying infant bilingualism. In: Schwieter JW, ed. The Cambridge Handbook of Bilingual Processing. Cambridge Handbooks in Language and Linguistics. Cambridge University Press; Chapter 5: 133-154.

Chobert, J., Marie, C., François, C., Schön, D., Besson, M. (2011). Enhanced passive and active processing of syllables in musician children. *Journal of cognitive neuroscience*, 23(12), 3874-3887.

Chobert, J., François, C., Habib, M., Besson, M. (2012). Deficit in the preattentive processing of syllabic duration and VOT in children with dyslexia. *Neuropsychologia*, 50(8), 2044-2055.

Chobert, J., François, C., Velay, J.L., Besson, M. (2014). Twelve months of active musical training in 8-to 10-year-old children enhances the preattentive processing of syllabic duration and voice onset time. *Cerebral cortex*, 24(4), 956-967.

Cohen, L.B., Atkinson, D.J., Chaput, H.H. (2004). Habit X: A new program for obtaining and organizing data in infant perception and cognition studies (Version 1.0). Austin, TX: University of Texas.

Eimas, P.D., Siqueland, E.R., Jusczyk, P., Vigorito, J. (1971). Speech perception in infants. *Science*, 171(3968), 303-306.

Elliott, L.L., Busse, L. A., Partridge, R., Rupert, J., De Graaff, R. (1986). Adult and child discrimination of CV syllables differing in voicing onset time. *Child development*, 628-635.

Escudero, P., Kalashnikova, M. (2020). Infants use phonetic detail in speech perception and word learning when detail is easy to perceive. *Journal of experimental child psychology*, 190, 104714.

Fennell, C.T., Werker, J.F. (2003). Early word learners' ability to access phonetic detail in well-known words. *Language and speech*, 46(2-3), 245-264.

Fennell, C.T., Byers-Heinlein, K., Werker, J.F. (2007). Using speech sounds to guide word learning: The case of bilingual infants. *Child development*, 78(5), 1510-1525.

Fernald, A., Pinto, J.P., Swingley, D., Weinberg, A., McRoberts, G.W. (1998). Rapid gains in speed of verbal processing by infants in the 2nd year. *Psych. Sci.*, 9(3), 228-231.

Fernald, A., Zangl, R., Portillo, A. L., Marchman, V.A. (2008). Looking while listening: Using eye movements to monitor spoken language. *Developmental psycholinguistics: On-line methods in children's language processing*, 44, 97.

Fernald, A., Marchman, V.A. (2012). Individual differences in lexical processing at 18 months predict vocabulary growth in typically developing and late-talking toddlers. *Child development*, 83(1), 203-222.

Fernández Planas, A. M. (2005). *Así se habla: nociones fundamentales de fonética general y española: apuntes de catalán, gallego y euskara*. Barcelona: Honsori.

François, C., Rodriguez-Fornells, A., Teixido, M., Agut, T., Bosch, L. (2021). Attenuated brain responses to speech sounds in moderate preterm infants at term age. *Dev. Sci.*, e12990.

Foster-Cohen, S., Edgin, J.O., Champion, P.R., Woodward, L.J. (2007). Early delayed language development in very preterm infants: evidence from the MacArthur-Bates CDI. *Journal of child language*, 34(3), 655-675.

Gervain, J. (2015). Plasticity in early language acquisition: the effects of prenatal and early childhood experience. *Current opinion in neurobiology*, 35, 13-20.

Gouyon, J.B., Iacobelli, S., Ferdynus, C., Bonsante, F. (2012). Neonatal problems of late and moderate preterm infants. *Seminars in Fetal and Neonatal Medicine*, 17(3), 146-152.

Graf-Estes, K., Hurley, K. (2013). Infant-directed prosody helps infants map sounds to meanings. *Infancy*, 18(5), 797–824.

Gray, S. (2005). Word learning by preschoolers with specific language impairment: effect of phonological or semantic cues. *J. Speech, Lang. Hear. Res.*, 48, 1452-1467.

Hayes, B., Sharif, F. (2009). Behavioural and emotional outcome of very low birth weight infants–literature review. *The journal of maternal-fetal & neonatal medicine*, 22(10), 849-856.

Höhle, B., Fritzsche, T., Meß, K., Philipp, M., Gafos, A. (2020). Only the right noise? Effects of phonetic and visual input variability on 14-month-olds' minimal pair word learning. *Dev. Sci.*, e12950.

Hornman, J., de Winter, A. F., Kerstjens, J. M., Bos, A. F., & Reijneveld, S. A. (2017). Stability of Developmental Problems after School Entry of Moderately-Late Preterm and Early Preterm-Born Children. *The Journal of Pediatrics*, 187, 73-79.

Hurtado, N., Marchman, V.A., Fernald, A. (2008). Does input influence uptake? Links between maternal talk, processing speed and vocabulary size in Spanish-learning children. *Dev. Sci.*, 11(6), F31-F39.

Jansson-Verkasalo, E., Valkama, M., Vainionpää, L., Pääkkö, E., Ilkko, E., Lehtihalmes, M. (2004). Language development in very low birth weight preterm children: a follow-up study. *Folia Phoniatrica et Logopaedica*, 56(2), 108-119.

Jin, J. H., Yoon, S. W., Song, J., Kim, S. W., & Chung, H. J. (2020). Long-term cognitive, executive, and behavioral outcomes of moderate and late preterm at school age. *Clinical and Experimental Pediatrics*, 63(6), 219.

Kern, S., Gayraud, F. (2007). Influence of preterm birth on early lexical and grammatical acquisition. *First Language*, 27, 159-173.

Law, F., Edwards, J.R. (2015). Effects of vocabulary size on online lexical processing by preschoolers. *Language Learning and Development*, 11(4), 331-355.

Lisker, L., Abramson, A.S. (1967). Some effects of context on voice onset time in English stops. *Language and Speech*, 10(1), 1–28.

López-Ornat, S., Gallego, C., Gallo, P., Karousou, A., Mariscal, S., Martínez, M. (2005). Inventarios de desarrollo comunicativo MacArthur. Manual. Madrid, España: TEA, Ediciones.

Marchman, V.A., Ashland, M.D., Loi, E.C., Munevar, M., Shannon, K.A., Fernald, A., Feldman, H.M. (2023). Associations between early efficiency in language processing and language and cognitive outcomes in children born full term and preterm: similarities and differences. *Child Neuropsychology*, 1-20.

Marchman, V.A., Fernald, A. (2008). Speed of word recognition and vocabulary knowledge in infancy predict cognitive and language outcomes in later childhood. *Developmental science*, 11(3), F9-F16.

Mariani, B., Nicoletti, G., Barzon, G., Ortiz Barajas, M. C., Shukla, M., Guevara, R., ... & Gervain, J. (2023). Prenatal experience with language shapes the brain. *Science Advances*, 9(47), eadj3524.

Martínez-Nadal, S., Bosch, L. (2021). Cognitive and learning outcomes in late preterm infants at school age: A systematic review. *International Journal of Environmental Research and Public Health*, 18(1), 74.

Meteyard, L., & Davies, R. A. (2020). Best practice guidance for linear mixed-effects models in psychological science. *Journal of Memory and Language*, 112, 104092.

Narayan, C.R. (2019). An acoustic perspective on 45 years of infant speech perception, Part 1: Consonants. *Language and Linguistics Compass*, 13(10), e12352.

Natarajan, G., & Shankaran, S. (2016). Short-and long-term outcomes of moderate and late preterm infants. *American journal of perinatology*, 305-317.

Ortiz-Mantilla, S., Choudhury, N., Leevers, H., Benasich, A.A. (2008). Understanding language and cognitive deficits in very low birth weight children. *Dev. psychobiol*, 50(2), 107-126.

Paquette, N., Vannasing, P., Tremblay, J., Lefebvre, F., Roy, M.S., McKerral, M., Lepore, F., Lassonde, M., Gallagher, A. (2015). Early electrophysiological markers of atypical language processing in prematurely born infants. *Neuropsychologia*, 79, 21-32.

Peña, M., Mehler, J., Nespor, M. (2011). The role of audiovisual processing in early conceptual development. *Psych. Sci*. 22(11), 1419-1421.

Peña, M., Werker, J.F., Dehaene-Lambertz, G. (2012). Earlier speech exposure does not accelerate speech acquisition. *Journal of neuroscience*, 32(33), 11159-11163.

Peter, M.S., Durrant, S., Jessop, A., Bidgood, A., Pine, J.M., Rowland, C.F. (2019). Does speed of processing or vocabulary size predict later language growth in toddlers? *Cognitive psychology*, 115, 101238.

Pritchard, V.E., Clark, C.A., Liberty, K., Champion, P.R., Wilson, K., Woodward, L.J. (2009). Early school-based learning difficulties in children born very preterm. *Early Hum Dev.*, 85(4), 215-224.

Putnick, D. L., Bornstein, M. H., Eryigit-Madzwamuse, S., & Wolke, D. (2017). Long-term stability of language performance in very preterm, moderate-late preterm, and term children. *The Journal of pediatrics*, 181, 74-79.

Quam, C., Clough, L., Knight, S., Gerken, L. (2021). Infants' discrimination of consonant contrasts in the presence and absence of talker variability. *Infancy*, 26(1), 84-103.

Ragó, A., Honbolygó, F., Róna, Z., Beke, A., & Csépe, V. (2014). Effect of maturation on suprasegmental speech processing in full and preterm infants: A mismatch negativity study. *Research in Developmental Disabilities*, 35(1), 192–202.

Ramon-Casas, M., Fennell, C.T., Bosch, L. (2017). Minimal-pair word learning by bilingual toddlers: the Catalan/e/-/ɛ/contrast revisited. Bilingualism: *Language and cognition*, 20(3), 649-656.

Ramos-Escobar, N., Segura, E., Olivé, G., Rodriguez-Fornells, A., & François, C. (2021). Oscillatory activity and EEG phase synchrony of concurrent word segmentation and meaning-mapping in 9-year-old children. *Developmental Cognitive Neuroscience*, 51, 101010.

Rost, G.C., McMurray, B. (2009). Speaker variability augments phonological processing in early word learning. *Dev. Sci*., 12(2), 339-349.

Rost, G.C., McMurray, B. (2010). Finding the signal by adding noise: The role of noncontrastive phonetic variability in early word learning. *Infancy*, 15(6), 608-635.

Sanchez-Alonso, S., Aslin, R.N. (2022). Towards a model of language neurobiology in early development. *Brain and Language*, 224, 105047.

Sansavini, A., Guarini, A., Justice, L. M., Savini, S., Broccoli, S., Alessandroni, R., & Faldella, G. (2010). Does preterm birth increase a child's risk for language impairment? *Early human development*, 86(12), 765-772.

Sansavini, A., Guarini, A., Savini, S., Broccoli, S., Justice, L., Alessandroni, R., & Faldella, G. (2011). Longitudinal trajectories of gestural and linguistic abilities in very preterm infants in the second year of life. *Neuropsychologia*, 49(13), 3677-3688.

Seidl, A.H., Indarjit, M., Borovsky, A. (2023). Touch to learn: Multisensory input supports word learning and processing. *Dev. Sci.*, e13419.

Shapiro-Mendoza, C.K., & Lackritz, E.M. (2012). Epidemiology of late and moderate preterm birth. *Seminars in Fetal and Neonatal Medicine*, 17(3), 120-125.

Shukla, M., Wen, J., White, K.S., Aslin, R.N. (2011). SMART-T: A system for novel fully automated anticipatory eye-tracking paradigms. Behavior Research Methods, 43, 384-398.

Singh, L., Fu, C. S., Tay, Z. W., & Golinkoff, R. M. (2018). Novel word learning in bilingual and monolingual infants: Evidence for a bilingual advantage. *Child Development*, 89(3), e183-e198.

Stager, C. L., & Werker, J. F. (1997). Infants listen for more phonetic detail in speech perception than in word-learning tasks. *Nature*, 388(6640), 381-382.

Ståhlberg-Forsén, E., Latva, R., Leppänen, J., Lehtonen, L., Stolt, S. (2022). Eye tracking based assessment of lexical processing and early lexical development in very preterm children. *Early human development*, 170, 105603.

Swingley, D., Aslin, R.N. (2000). Spoken word recognition and lexical representation in very young children. *Cognition*, 76(2), 147–166.

Swingley D, Aslin RN. (2002). Lexical neighborhoods and the word-form representations of 14-month-olds. *Psych. Sci.,* 13, 480-484.

Thiele, M., Hepach, R., Michel, C., Gredebäck, G., Haun, D.B. (2021). Social interaction targets enhance 13-month-old infants' associative learning. *Infancy*, 26(3), 409-422.

Tomalski, P., Pérez, D.L., Radkowska, A., Malinowska-Korczak, A. (2022). Dyadic interactions during infant learning: Exploring infant-parent exchanges in experimental eye-tracking studies. *Infant Behavior and Development*, 69, 101780.

Tsui, A.S.M., Byers-Heinlein, K., Fennell, C.T. (2019). Associative word learning in infancy: A meta-analysis of the switch task. *Developmental psychology*, 55(5), 934.

van Noort-van der Spek, I.L., Franken, M.C., Wieringa, M.H., Weisglas-Kuperus, N. (2010). Phonological development in very-low-birth weight children: an exploratory study. *Dev. Med. Child. Neurol.*, 52(6), 541-546.

Vohr, B.R., Garcia Coll, C., Oh, W. (1988). Language development of low-birth weight infants at two years. *Dev. Med. Child. Neurol.*, 30(5), 608-615.

Vohr, B. (2013). Long-term outcomes of moderately preterm, late preterm, and early term infants. *Clinics in perinatology*, 40(4), 739-751.

Waxman, S.R., Gelman, S.A. (2009). Early word-learning entails reference, not merely associations. *Trends in cognitive sciences*, 13(6), 258-263.

Weismer, S.E., Evans, J.L. (2002). The role of processing limitations in early identification of specific language impairment. *Topics in Language Disorders*, 22(3), 15-29.

Werker, J.F., Cohen, L.B., Lloyd, V.L., Stager, C.L., Casasola, M. (1998). Acquisition of word-object associations by 14-month old infants. *Dev. Psychol.*, 34, 1289-1309

Werker, J. F., Fennell, C. T., Corcoran, K. M., & Stager, C. L. (2002). Infants' ability to learn phonetically similar words: Effects of age and vocabulary size. *Infancy*, *3*(1), 1-30.

Werker, J.F., Hensch, T.K. (2015). Critical periods in speech perception: New directions. *Annual review of psychology*, 66, 173-196.

Wolke, D., Samara, M., Bracewell, M., Marlow, N., EPICure Study Group. (2008). Specific language difficulties and school achievement in children born at 25 weeks of gestation or less. *J. Pediatrics*, 152, 256-262.

Yoshida, K.A., Fennell, C.T., Swingley, D., Werker, J.F. (2009). Fourteen month-old infants learn similar sounding words. *Dev. Sci.*, 12(3), 412-418.

Sampling units	N total obs = 72 N Participants = 36							
Model specification	**Model name**	**Nested / simpler model**	**Model fit**				**LRT test against nested**	
			AIC	**BIC**	**LL**	**df**	**df**	**χ^2**
Condition x Language of test + Gender + Linguistic background + Scaled Age of test + (1 | Participant)	Null	---	-97.1	-76.6	57.5	9	---	---
Condition x Group + Condition x Language of test + Gender + Linguistic background + Scaled Age of test + (1 | Participant)	Full	Null	-99.6	-74.5	60.8	11	2	6.49*

Final model:

Fixed effects				
	Estimate	**SE**	**z**	**p**
Intercept	0.3781	0.1491	2.536	.0112*
Condition (Novel)	0.0177	0.1419	0.125	.9007
Group (L-MPT)	0.0784	0.1535	0.511	.6092
Language of test (ESP)	-0.0864	0.1552	-0.556	.5779
Gender (F)	-0.0501	0.1220	-0.411	.6811
Bilingual (Bilingual)	-0.0248	0.1263	-0.196	.8445
Scale(Age of test)	0.0994	0.0632	1.571	.1161
Condition x Group	-0.4331	0.1828	-2.369	.0178*
Condition x Language of test	0.1572	0.1853	0.848	.3963
Random Effects				
	Variance	**SD**		
Participant (Intercept)	0.0466	0.216		
Model Fit				
R^2	**Marginal**	**Conditional**		
	0.20	0.47		
Model equation: Mean PLT ~ Condition x Group + Condition x Language of test + Gender + Linguistic background + scale(Age of test) + (1|Participant)				

Table 1. Model comparison. Table structure is inspired by Meteyard & Davies (2020) guidelines for reporting linear mixed-effects models in social sciences. The first column (model specification) outlines the current model and its components. The second column (model name) identifies the created model. The third column (nested/simpler model) designates the model against which the current model is compared, identified by its model name. The fourth column (model fit) provides information on model fit, including the Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), Log Likelihood (LL), and degrees of freedom (df). The last column (LRT test against nested) presents Likelihood Ratio Test (LRT) results for the current model against the nested/simpler model, including degrees of freedom (df) and Chi-square values (χ^2).