

Input Quality and Speech Perception Development in Bilingual Infants' First Year of Life

Marina Kalashnikova^{1,2} & Manuel Carreiras^{1,2,3}

¹Basque Center on Cognition, Brain and Language, San Sebastian, Spain.

²IKERBASQUE, Basque Foundation of Science, Bilbao, Spain.

³University of the Basque Country UPV/EHU, Spain.

Corresponding Author:

Marina Kalashnikova

BCBL. Basque Center on Cognition, Brain and Language

Mikeletegi Pasealekua, 69

Donostia, Gipuzkoa, 20009, Spain

ph: +34 943 309 300

email: m.kalashnikova@bcbl.eu

Acknowledgments: This research is supported by the Basque Government through the BERC 2018-2021 program, and by the Agencia Estatal de Investigación through BCBL Severo Ochoa excellence accreditation SEV-2015-0490. The first author receives support from the European Union's Horizon 2020 Marie Skłodowska-Curie individual Fellowships European Programme under Grant Agreement No 798908 'Optimising IDS' and by the Spanish Ministry of Science and Innovation through the Ramon y Cajal Research Fellowship, PID2019-105528GA-I00. We are grateful to all the infants and families who participated in this research.

Individual differences in infants' native phonological development have been linked to the quantity and quality of infant-directed speech (IDS). The effects of parental and infant bilingualism on this relation in 131 five- and nine-month-old monolingual and bilingual Spanish and Basque infants (72 male; 59 female; from white middle-class background) were investigated. Bilingualism did not affect the developmental trajectory of infants' native and non-native speech perception and the quality of maternal speech. In both language groups, vowel exaggeration in IDS was significantly related to speech perception skills for nine-month-olds ($r=-.30$), but not for five-month-olds. This demonstrates that bilingual and monolingual caregivers provide their infants with speech input that assists their task of learning the phonological inventory of one or two languages.

Input Quality and Speech Perception Development in Bilingual Infants' First Year of Life

During their first year of life, infants undergo a process of perceptual learning and acquire specialised knowledge about the phonetic, phonotactic, and lexical properties of their native language (Kuhl, 2004; Werker, 2018 for reviews). These properties are extracted from the rich linguistic input that infants receive through dynamic communicative interactions with adults and other children in their environment (Kuhl, 2004). The majority of infants around the world grow up in multilingual communities, so they face the additional challenge of extracting the properties relevant for each of their languages from more complex and varied linguistic input (Byers-Heinlein & Fennell, 2014). Previous research has made significant progress in understanding how the quantity and quality of early linguistic input influence the earliest stages of language development in monolingual infants (Kalashnikova & Burnham, 2018; Liu et al., 2003), but this relation continues to be poorly understood in the case of multilingual infants. Here we investigate the relation between bilingual infants' development of native speech perception abilities and the acoustic properties of their caregivers' speech. The main purpose of this study is to determine the extent to which caregivers of bilingual and monolingual infants adjust the properties of their speech according to their infants' linguistic background and linguistic needs, and how infants in turn use this information to shape their own linguistic development.

Early Speech Perception Abilities in Monolingual and Bilingual Infants

At birth, infants' speech perception abilities appear to be universal. In their first months, infants can discriminate phonetic differences that are and are not part of their native language's phonetic inventory, preparing them to learn any language in the world (Kuhl, 2004). As their immersion into the native language environment continues, this ability narrows down, and language-specific speech perception abilities are developed. This process known as perceptual

attunement is marked by maintained or increased sensitivity to native phonetic contrasts and a decrease in sensitivity to non-native contrasts (Kuhl et al., 2006), and it occurs around four to six months for vowel categories, and around 10-11 months for consonants (Polka & Bohn, 2011; Polka & Werker, 1994; Stager & Werker, 1997; Werker & Tees, 1984).

Infants have been proposed to enlist domain-general statistical learning mechanisms during this process (Maye et al., 2002, 2008). According to this view, infants track the frequencies with which speech sounds occur in their linguistic input, and this leads to the identification of the distributional properties of the phonetic categories of their native language. While this is a powerful mechanism, more recent research has proposed that it may not be sufficient to fully account for the early development of native speech perception skills (McMurray et al., 2009; Schatz et al., 2021; Werker et al., 2012). While infants are learning the sounds of their language, they have access to other rich acoustic and contextual information in the input, and they can rely on their emerging lexical knowledge in this process (Swingley, 2009). In their first year of life, infants learn the sounds but also words of their native language (Bergelson & Swingley, 2012), and early knowledge of word forms can guide them in finding lexically relevant speech contrasts in their language, especially for cases when the distributional properties of specific sound categories are not easily differentiated in the input (Swingley & Alarcon, 2018).

Unlike monolinguals, bilingual infants must simultaneously acquire two phonetic inventories. To do so, bilinguals first must be able to differentiate the two languages that they hear in their environment. This ability appears to be available very early on – at birth for bilinguals acquiring languages from different rhythm classes (Byers-Heinlein et al., 2010), and around five months for bilinguals acquiring rhythmically-similar languages (Bosch & Sebastián-Gallés, 2001; Molnar, Gervain, & Carreiras, 2014). This suggests that bilinguals may be able to successfully separate the two languages in their input and acquire the phonetic

categories of each language along a similar timeline as their monolingual peers. This proposal was first supported by Burns and colleagues (Burns et al., 2007) who assessed the ability to discriminate native consonant contrasts in monolingual English and bilingual English-French infants at six-to-eight, 10-to-12, and 14-to-20 months. English and French both include a voicing consonant contrast in their phonemic inventories, but the specific categories are marked by different boundaries placed along the Voice Onset Time (VOT) continuum. In English, the voiced category is placed around the 0 VOT point, and the voiceless category is marked by a positive VOT lag perceived as aspiration. In French, on the other hand, the voiced category is marked by a negative VOT, while the voiceless category is placed closer to the 0 VOT point. Thus, bilingual infants acquiring English and French must keep track of the statistical distributions of these categories separately for each one of their languages and adjust the different VOT boundaries accordingly. Monolingual and bilingual infants were reported to discriminate both contrasts at six-to-eight months according to the universal pre-attunement perceptual abilities expected at this age. However, at 10-to-12 and 14-to-20 months, monolingual infants only discriminated their native English contrast, while bilinguals discriminated both contrasts demonstrating native perceptual abilities for both their languages.

In some cases, a temporary delay has been observed in bilingual infants' attunement to native phonetic categories, in particular, when the distributional properties of a category were different when tracked within one language or across two languages. Such delay was reported by Bosch and Sebastián-Gallés (2003) for Catalan-Spanish bilingual infants' discrimination of the /e/-/ɛ/ vowel contrast. This contrast is phonemic in Catalan, whereas Spanish only employs the /e/ category. Crucially, the Spanish category falls between the Catalan /e/ and /ɛ/ in vocalic space, and it occurs with significantly greater frequency compared to both Catalan vowels. Thus, when the distributions of these three categories are tracked in bilingual infants' input, they amount to a single unimodal distribution corresponding to one category. In addition, the

two languages share numerous cognates that are primarily differentiated by vowel changes, which can lead infants to accept contrasts between acoustically similar vowels as mispronunciations rather than lexically meaningful contrasts (Ramon-Casas et al., 2009), and can make the cues to category membership extracted from infants' early-learned word forms less reliable. Bosch and Sebastián-Gallés showed that Spanish and Catalan monolinguals discriminated the /e/-/ɛ/ contrast at four months, but only Catalan monolinguals did so at eight months. Spanish-Catalan bilinguals also discriminated the contrast at four months, failed to discriminate it at eight months, and then showed a resurgence of discrimination at 12 months. This U-shaped developmental trajectory in the bilingual sample suggests that bilingual infants required additional accumulated exposure to the two languages in order to track the challenging distributional properties of these vowel categories (see also Sebastián-Gallés & Bosch, 2009). Importantly, however, bilinguals do not show these delays under conditions that assist discrimination, which can depend on their specific language pair (Sundara & Scutellaro, 2011), the phonetic contrast used at test (Sundara et al., 2008), or the specific experimental task (Albareda-Castellot et al., 2011). Therefore, this research indicates that monolingual and bilingual infants follow similar trajectories in speech perception development, but the exact timeline for the acquisition of specific native contrasts by bilingual infants may be impacted by the degree to which their languages' phonetic and lexical inventories overlap.

An alternative view is that early bilingual experience may affect the developmental trajectory of speech perception regardless of the phonetic inventories of the specific language pairs acquired by bilingual infants. Bilinguals experience a significantly different language environment compared to monolinguals (Byers-Heinlein & Fennell, 2014); their language input in each language is reduced, and it tends to be more variable since the two languages are commonly mixed within and across utterances (Orena et al., 2019). A protracted period of perceptual openness manifested in later attunement to native categories and persistent ability

to discriminate non-native categories may thus emerge as bilingual infants' perceptual systems adapt to these environmental factors (Petitto et al., 2012). For instance, bilingual infants have been shown to maintain sensitivity to non-native speech contrasts that do not occur in their language input (i.e., non-native to both their languages) after their first birthday (Burnham et al., 2018; Graf Estes & Hay, 2015; Liu & Kager, 2017; Singh, 2018). These processing differences have also been observed in infants' brain responses to native and non-native speech sounds. Monolingual English-learning infants exhibit a negative mismatch response (MMR) to both native and non-native phonetic contrasts at seven months of age, but only to native contrasts at 11 months (Rivera-Gaxiola et al., 2005). Bilingual infants, on the other hand, do not show MMRs at seven months, and only do so at 11 months (Garcia-Sierra et al., 2011). More recent analyses of bilingual neural responses to speech sounds suggest that bilinguals exhibit a positive mismatch response to native stimuli from six to 11 months (Ferjan Ramírez et al., 2017; Garcia-Sierra et al., 2016), a component that is typically associated with acoustic-level processing (as opposed to the phonemic-level associated with negative MMRs) and more effortful and immature speech perception (e.g., Friederici et al., 2007). Interestingly, the polarity and amplitude of infants' MMRs to native language contrasts significantly relates to the amount of linguistic input that they receive in daily interactions with their caregivers, and to bilinguals' amount of exposure to the language of the test phonetic contrast (Garcia-Sierra et al., 2016; Ramírez-Esparza et al., 2017).

This evidence suggests that monolingual and bilingual infants' neural commitment to their native language or languages and maturation of speech processing abilities is supported by the quantity of language input that they receive in their environment. In the case of bilinguals, a protracted period of perceptual attunement may be a useful adaptation to the properties of bilingual language input – greater perceptual openness and flexibility can allow bilinguals to continue sampling information from their environment (D'Souza et al., 2020),

thus mitigating the challenge of extracting meaningful linguistic information from more reduced and complex input compared to monolinguals (Singh, 2021). However, input quantity is not all that infants need; the *quality* of speech that infants hear in their environment can also impact the developmental trajectory of speech perception skills.

Infant-directed Speech and Early Language Development

When adults interact with young infants, they use infant-directed speech (IDS), a speech register that compared to adult-directed speech (ADS) is characterized by grammatical and lexical simplification (Soderstrom, 2007), slow speech rate (Narayan & McDermott, 2016), positive warm affect (Kitamura & Burnham, 2003), and exaggerated pitch height and range (Fernald & Simon, 1984). In addition, adults exaggerate the acoustic properties of speech sounds, particularly vowels, in IDS compared to ADS (Burnham et al., 2002; Kuhl et al., 1997). This is typically assessed by extracting the first and second formant (F1, F2) values for the three corner vowels /i/, /u/, /a/ from speech recordings, plotting them in two-dimensional space, and measuring the area of the resulting vowel triangle. A larger triangle area represents more distinct corner vowels, and it is characteristic of more intelligible and clear speech (Bradlow et al., 1996). Accordingly, vowel triangle areas tend to be larger in IDS than ADS, so it appears that caregivers (albeit unconsciously) facilitate infants' language acquisition by providing them with clear and more intelligible speech.

In support of this claim, there is evidence that exposure to acoustically exaggerated vowels enhances speech processing in young infants. In experimental tasks when speech stimuli contain acoustically exaggerated vowels, infants show more mature processing of vowel categories (Peter et al., 2016) and are faster at recognising familiar words (Song et al., 2010). Critically, the degree to which individual caregivers exaggerate vowels in IDS is significantly related to their infants' linguistic ability. Liu and colleagues (Liu et al., 2003) assessed discrimination of a native consonant contrast in six-to-eight and 10-to-12-month-old

infants acquiring Mandarin Chinese and recorded the IDS of these infants' mothers. Results showed that mothers who produced larger vowel space areas in IDS had infants who required significantly fewer trials to reach a pre-established criterion in the consonant discrimination task, which is an index of more mature native speech perception abilities. This relation appears to extend beyond speech perception as the degree to which individual mothers exaggerate vowels in IDS has also been shown to relate to their infants' concurrent and future vocabulary size from nine to 19 months (Hartman et al., 2017; Kalashnikova & Burnham, 2018) and even later language proficiency assessments at four years of age (Dilley et al., 2020).

A recent study suggests that a similar relation between vowel exaggeration in caregiver IDS and infants' early speech processing capacities may be present for bilingual mother-infant dyads. Garcia-Sierra et al. (Garcia-Sierra et al., 2021) assessed the vowel quality in IDS by the caregivers of 17 Spanish-English 11- and 14-month-old infants as well as their infants' neural responses to native and non-native phonetic contrasts using EEG. Both native and non-native stimuli elicited positive MMRs in infants (Ferjan Ramírez et al., 2017; Garcia-Sierra et al., 2011), but larger and more widely distributed MMRs were observed for native compared to non-native stimuli. Furthermore, only the strength of the MMR response to the native contrast was positively related to the vowel triangle area of caregivers' IDS. This study was the first to suggest that despite potential differences in the developmental timeline for perceptual attunement in monolingual and bilingual infants, this process is similarly supported by the quality of infants' early linguistic input. However, it left several unanswered questions that are essential for understanding the nature of this input-output relation in bilinguals. First, the lack of a monolingual comparison group does not allow us to specify whether the observed correlation reflects individual differences in parental speech production that manifest regardless of their own or their infants' bilingualism, or if acoustic adjustments in bilingual IDS differ from monolingual IDS. Second, it leaves unknown how bilingual infants' *dual*

language input quality relates to speech perception development since IDS and ADS were only assessed in caregivers' and infants' dominant language.

The Current Study

This study aims to address these questions by investigating the development of native and non-native speech perception abilities in monolingual and bilingual infants, and the relation between infants' performance and the degree of vowel exaggeration in their caregivers' IDS in each language. For this purpose, this study included caregiver-infant dyads with infants aged five and nine months, ages that correspond to timepoints before and during the process of phonological attunement (Werker, 2018). Infants completed a speech perception task that assessed their ability to discriminate a native and a non-native phonetic contrast. In addition, infants' mothers were recorded in interactions with their infants and with an adult, which were used to calculate the vocalic triangle area in their IDS compared to their own ADS. This study included a combination of confirmatory and exploratory efforts. As explained below, specific predictions were constructed for infants' performance in the speech perception task based on the extensive existing research on speech perception development in monolingual (see Werker, 2018 for a review) and bilingual infants (see Höhle, Bijeljac-Babic, & Nazzi, 2019 for a review). On the other hand, the assessment of the properties of Spanish and Basque IDS and their relation to infant speech perception were exploratory given the highly limited previous evidence about IDS properties in these two specific languages and IDS properties to bilingual infant populations in general.

The monolingual infants in this study were acquiring Spanish or Basque, and bilinguals were acquiring Spanish and Basque, which comprise an optimal language pair for isolating the effects of bilingualism on infant speech perception. Spanish and Basque are typologically distinct languages, and their grammatical and lexical structures are highly different. Infants acquiring these languages in monolingual and bilingual contexts show the ability to

discriminate them based on their prosodic patterns (Molnar et al., 2014). Despite these differences, the phonetic inventories of Spanish and Basque are largely overlapping, particularly in the case of stop consonants, which were tested here (Hualde, 2004; Rosner, López-Bascuas, García-Albea, & Fahey, 2000). Specifically, Basque and Spanish both employ a two-way VOT distinction between voiced and voiceless plosives with no aspiration (this is the case for the dialect of Basque studied here, though some cross-dialectal differences have been reported in the literature) (Hualde et al., 2011; Keating et al., 1983). Second, the Spanish-Basque bilingual population enables us to assess the effects of *infant* bilingualism on caregivers' speech while maintaining caregivers' language status identical across groups. All dyads in this study were recruited in the Basque country where these languages are spoken and share the official language status (Lasagabaster, 2018). All adult speakers of Basque are also proficient speakers of Spanish, so all caregivers in this study were proficient native bilinguals, who have decided to raise their infants either as monolingual or as bilingual.

Separate predictions were constructed for the younger and older age groups' performance in the speech perception task in this study. First, for the five-month age group, we predicted that monolingual and bilingual infants would show discrimination of the native and non-native contrasts in our speech perception task, which is consistent with the language-general abilities prior to perceptual attunement (Werker & Tees, 1984). Second, for the nine-month group, we predicted greater discrimination ability of the native contrast compared to the non-native contrast. Moreover, in line with the proposal that bilingual experience impacts infants' speech perception development, this decline in non-native contrast discrimination was expected in the monolingual but not the bilingual group, indicating a protracted period of phonological attunement in bilingual infants (Ferjan Ramírez et al., 2017; Garcia-Sierra et al., 2011; Petitto et al., 2012).

In the case of our IDS measure, we expected that caregivers of monolingual and bilingual five- and nine-month-olds would exhibit vowel exaggeration in IDS compared to their own ADS. This would provide the first evidence for this acoustic adjustment in Basque IDS. Comparable degrees of vowel exaggeration in Spanish and Basque and in caregivers of monolingual and bilingual 5- and 9-month-old infants would support the view that this adjustment is manifested across a variety of languages and is independent of infants' age (Kalashnikova & Burnham, 2018; Liu et al., 2009), and that phonetic properties of bilingual and monolingual caregivers' IDS are comparable (Danielson et al., 2014). Furthermore, individual indices of acoustic vowel exaggeration in IDS were expected to relate to infants' native speech discrimination abilities across the entire sample including monolingual and bilingual infants (Garcia-Sierra et al., 2021; Liu et al., 2003).

Method

Participants

The final sample included 131 infants (72 male and 59 female) and their primary caregivers (mothers). Seventy-seven infants were monolingual Spanish or Basque, and 54 infants were bilingual Spanish-Basque. See Table 1 for sample sizes, infants' age, and sex distribution for each age and language sub-group. Infants were not premature, were in good physical health at the time of the experiment and did not have any reported sensory deficits or risk for developmental disorders. Participants were white and lived in the Basque Country in Spain (all infants were born in this region). All caregivers who took part in the interaction sessions (IDS and ADS) were infants' biological mothers (M age = 35.13 years, SD = 4.08). Maternal education ranged from a high school to a doctorate degree with a university degree as the Median, which did not differ between the monolingual and bilingual groups, Mann-Whitney $U = 2531.00$, $p = .792$. Additional 18 mother-infant dyads took part in the study but were excluded due to failure to contribute data for either task (6), and failure to comply with

the language exposure criteria for inclusion in the monolingual and bilingual groups (12). The sample size included in this study was based on previous research on speech perception comparing monolingual and bilingual infants' performance (e.g., Albareda-Castellot et al., 2011; Bosch & Sebastián-Gallés, 2003; Sundara & Scuttelaro, 2011). This study received approval from the [blind for review] Ethics Committee (approval number: 22102018M). Data for this study were collected from 01 November 2018 to 01 March 2020.

Bilingual exposure and language group inclusion criteria. Mothers completed the Language Exposure Questionnaire in which they reported their infants' patterns of exposure to Spanish and Basque from birth until the day of testing, their own language proficiency levels in Spanish and Basque, and patterns of language use in the home (Bosch & Sebastián-Gallés, 2001; Molnar et al., 2014). Infants were included in the monolingual group if they received at least 90% of exposure to Spanish or Basque, and they were included in the bilingual group if their Spanish-Basque exposure ranged from 50-50% to 75-25% with no more than 5% exposure to a third language (see Byers-Heinlein, 2015 for a discussion on defining criteria for bilingual group inclusion). In the bilingual group, both parents used both languages with the infant in 17 families, parents used the one parent-one language strategy in 27 families, and in 8 families, both parents reported to primarily use one language in interactions with their infant and the other language in interactions with each other (the latter group of infants were included despite appearing to have primary exposure to one language since their estimated percentages of exposure to the two languages were comparable to the rest of the bilinguals in the sample and fit our language exposure criteria; this is consistent with research findings that bilingual infants receive similar exposure to a non-dominant language when their parents report using it to address the infant or only to address each other, Orena et al., 2019). Table 1 presents detailed information about infants' language background.

Six mothers were Spanish monolinguals with no knowledge of Basque. As described below, their infants were included in the Speech Perception sample, but their IDS and ADS data were not analysed. The remaining mothers were Spanish-Basque bilinguals who used their dominant language on average for 65% ($SD = 13.78$) of time, and with an average age of acquisition of the non-dominant language of 4 years ($SD = 3.86$). All mothers of bilingual infants confirmed that they used both Spanish and Basque in interactions with their infants at least to some extent (regardless of the language use patterns in their family), and that they felt comfortable speaking Spanish and Basque during a play session with their infant.

Native and Non-Native Speech Perception Task

This task assessed infants' ability to discriminate three categories of bilabial consonants placed at different points along the VOT continuum: two categories belonged to the phonetic inventories of both Basque and Spanish – voiced /b/ with a long negative VOT and voiceless /p/ with a brief positive VOT, and one category did not belong to the inventories of either language – aspirated voiceless /p^h/ with a long positive VOT. A female English speaker was recorded producing the three syllables (an English native was chosen to obtain productions of the syllable /p^ha/, which is native to English but non-native to Spanish and Basque). A single token was chosen for each syllable, and these tokens were subject to manipulations in Praat (Boersma & Weenink, 2010) to obtain the VOT values that are native and non-native to Spanish and Basque and to equate the intensity and pitch of the tokens. The VOT values were -60 msec for /ba/ (duration = .509 sec), +16 msec for /pa/ (duration = .439 sec), and +48 msec for /p^ha/ (duration = .452 sec). The syllables had an intensity of 75 dB and pitch of 145.35 Hz. Repetitions of each syllable were used to create three 30 second strings (one for each syllable) with varying inter-stimulus intervals that ranged from .850 to 1.050 sec.

These auditory stimuli were paired with a static image of a colourful bullseye presented on a white background. In addition to the habituation and test stimuli, the task included a pre-

test and a post-test trial that consisted of a video of a moving water wheel toy and audio recordings of nonwords produced by two female speakers in lively infant-directed speech (e.g., tuda, lamu, tagu, muba). These two trials were presented at the start of the task (before the first habituation trial) and at the end of the task (after the last test trial), respectively, in order to assess infants' overall engagement in the task. Comparable looking times to the pre-test and post-test trials would indicate that any decrease in looking times observed during the task (habituation and test phases) reflected infants' responses to the task stimuli and not to an overall decrease in attention due to fatigue or boredom.

Visual stimuli were presented on a 52 inches television screen, and auditory stimuli were presented via a centrally placed Electro Voice EVID 4.2 loudspeaker hidden under the screen. Habit X (Cohen et al., 2004) software was used to control stimulus presentation and to record infants' gaze behaviors during the task. Infants sat on their caregiver's lap facing the television screen inside a dimly lit child-friendly laboratory room. The caregiver listened to masking music during the task and was instructed not to speak to their infant and not to point to the screen to avoid interfering with their infant's performance. A webcam placed under the screen was used to transmit a live feed of the infant's face, which was monitored by an experimenter who sat in an adjoining control room.

All trials in this task were infant controlled, so each trial started when the infant has fixated the center of the screen for 1 second and terminated when the infant looked away for 2 seconds or when the maximum duration of the trial (30 seconds) was reached. An attention getter (a circle expanding and contracting in the center of the screen) was used to re-direct infants' attention to the screen between trials. First, infants completed a pre-test trial. Next, the habituation phase began. In this phase, infants were presented with /pa/ trials until they reached the habituation criterion or until completing the maximum of 24 trials. The habituation criterion was reached when infants' average looking duration on three consecutive trials was 50% below

their average looking duration on the first three habituation trials. After habituation, infants proceeded to the test phase that consisted of 4 trials: 2 native /ba/ trials and 2 non-native /p^ha/ trials. The native and non-native trials were presented in alternating order, and the selection of the first trial was counterbalanced across infants. The final trial of the task was a post-test trial.

The dependent variable consisted of the infants' looking duration in seconds on each native (looking duration to /ba/) and non-native (looking duration to /p^ha/) test trial and on the final two habituation trials (looking duration to /pa/). The following three *a priori* exclusion criteria were applied in this task: (1) extreme fussiness, crying, and/or failure to complete the task (5 infants), (2) failure to reach the habituation criterion (4 infants), and (3) fussiness manifested in a significant decrease in attention throughout the task (3 infants who showed a 70% or higher decrease in looking time to the post-test compared to the pre-test trial). These criteria enabled us to ensure that all infants in the final sample were attentive during the task, successfully habituated to the initial native stimulus, and any decreases in looking time during habituation and test trials could not be attributed to an overall decrease in attention to the task. The sample used for analyses included 119 infants.

Infant- and Adult-Directed Speech

Each mother completed two interactive sessions: one with their infant (IDS) and one with an experimenter (ADS). In order to elicit the three corner vowels /a/, /i/, /u/ in the two registers, mothers were provided with two sets of toys and images, one for Spanish and one for Basque, to elicit three target words containing these vowels. For Spanish, the target words were 'pato' (duck), 'bicho' (bug), and 'pulpo' (octopus), and for Basque, the target words were 'bale' (whale), 'piku' (fig), and 'puzle' (puzzle). To elicit IDS, mothers were asked to interact with their baby naturally as they do during play time at home. They were instructed to tell their baby a story about each object and each image. To elicit ADS, the experimenter interviewed

the mother and asked her to describe what she told her baby about each object and image. The experimenter was a female native bilingual speaker of Spanish and Basque.

The sessions were conducted inside a sound-proof booth in an infant laboratory. During IDS sessions, mothers sat on a chair facing their infant who sat in a highchair. During ADS sessions, the infant was not present in the booth, and mothers sat facing the experimenter. Mothers' voice was recorded using a head mounted microphone (Audio Technica AT892CWTH) connected to a recorder (Digital audio recorder Zoom H4n). Mothers of monolingual infants completed these sessions in their infant's only language, and mothers of bilingual infants completed each session in the two languages with the order of Spanish and Basque counterbalanced across participants (e.g., IDS Spanish, IDS Basque, ADS Spanish, ADS Basque).

The following *a priori* exclusion criteria were applied for these sessions: (1) the mother was not a bilingual Spanish-Basque speaker regardless of their infant's group assignment (5 dyads), (2) the mother failed to contribute at least 3 analysable tokens of each vowel for each register and each language (7 dyads), (3) the IDS recording could not be completed because the infant was fussy or crying (2 dyads), (4) in the case of bilinguals, the mother failed to contribute IDS and ADS recordings in both languages (4 dyads). These criteria enabled us to ensure that a full set of recordings was available for each mother-infant dyad, that the recordings were of high quality and free from excessive infant or environmental noises, and critically, that maternal language background was matched across the monolingual and bilingual groups. Note that the mothers who were not Spanish-Basque bilinguals, were Spanish monolinguals who had immigrated to the Basque Country from other regions of Spain or from South America. Therefore, it was not only the case that they were not Basque speakers, but they spoke a different dialectal variety of Spanish, so their vowel production could not be

reliably compared to the rest of the sample. The sample used for analyses included 109 mother-infant dyads.

Acoustic analyses. Praat software (Boersma & Weenink, 2010) was used to identify the target words in each recording and manually segment the target corner vowels /i/, /u/, /a/ in each word. All vowel segments were reviewed by a second coder, and any disagreements in segmenting were resolved by a third independent coder. Only vowel segments that were free from noises from the infant or the environment and were not produced in whispered or sung speech were included. Praat scripts were then used to calculate the mean F0, duration, and mean value in Hz of the first and second formants (F1, F2) for the period between the 40% and the 80% points of each vowel's duration. The output was inspected, and outliers were removed (F1 and F2 values that were 3 SD above or below the mean for each register in each language). Next, mean F1 and F2 values were calculated for each target vowel for each mother's IDS and ADS in each language, which were used to calculate the vowel triangle area using the formula: $ABS \frac{1}{2} \times [(F1/a/ \times (F2/i/ - F2/u/) + F1/i/ \times (F2/u/ - F2/a/) + F1/u/ \times (F2/a/ - F2/i/)]$, where F1/a/ refers to the average value in Hz of the first formant for the vowel /a/, F2/i/ to the average value in Hz of the second formant for the vowel /i/, and so forth.

Plots of the extracted vowels and detailed information about the number of vowel tokens used for analyses, their duration, and F0 are presented in Supplementary Materials. As can be seen in Supplementary Tables S1 and S2, the expected IDS and ADS differences were observed in mothers' vowel productions across language and age groups in Spanish and Basque. Vowels in each category were longer in IDS than in ADS, and vowels were produced with higher pitch in IDS than in ADS.

Results

Native and Non-Native Speech Perception Task

Two sets of analyses, each using a different dependent variable, were used to assess infants' performance in this task. The first set used difference scores for each contrast computed as average looking time during the two native test trials minus the last two habituation trials (native difference score) and average looking time during the two non-native test trials minus the last two habituation trials (non-native difference score). These scores were used in one-sample *t*-test analyses that compared each score to 0 (a difference score of 0 would denote no recovery in looking time at test compared to habituation). Difference scores capture the magnitude of looking time recovery in response to each type of test trial compared to habituation, but they are based on averages from the two trials of each type, thus obscuring individual and inter-trial variability in infants' performance. To counteract this limitation and to directly compare performance across the two age and language groups, the second set of analyses used infants' raw looking times to the three trial types, habituation, native, and non-native, in a Linear Mixed Effects model. The model was followed by pairwise comparisons that assessed infants' looking time recovery in each type of test trial compared to habituation. Given that separate performance patterns were predicted for the two age groups, all *t*-tests and pairwise comparisons were conducted separately for the 5-month and 9-month data. Finally, infants' performance in the habituation phase of the speech perception task was also analysed, and these results and discussion are reported in the Supplementary Materials.

First, the one-sample *t*-test analyses indicated that in the 5-month group, difference scores were significantly above 0 for the native, $t(56) = 4.554, p < .001, d = .603$, and the non-native test trials, $t(56) = 5.070, p < .001, d = .672$. This was also the case in the 9-month-group for the native, $t(61) = 2.905, p = .005, d = .369$, and the non-native test trials, $t(61) = 3.172, p = .002, d = .403$. Difference scores for the native and non-native test trials did not differ within the 5-month, $t(56) = -.330, p = .743, d = -.044$, and 9-month groups, $t(61) = .167, p = .868, d =$

.021. Therefore, infants in the two age groups showed significant recovery in looking times in the test trials compared to the habituation trials.

Second, a Linear Mixed Effects (LME) model was constructed to compare performance across Trial Types (Habituation /pa/, Native /ba/, and Non-Native /p^ha/), Language Groups (monolingual, bilingual), and Age Groups (5 months, 9 months). The model included the following two-way and three-way interactions: Trial Type × Age Group, Trial Type × Language Group, Trial Type × Age Group × Language Group. In addition, test order (native trial first, non-native trial first) was included as a main effect, and random intercepts were specified per participant. The resulting model was specified as follows:

$$\begin{aligned} \text{Looking time} \sim & \text{Trial Type} + \text{Age Group} + \text{Lang Group} + \text{Test Order} + \\ & \text{Trial Type} * \text{Age Group} + \text{Trial Type} * \text{Lang Group} + \\ & \text{Trial Type} * \text{Age Group} * \text{Lang Group} \\ & + (1|\text{Participant}) \end{aligned}$$

The continuous variables used in the analyses were scaled and centered around zero to assist with model convergence. Simple effects coding was used for the categorical variables. To directly assess infants' native and non-native discrimination ability and test our predictions constructed for each age group, the LME model was followed by pairwise comparisons of infants' looking times to the native test trials minus habituation trials (looking time recovery to the native contrast) and to the non-native test trials minus habituation trials (looking time recovery to the non-native contrast). Analyses were conducted using the lme4 (Bates, 2005) and lmerTest (Kuznetsova et al., 2015) packages in R.

Infants' looking times to the three trial types are displayed in Figure 1, and Table 2 presents the output of the LME model. The main effects of Trial Type and Language Group were significant. Overall, compared to habituation, infants produced longer looking times when presented with the native contrast (/pa/ vs. /ba/), and when presented with the non-native

contrast (/pa/ vs. /p^ha/). The main effect of Language Group also revealed that bilingual infants overall produced longer looking times compared to monolingual infants across trial types. This effect was qualified by an Age Group by Language Group interaction. Across trial types, bilingual 5-month-olds produced longer looking times than monolingual 5-month-olds, $\beta = -.431$, $SE = .166$, $CI[-.761, -.102]$, $t = -2.595$, $p = .011$, but this difference was not statistically significant for the 9-month-olds, $\beta = .088$, $SE = .156$, $CI[-.221, .396]$, $t = .563$, $p = .574$.

The model also yielded a Trial Type by Age Group interaction indicating that looking times to the non-native trials by 9-month-olds differed compared to the intercept, and an Age Group by Language Group interaction indicating that looking times by 9-month-old bilinguals differed compared to the intercept. To further understand the sources of these interactions, and to test our predictions constructed for each age group, pairwise comparisons were conducted (Bonferroni corrections were used to adjust the p -values for multiple comparisons, resulting in an adjusted significance threshold of 0.0125 for four tests). These comparisons revealed that five-month-olds produced longer looking times to native test trials compared to habituation trials, $\beta = -.463$, $SE = .115$, $CI[-.690, -.237]$, $t = -4.018$, $p < .001$, and to non-native test trials compared to habituation trials, $\beta = -.516$, $SE = .115$, $CI[-.743, -.290]$, $t = -4.475$, $p < .001$. Nine-month-olds also produced longer looking times to native test trials compared to habituation trials, $\beta = -.288$, $SE = .108$, $CI[-.500, -.076]$, $t = -2.663$, $p = .008$, but the difference between non-native test trials and habituation trials did not reach significance, $\beta = -.243$, $SE = .108$, $CI[-.456, -.030]$, $t = -2.245$, $p = .025$. Therefore, we observed significant recovery in infants' looking times in the native and non-native test trials at the group level (as shown by the one-sample t -tests on difference scores), but evidence for this recovery was not significant in the 9-month-old group when individual and trial-by-trial variability were taken into account.

Infant-directed Speech

Figure 2 displays the vowel triangles for IDS and ADS for each language and age group. For analyses, vowel hyper-scores were computed for each mother by dividing the vowel area in IDS by the vowel area in ADS. Contrary to the raw vowel area values in each register, vowel hyper-scores can be used as an index of each mother's degree of vowel exaggeration in IDS using her own ADS as a baseline and thus accounting for individual differences in vowel production. These hyper-scores were first subject to one-sample *t*-test analyses that compare them to the value of 1. If vowel hyper-scores are > 1 , this indicates larger vowel area in IDS compared to ADS, vowel hyper-scores = 1 indicate vowel areas in IDS that are comparable to ADS, and vowel hyper-scores < 1 indicate smaller vowel areas in IDS compared to ADS. Next, to assess IDS production across age and language groups, these hyper-scores were used as dependent variables in univariate ANOVAs. These analyses were conducted separately for Spanish and Basque to avoid any language-specific vowel production effects on the observed results patterns.

Mothers produced hyper-scores significantly above 1 in Spanish IDS ($M = 2.100$, $SD = 2.45$), $t(64) = 3.626$, $p < .001$, and in Basque IDS ($M = 1.636$, $SD = 1.642$), $t(88) = 3.655$, $p < .001$. Univariate ANOVAs did not yield significant effects of Age or Language Group on mothers' hyper-scores. This was the case for Spanish (age group, $F(1, 61) = .160$, $p = .691$, language group, $F(1, 61) = .009$, $p = .926$, age group \times language group, $F(1, 61) = 1.211$, $p = .276$), and for Basque (age group, $F(1, 85) = 1.029$, $p = .313$, language group, $F(1, 85) = 2.163$, $p = .145$, age group \times language group, $F(1, 85) = .0002$, $p = .988$).

Mothers of bilingual infants produced IDS in both their infant's dominant and non-dominant language. To assess whether infants' language dominance had an impact on maternal vowel production, additional independent-samples *t*-tests were conducted within the bilingual group for each language. In Spanish and Basque, there were no statistical differences between

hyper-scores in infants' dominant and non-dominant language (Spanish, $t(44) = -.188, p = .852$; Basque, $t(44) = -0.718, p = .477$).

Relation between IDS and native and non-native speech perception performance

Finally, the relations between maternal IDS quality and infants' speech perception performance were assessed. Native and non-native difference scores were used for these analyses. Maternal vowel hyper-scores were used for monolingual infants' only language and for bilingual infants' dominant language. Given that different speech perception patterns emerged for 5- and 9-month-old groups, separate correlations were conducted by age group. The analyses for 5-month-olds showed no significant relations between infants' native ($r = .018, p = .904$) and non-native difference scores ($r = .181, p = .228$), and the hyper-scores in their mothers' IDS. For 9-month-olds, there was also no significant relation between native difference scores and hyper scores in IDS ($r = -.063, p = .665$), but there was a significant negative relation between infants' non-native difference scores and hyper-scores in maternal IDS ($r = -.303, p = .032$). That is, 9-month-old infants who directed less attention to the non-native contrast had mothers who exaggerated vowels to a greater extent in IDS relative to ADS.

For the subset of bilingual infants, additional analyses were conducted to assess the relations between hyper-scores in infants' non-dominant language and their speech perception performance. In this case, no relations were significant for 5-month-olds (native scores, $r = .174, p = .534$; non-native scores, $r = -.162, p = .535$) and for 9-month-olds (native scores, $r = -.072, p = .752$; non-native scores, $r = .148, p = .510$).

Discussion

This study investigated the relation between native speech perception abilities in monolingual and bilingual infants and the quality of their early speech input. Monolingual and bilingual infants' performance did not differ significantly in this speech perception task. This study included 5- and 9-month-old infants, which mark timepoints prior to the completion of

perceptual attunement. Accordingly, in both age groups, infants showed discrimination of the native and non-native contrast. However, only for nine-month-olds, the difference in looking times between the test and habituation trials reached significance for the native and not the non-native contrast, which is indicative of a decrease in sensitivity to non-native phonetic contrasts at this age. In addition, our study was the first to compare the properties of bilingual mothers' IDS addressed to monolingual and bilingual infants, and we showed that mothers produced acoustically exaggerated vowels in IDS compared to ADS, and this adjustment did not differ across the two infant age and language background groups, and two typologically distinct languages, Spanish and Basque. Most importantly, the degree to which individual mothers exaggerated vowels in IDS significantly related to nine-month-old infants' non-native speech perception abilities. These findings suggest that even though monolingual and bilingual infants experience different types of language environments, they receive exposure to similar acoustic adjustments in their caregivers' speech, which support infants' challenging task of navigating their early linguistic input and learning the specific properties of one or two languages.

The results of this study did not reveal significant differences in the developmental trajectory for phonological attunement in monolingual and bilingual infants. That is, in a linguistic environment in which the two languages are phonetically similar, bilingual infants demonstrated universal speech discrimination abilities at 5 months and the emergence of language specific abilities at 9 months, which were comparable to their monolingual peers. This implies that early bilingualism *per se* does not impact the attunement of infants' language-specific speech discrimination skills, but this process is rather shaped by infants' ability to accrue sufficient exposure to the phonetic categories belonging to each language in their input.

This explanation matches previous findings in the literature. Bilingual infants have been demonstrated to exhibit apparent delays in native speech contrast discrimination compared to

monolinguals when tested on contrasts that are not shared by their two languages, or which distributions are affected by the co-occurrence of the two languages in bilingual input. This is the case for the vowel discrimination patterns reported for Catalan-Spanish bilinguals who showed delays in discriminating the challenging /e/-/ɛ/ contrast (Bosch & Sebastián-Gallés, 2003). However, when the same population was tested on the /e/-/o/ contrast phonemic to both their languages, no such delay was observed (Sebastián-Gallés & Bosch, 2009). Similarly, in this study Spanish-Basque bilinguals show discrimination of a consonant contrast native to the two languages comparable to their monolingual peers. However, this population was shown to require several extra months of language exposure to discriminate a low frequency fricative contrast native only to Basque but not Spanish (Larraza et al., 2020). Therefore, we propose that these findings reflect differences in monolingual and bilingual infants' patterns of exposure to one vs. two languages and the degrees to which the phonemic inventories of the two languages overlap, but not to effects of bilingualism on processes underlying speech perception.

This proposal is further supported by infants' discrimination of the non-native contrast in this study. Bilingual infants did not show greater discrimination of the non-native contrast compared to monolinguals at either age, which is contrary to the predictions by the account that bilingualism leads to a protracted period of perceptual openness in this population. This result is contrary to several previous findings that have reported greater sensitivity to non-native contrasts in monolinguals compared to bilinguals, which can persist into infants' second year of life (Burnham et al., 2018; Graf Estes & Hay, 2015b; Liu & Kager, 2017; Singh, 2018; Singh et al., 2017). It is possible that a lack of language group effect here is due to the age of the infants. We deliberately chose the age groups of five and nine months given our interest in capturing individual differences in native and non-native speech perception at ages when infants have not completed the process of perceptual attunement. Accordingly, nine-month-old

monolinguals and bilinguals in this study continued to show discrimination of the non-native contrast. Therefore, it is possible that group effects would emerge in older infants closer to or beyond the age of 10-12 months, with monolinguals showing a more pronounced decline in non-native perception compared to bilinguals. This possibility is interesting, and we leave it for future research, but so far, our findings suggest that between five and nine months of age, monolingual and bilingual infants follow comparable trajectories for the development of perception abilities.

In addition to testing infants' speech perception skills, we also assessed the properties of their mothers' IDS. As expected, mothers produced heightened pitch and exaggerated vowels in IDS compared to ADS regardless of whether they spoke Spanish or Basque, or if they addressed monolingual or bilingual infants. Most importantly, the degree to which mothers exaggerated vowels significantly related to their nine-month-old infants' speech perception performance. Specifically, mothers who produced larger vowel triangles in IDS relative to their own ADS had infants who were less likely to discriminate a consonant contrast that was non-native to both Spanish and Basque. Thus, enhanced clarity in maternal speech supports their infants' speech perception development, and its role is most prominent when infants are in the process of attuning to their native language or languages, precisely as their sensitivity to native categories is maintained while sensitivity to non-native categories is in progressive decline.

While our results support the proposed linguistic function for vowel exaggeration in IDS, we acknowledge that it is controversial and has received criticism for two main reasons. First, vowel exaggeration has not been found in IDS of several languages such as Dutch (Benders, 2013), Norwegian (Englund & Behne, 2005), and German (Audibert & Falk, 2018). This is contrary to the existing findings from English (Burnham et al., 2002), Russian, Swedish (Kuhl et al., 1997), Mandarin Chinese (Liu et al., 2003), and now Spanish and Basque. The

source of these cross-linguistic differences requires further study since it is plausible that instead of or in addition to vowels, acoustic exaggeration may be evidenced in other phonetic categories. For instance, this has been seen for Norwegian where there is consonant but not vowel exaggeration (Englund, 2005). Second, it has been proposed that expansion of the vowel triangle is accompanied by increased variability within categories and reduced distance between non-corner vowel categories, which would be expected to increase the difficulty of discriminating these contrasts for young infants instead of facilitating it (Cristia & Seidl, 2014). This argument has received support from computational models of vowel categorisation from IDS (e.g., Martin et al., 2015; McMurray et al., 2013), but it is at odds with the growing evidence that infants' speech processing is enhanced when they encounter acoustically exaggerated vowels (Peter et al., 2016; Song et al., 2010), and that their language processing skills are correlated with vowel exaggeration in their mothers' IDS (this study and Dilley et al., 2020; Garcia-Sierra et al., 2021; Hartman et al., 2017; Kalashnikova & Burnham, 2018; Liu et al., 2003). These contradictory findings point to the need for further research to understand the interactions between the specific acoustic adjustments in IDS (e.g., segmental exaggeration, segmental variability, pitch raising and variability, among others), and how infants perceive and employ these components of their language input at different stages of language development.

Our findings further confirm the complexity of the relation between maternal IDS and infants' speech perception abilities, but they remain silent about the exact source of this relation. Unlike Liu et al. (2003), we show that this relation was only significant for the older infants in our sample (but note that the younger infants in this study were younger than the younger group in Liu et al.; 5 months vs. 6 to 8 months). Previous research has proposed that caregivers adjust the properties of their IDS according to their infants' linguistic needs (Kalashnikova et al., 2018; Kitamura & Burnham, 2003) and the communicative context of the

interaction (Lam & Kitamura, 2012; Smith & Trainor, 2008); and infants' reliance on specific IDS cues changes with age (Kalashnikova & Burnham, 2018). As demonstrated here and in previous research, the degree of vowel hyperarticulation remains stable across infants' first two years of life (Burnham et al., 2015; Kalashnikova & Burnham, 2018; Liu et al., 2009), but it is possible that infants benefit from accumulated exposure to this IDS component when crossing significant developmental milestones. This would explain why a correlation between this component of maternal IDS and infants' speech perception is observed only for non-native discrimination scores in nine-month-olds, which is the only test that revealed emerging differences between younger and older infants' speech perception performance in this study.

The conclusions of our study are limited by the analysis of vowel area hyper-scores as the sole acoustic property of IDS. The use of this measure has several benefits. First, acoustic exaggeration of vowels has been demonstrated to be one of the most characteristic features of clear speech and a robust correlate of speech intelligibility (Bradlow et al., 1996; Smiljanic & Bradlow, 2009). Furthermore, measures of the acoustic space between the three corner vowels allow for direct cross-linguistic comparisons as these vowels are found in the phonological inventories of most languages spoken around the world. However, it is highly likely that additional interactional, lexical, prosodic, and phonetic components of IDS support the development of speech perception during different stages of this process (Rowe & Snow, 2020). The study of these components is particularly relevant to understanding the role of IDS in early bilingual language acquisition since it is plausible that bilingual caregivers make additional language-specific prosodic and acoustic adjustment in IDS of each language, which would assist their infants' learning of the two languages' phonological inventories.

The correlational nature of our design does not allow us to discard other potential explanations for this input-output relation. For instance, it is possible that a shared genetic trait accounts for both speech production clarity in the mother and speech perception ability in the

infant. However, if this were the case, we should have observed similar correlations in our five- and nine-month-old groups. Our results also raise the important question of how quantity of IDS exposure supports the emergence of these relations. In our bilingual sample, IDS properties did not differ when mothers spoke in their infants' dominant and non-dominant language, but only IDS in the dominant language correlated with infants' speech perception. Recent evidence from large-scale studies investigating infants' preference to IDS over ADS suggests that quantity of exposure does play a role in infants' processing of IDS (Byers-Heinlein et al., 2021; ManyBabies Consortium, 2020). Monolingual and bilingual infants show greater preferences for IDS produced in their native language (and particular dialectal variety), and in bilinguals specifically, the size of IDS preference in one of infants' languages is positively related to the amount of exposure to that language. Indeed, greater attention to IDS has been suggested to lead to successful encoding of linguistic information conveyed in this register (Kalashnikova et al., 2018), so it may be the case that as bilingual infants develop greater preference for IDS in the language that they encounter more often in their environment, they extract greater benefit from this type of input. This also converges with previously shown significant relations between the quantity of single language exposure and bilingual children's neural discrimination of native speech sounds and emerging lexical skills (Garcia-Sierra et al., 2016; Hurtado et al., 2014), so it is possible that the role that IDS quality plays in bilingual infants' language development is modulated by the amount of infants' exposure to IDS in that language. This is left as an open question until the quantity and quality of IDS can be assessed simultaneously in this population, and until the quality of speech by more than one caregiver is assessed in order to capture a complete picture of bilingual infants' dual language environment, especially in the case of infants growing up in households where the one parent-one language type of exposure is preferred.

To conclude, these findings indicate that early speech perception development follows a comparable trajectory for monolingual and bilingual infants. Bilingual infants' linguistic experience is divided into two languages, and they face the additional challenge of acquiring not one but two linguistic systems. Nevertheless, our results suggest that despite these differences, the early language input experienced by monolingual and bilingual infants is similar in quality, specifically that caregivers produce acoustically exaggerated speech sounds when addressing their infants regardless of infants' bilingualism status. This adjustment, in turn, is related to infants' developing native language speech perception skills as mothers who produce clearer IDS have infants who exhibit more mature speech perception ability manifested in reduced sensitivity to a non-native phonetic contrast. These results support the linguistic function of IDS demonstrating that caregivers unconsciously optimise their speech according to their infants' linguistic needs, and thus provide support for their infants' phonological acquisition in one or two languages.

References

- Albareda-Castellot, B., Pons, F., & Sebastián-Gallés, N. (2011). The acquisition of phonetic categories in bilingual infants: New data from an anticipatory eye movement paradigm. *Developmental Science, 14*, 395–401. <https://doi.org/10.1111/j.1467-7687.2010.00989.x>
- Audibert, N., & Falk, S. (2018). Vowel space and f0 characteristics of infant-directed singing and speech. *Proceedings of the International Conference on Speech Prosody, 2018-June*, 153–157. <https://doi.org/10.21437/SpeechProsody.2018-31>
- Bates, D. (2005). Fitting linear mixed models in R. *R News, 5*, 27–30.
- Benders, T. (2013). Mommy is only happy! Dutch mothers' realisation of speech sounds in infant-directed speech expresses emotion, not didactic intent. *Infant Behavior and Development, 36*, 847–862. <https://doi.org/10.1016/j.infbeh.2013.09.001>
- Bergelson, E., & Swingle, D. (2012). At 6-9 months, human infants know the meanings of many common nouns. *Proceedings of the National Academy of Sciences, 109*, 3253–3258. <https://doi.org/10.1073/pnas.1113380109>
- Boersma, P., & Weenink, D. (2010). *Praat: Doing phonetics by computer (version 5.3.16)*.
- Bosch, L., & Sebastián-Gallés, N. (2003). Language Experience and the Perception of a Voicing Contrast in Fricatives: Infant and Adult Data. *Proceedings of the 15th International Congress of the Phonetic Sciences, Barcelona, 1987–1990*.
- Bosch, L., & Sebastián-Gallés, N. (2001). Evidence of Early Language Discrimination Abilities in Infants from Bilingual Environments. *Infancy, 2*, 29–49. https://doi.org/10.1207/S15327078IN0201_3
- Bosch, L., & Sebastián-Gallés, N. (2003). Simultaneous Bilingualism and the Perception of a Language-Specific Vowel Contrast in the First Year of Life. *Language and Speech, 46*, 217–243. <https://doi.org/10.1177/00238309030460020801>
- Bradlow, A. R., Torretta, G. M., & Pisoni, D. B. (1996). Intelligibility of normal speech I:

- Global and fine-grained acoustic-phonetic talker characteristics. *Speech Communication*, 20, 255–272. [https://doi.org/10.1016/S0167-6393\(96\)00063-5](https://doi.org/10.1016/S0167-6393(96)00063-5)
- Burnham, D., Kitamura, C., & Vollmer-Conna, U. (2002). What's new, pussycat? On talking to babies and animals. *Science (New York, N.Y.)*, 296, 1435. doi.org/10.1126/science.1069587
- Burnham, D., Singh, L., Mattock, K., Woo, P. J., & Kalashnikova, M. (2018). Constraints on Tone Sensitivity in Novel Word Learning by Monolingual and Bilingual Infants: Tone Properties Are More Influential than Tone Familiarity. *Frontiers in Psychology*, 8, 1–14. <https://doi.org/10.3389/fpsyg.2017.02190>
- Burnham, E. B., Wieland, E. A., Kondarova, M. V., McAuley, J. D., Bergeson, T. R., & Dilley, L. C. (2015). Phonetic modification of vowel space in storybook speech to infants up to 2 years of age. *Journal of Speech Language and Hearing Research*, 58, 241–253.
- Burns, T. C., Yoshida, K., Hill, K., & Werker, J. F. (2007). The development of phonetic representation in bilingual and monolingual infants. *Applied Psycholinguistics*, 28, 455–474. <https://doi.org/10.1017/S0142716407070257>
- Byers-Heinlein, K. (2015). Methods for studying infant bilingualism. *Cambridge Handbook of Bilingual Processing*, 133–154.
- Byers-Heinlein, K., Burns, T. C., & Werker, J. F. (2010). The roots of bilingualism in newborns. *Psychological Science*, 21, 343–348. doi.org/10.1177/0956797609360758
- Byers-Heinlein, K., & Fennell, C. T. (2014). Perceptual narrowing in the context of increased variation: Insights from bilingual infants. *Developmental Psychobiology*, 56, 274–291.
- Byers-Heinlein, K., Tsui, A. S. M., Bergmann, C., Black, A. K., Brown, A., Carbajal, M. J., Durrant, S., Fennell, C. T., Fiévet, A.-C., Frank, M. C., Gampe, A., Gervain, J., Gonzalez-Gomez, N., Hamlin, J. K., Havron, N., Hernik, M., Kerr, S., Killam, H., Klassen, K., ... Wermelinger, S. (2021). A Multilab Study of Bilingual Infants: Exploring the Preference

- for Infant-Directed Speech. *Advances in Methods and Practices in Psychological Science*, 4. <https://doi.org/10.1177/2515245920974622>
- Cohen, L. B., Atkinson, H. H., & Chaput, H. H. (2004). *Habit X: A new program for obtaining and organizing data in infant perception and cognition studies (Version 1.0)*. University of Texas.
- Cristia, A., & Seidl, A. (2014). The hyperarticulation hypothesis of infant-directed speech. *Journal of Child Language*, 41, 913–934. <https://doi.org/10.1017/S0305000912000669>
- D’Souza, D., Brady, D., Haensel, J. X., & D’Souza, H. (2020). Is mere exposure enough? The effects of bilingual environments on infant cognitive development. *Royal Society Open Science*, 7, 180191. [dx.doi.org/10.1098/rsos.180191](https://doi.org/10.1098/rsos.180191)
- Danielson, D., Seidl, A., Onishi, K. H., Alamian, G., & Cristia, A. (2014). The acoustic properties of bilingual infant-directed speech. *The Journal of the Acoustical Society of America*, 135, EL95–EL101. <https://doi.org/10.1121/1.4862881>
- Dilley, L., Lehet, M., Wieland, E. A., Arjmandi, M. K., Kondaurova, M., Wang, Y., Reed, J., Svirsky, M., Houston, D., & Bergeson, T. (2020). Individual differences in mothers’ spontaneous infant-directed speech predict language attainment in children with cochlear implants. *Journal of Speech, Language, and Hearing Research*, 63, 2453–2467. https://doi.org/10.1044/2020_JSLHR-19-00229
- Englund, K. T. (2005). Voice onset time in infant directed speech over the first six months. *First Language*, 25, 219–234. <https://doi.org/10.1177/0142723705050286>
- Englund, K. T., & Behne, D. M. (2005). Infant directed speech in natural interaction—norwegian vowel quantity and quality. *Journal of Psycholinguistic Research*, 34, 259–280. <https://doi.org/10.1007/s10936-005-3640-7>
- Ferjan Ramírez, N., Ramírez, R. R., Clarke, M., Taulu, S., & Kuhl, P. K. (2017). Speech discrimination in 11-month-old bilingual and monolingual infants: a

- magnetoencephalography study. *Developmental Science*, *20*, 1–16.
<https://doi.org/10.1111/desc.12427>
- Fernald, A., & Simon, T. (1984). Expanded intonation contours in mothers' speech to newborns. *Developmental Psychology*, *20*, 104–113. <https://doi.org/10.1037/0012-1649.20.1.104>
- Friederici, A. D., Friedrich, M., & Christophe, A. (2007). Brain responses in 4-month-old infants are already language specific. *Current Biology*, *17*, 1208–1211.
doi.org/10.1016/j.cub.2007.06.011.
- Garcia-Sierra, A., Ramírez-Esparza, N., & Kuhl, P. K. (2016). Relationships between quantity of language input and brain responses in bilingual and monolingual infants. *International Journal of Psychophysiology*, *110*, 1–17.
doi.org/10.1016/j.ijpsycho.2016.10.004
- Garcia-Sierra, A., Ramirez-Esparza, N., Wig, N., & Robertson, D. (2021). *Brain and Language Language learning as a function of infant directed speech (IDS) in Spanish : Testing neural commitment using the positive-MMR. 212.*
<https://doi.org/10.1016/j.bandl.2020.104890>
- Garcia-Sierra, A., Rivera-Gaxiola, M., Percaccio, C. R., Conboy, B. T., Romo, H., Klarman, L., Ortiz, S., & Kuhl, P. K. (2011). Bilingual language learning: An ERP study relating early brain responses to speech, language input, and later word production. *Journal of Phonetics*, *39*, 546–557. <https://doi.org/10.1016/j.wocn.2011.07.002>
- Graf Estes, K., & Hay, J. F. (2015a). Flexibility in Bilingual Infants' Word Learning. *Child Development*, *86*, 1371–1385. <https://doi.org/10.1111/cdev.12392>
- Hartman, K. M., Ratner, N. B., & Newman, R. S. (2017). Infant-directed speech (IDS) vowel clarity and child language outcomes. *Journal of Child Language*, *44*, 1140–1162.
<https://doi.org/10.1017/S0305000916000520>

- Höhle, B., Bijeljac-babic, R., & Nazzi, T. (2019). Variability and stability in early language acquisition : Comparing monolingual and bilingual infants ' speech perception and word recognition. *Bilingualism: Language and Cognition*, *23*, 56-71.
doi.org/10.1017/S1366728919000348
- Hualde, J. I. (2004). *Basque phonology*. Routledge.
- Hualde, J. I., Simonet, M., & Nadeu, M. (2011). Consonant lenition and phonological recategorization. *Laboratory Phonology*, *2*, 301–329.
<https://doi.org/10.1515/labphon.2011.011>
- Hurtado, N., Grüter, T., Marchman, V. A., & Fernald, A. (2014). Relative language exposure, processing efficiency and vocabulary in Spanish-English bilingual toddlers. *Bilingualism*, *17*, 189–202. <https://doi.org/10.1017/S136672891300014X>
- Kalashnikova, M., & Burnham, D. (2018). Infant-directed speech from seven to nineteen months has similar acoustic properties but different functions. *Journal of Child Language*, *45*, 1035–1053. <https://doi.org/10.1017/S0305000917000629>
- Kalashnikova, M., Goswami, U., & Burnham, D. (2018). Mothers speak differently to infants at-risk for dyslexia. *Developmental Science*, *21*, 1–15. doi.org/10.1111/desc.12487
- Kalashnikova, M., Peter, V., Di Liberto, G. M., Lalor, E. C., & Burnham, D. (2018). Infant-directed speech facilitates seven-month-old infants' cortical tracking of speech. *Scientific Reports*, *8*, 1-8.
- Keating, P., Linker, W., & Huffman, M. (1983). Patterns in allophone distribution for voiced and voiceless stops. *Journal of Phonetics*, *11*, 277–290. [https://doi.org/10.1016/s0095-4470\(19\)30827-7](https://doi.org/10.1016/s0095-4470(19)30827-7)
- Kitamura, C., & Burnham, D. (2003). Pitch and Communicative Intent in Mother's Speech: Adjustments for Age and Sex in the First Year. *Infancy*, *4*, 85–110.
https://doi.org/10.1207/S15327078IN0401_5

- Kuhl, P. K. (2004). Early language acquisition: cracking the speech code. *Nature Reviews. Neuroscience*, 5, 831–843. <https://doi.org/10.1038/nrn1533>
- Kuhl, P. K., Andruski, J. E., Chistovich, I. A., Chistovich, L. A., Kozhevnikova, E. V., Ryskina, V. L., Stolyarova, E. I., Sundberg, U., & Lacerda, F. (1997). Cross-Language Analysis of Phonetic Units in Language Addressed to Infants. *Science (New York, N.Y.)*, 277, 684–686. <https://doi.org/10.1126/science.277.5326.684>
- Kuhl, P. K., Stevens, E., Hayashi, A., Deguchi, T., Kiritani, S., & Iverson, P. (2006). Infants show a facilitation effect for native language phonetic perception between 6 and 12 months. *Developmental Science*, 9, F13–F21.
- Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2015). Package “lmerTest.” *R Package Version*, 2(0).
- Lam, C., & Kitamura, C. (2012). Mommy, speak clearly: induced hearing loss shapes vowel hyperarticulation. *Developmental Science*, 15, 212–221. <https://doi.org/10.1111/j.1467-7687.2011.01118.x>
- Larraza, S., Molnar, M., & Samuel, A. G. (2020). Phonemic contrasts under construction? Evidence from Basque. *Infancy*, 25, 304–318. <https://doi.org/10.1111/infa.12330>
- Lasagabaster, D. (2018). Chapter 3. Language Use and Language Attitudes in the Basque Country. *Multilingualism in European Bilingual Contexts, December*, 65–89. <https://doi.org/10.21832/9781853599316-005>
- Liu, H.-M., Kuhl, P. K., & Tsao, F. M. (2003). An association between mothers’ speech clarity and infants’ speech discrimination skills. *Developmental Science*, 6, 1–10. <https://doi.org/10.1111/1467-7687.00275>
- Liu, H.-M., Tsao, F.-M., & Kuhl, P. K. (2009). Age-related changes in acoustic modifications of Mandarin maternal speech to preverbal infants and five-year-old children: A longitudinal study. *Journal of Child Language*, 36, 909.

- <https://doi.org/10.1016/j.jmb.2008.10.054>.The
- Liu, L., & Kager, R. (2017). Perception of tones by bilingual infants learning non-tone languages. *Bilingualism*, *20*, 561–575. <https://doi.org/10.1017/S1366728916000183>
- ManyBabies Consortium. (2020). Quantifying sources of variability in infancy research using the infant-directed-speech preference. *Advances in Methods and Practices in Psychological Science*, *3*, 24-52.
- Martin, A., Schatz, T., Versteegh, M., Miyazawa, K., Mazuka, R., Dupoux, E., & Cristia, A. (2015). Mothers Speak Less Clearly to Infants Than to Adults: A Comprehensive Test of the Hyperarticulation Hypothesis. *Psychological Science*, *26*, 341–347. <https://doi.org/10.1177/0956797614562453>
- Maye, J., Weiss, D. J., & Aslin, R. N. (2008). Statistical phonetic learning in infants: Facilitation and feature generalization. *Developmental Science*, *11*, 122–134. <https://doi.org/10.1111/j.1467-7687.2007.00653.x>
- Maye, J., Werker, J. F., & Gerken, L. (2002). Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition*, *82*, B101–B111.
- McMurray, B., Aslin, R. N., & Toscano, J. C. (2009). Statistical learning of phonetic categories: Insights from a computational approach. *Developmental Science*, *12*, 369–378. <https://doi.org/doi:10.1111/j.1467-7687.2009.00822.x>.
- McMurray, B., Kovack-Lesh, K. A., Goodwin, D., & McEchron, W. (2013). Infant directed speech and the development of speech perception: enhancing development or an unintended consequence? *Cognition*, *129*, 362–378. <https://doi.org/10.1016/j.cognition.2013.07.015>
- Molnar, M., Gervain, J., & Carreiras, M. (2014). Within-rhythm class native language discrimination abilities of basque-Spanish monolingual and bilingual infants at 3.5 months of age. *Infancy*, *19*, 326–337. <https://doi.org/10.1111/infa.12041>

- Narayan, C. R., & McDermott, L. C. (2016). Speech rate and pitch characteristics of infant-directed speech: Longitudinal and cross-linguistic observations. *The Journal of the Acoustical Society of America*, *139*, 1272–1281. <https://doi.org/10.1121/1.4944634>
- Orena, A. J., Byers-Heinlein, K., & Polka, L. (2019). What do bilingual infants actually hear? Evaluating measures of speech input to bilingual-learning 10-month-olds. *Developmental Science*.
- Peter, V., Kalashnikova, M., Santos, A., & Burnham, D. (2016). Mature neural responses to infant-directed speech but not adult-directed speech in pre-verbal infants. *Scientific Reports*, *6*. <https://doi.org/10.1038/srep34273>
- Petitto, L. A., Berens, M. S., Kovelman, I., Dubins, M. H., Jasinska, K., & Shalinsky, M. (2012). The “Perceptual Wedge” hypothesis as the basis for bilingual babies’ phonetic processing advantage: New insights from fNIRS brain imaging. *Brain and Language*, *121*, 130–143. <https://doi.org/10.1021/nl061786n>. Core-Shell
- Polka, L., & Bohn, O. S. (2011). Natural Referent Vowel (NRV) framework: An emerging view of early phonetic development. *Journal of Phonetics*, *39*, 467–478. <https://doi.org/10.1016/j.wocn.2010.08.007>
- Polka, L., & Werker, J. F. (1994). Developmental Changes in Perception of Nonnative Vowel Contrasts. *Journal of Experimental Psychology: Human Perception and Performance*, *20*, 421–435. <https://doi.org/10.1037/0096-1523.20.2.421>
- Ramírez-Esparza, N., García-Sierra, A., & Kuhl, P. K. (2017). The Impact of Early Social Interactions on Later Language Development in Spanish–English Bilingual Infants. *Child Development*, *88*, 1216–1234. <https://doi.org/10.1111/cdev.12648>
- Ramon-Casas, M., Swingley, D., Sebastián-Gallés, N., & Bosch, L. (2009). Vowel categorization during word recognition in bilingual toddlers. *Cognitive Psychology*, *59*, 96–121. <https://doi.org/10.1016/j.cogpsych.2009.02.002>

- Rivera-Gaxiola, M., Silva-Pereyra, J., & Kuhl, P. K. (2005). Brain potentials to native and non-native speech contrasts in 7- and 11 -month-old American infants. *Developmental Science*, 8, 162–172. <https://doi.org/10.1111/j.1467-7687.2005.00403.x>
- Rosner, B. S., López-Bascuas, L. E., García-Albea, J. E., & Fahey, R. P. (2000). Voice-onset times for Castilian Spanish initial stops. *Journal of Phonetics*, 28, 217–224. <https://doi.org/10.1006/jpho.2000.0113>
- Rowe, M. L., & Snow, C. E. (2020). Analyzing input quality along three dimensions : interactive , linguistic , and conceptual. *Journal of Child Language*, 47, 5–21. <https://doi.org/10.1017/S0305000919000655>
- Schatz, T., Feldman, N. H., Goldwater, S., Cao, X. N., & Dupoux, E. (2021). Early phonetic learning without phonetic categories: Insights from large-scale simulations on realistic input. *PNAS*, 118. <https://doi.org/10.1073/pnas.2001844118>
- Sebastián-Gallés, N., & Bosch, L. (2009). Developmental shift in the discrimination of vowel contrasts in bilingual infants: Is the distributional account all there is to it? *Developmental Science*, 12, 874–887. <https://doi.org/10.1111/j.1467-7687.2009.00829.x>
- Singh, L. (2018). Bilingual Infants Demonstrate Advantages in Learning Words in a Third Language. *Child Development*, 89(4), e397–e413. <https://doi.org/10.1111/cdev.12852>
- Singh, L. (2021). Evidence for an Early Novelty Orientation in Bilingual Learners. *Child Development Perspectives*, 0, 1–7. <https://doi.org/10.1111/cdep.12407>
- Singh, L., Loh, D., & Xiao, N. G. (2017). Bilingual infants demonstrate perceptual flexibility in phoneme discrimination but perceptual constraint in face discrimination. *Frontiers in Psychology*, 8. <https://doi.org/10.3389/fpsyg.2017.01563>
- Smiljanic, R., & Bradlow, A. R. (2009). Speaking and Hearing Clearly: Talker and Listener Factors in Speaking Style Changes. *Linguistics and Language Compass*, 3, 236–264. <https://doi.org/10.1111/j.1749-818X.2008.00112.x>.Speaking

- Smith, N. A., & Trainor, L. J. (2008). Infant-Directed Speech Is Modulated by Infant Feedback. *Infancy*, 13, 410–420. doi/abs/10.1080/15250000802188719
- Soderstrom, M. (2007). Beyond babytalk: Re-evaluating the nature and content of speech input to preverbal infants. *Developmental Review*, 27, 501–532.
<https://doi.org/10.1016/j.dr.2007.06.002>
- Song, J. Y., Demuth, K., & Morgan, J. (2010). Effects of the acoustic properties of infant-directed speech on infant word recognition. *The Journal of the Acoustical Society of America*, 128, 389–400. <https://doi.org/10.1121/1.3419786>
- Stager, C. L., & Werker, J. F. (1997). Infants listen for more phonetic detail in speech perception than in word-learning tasks. *Nature*, 388, 381–382.
<https://doi.org/10.1038/41102>
- Sundara, M., Polka, L., & Molnar, M. (2008). Development of coronal stop perception: Bilingual infants keep pace with their monolingual peers. *Cognition*, 108, 232–242.
<https://doi.org/10.1016/j.cognition.2007.12.013>
- Sundara, M., & Scutellaro, A. (2011). Rhythmic distance between languages affects the development of speech perception in bilingual infants. *Journal of Phonetics*, 39, 505–513.
<https://doi.org/10.1016/j.wocn.2010.08.006>
- Swingley, D. (2009). Contributions of infant word learning to language development. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364, 3617–3632.
<https://doi.org/10.1098/rstb.2009.0107>
- Swingley, D., & Alarcon, C. (2018). Lexical Learning May Contribute to Phonetic Learning in Infants: A Corpus Analysis of Maternal Spanish. *Cognitive Science*, 42, 1618–1641.
<https://doi.org/10.1111/cogs.12620>
- Werker, J. F. (2018). Perceptual beginnings to language acquisition. *Applied Psycholinguistics*, 39, 703–728. <https://doi.org/10.1017/S0142716418000152>

Werker, J. F., & Tees, R. C. (1984). Cross-language speech-perception - evidence for perceptual reorganization during the 1st year of life. *Infant Behavior and Development*, 7, 49–63. [https://doi.org/10.1016/S0163-6383\(84\)80022-3](https://doi.org/10.1016/S0163-6383(84)80022-3)

Werker, J. F., Yeung, H. H., & Yoshida, K. A. (2012). How Do Infants Become Experts at Native-Speech Perception? *Current Directions in Psychological Science*, 21, 221–226. <https://doi.org/10.1177/0963721412449459>

Table 1. *Monolingual and bilingual infants' language background and language exposure patterns.*

		Spanish- dominant (<i>N</i>)	Basque- dominant (<i>N</i>)	Female; Male	Age in months (<i>M, SD</i>)	Dominant language exposure (%)	Non- dominant language exposure (%)
Monolingual	5 months	15	25	16; 24	5.05 (.13)	97.97 (3.06)	1.99 (2.91)
	9 months	16	21	15; 22	9.08 (.14)	97.53 (3.09)	2.54 (3.10)
Bilingual	5 months	12	14	15; 11	5.11 (.24)	60.99 (7.36)	39.30 (7.77)
	9 months	13	15	13; 15	9.10 (.12)	63.83 (7.87)	35.95 (7.78)

Table 2. Output of the Linear Mixed Effects model on monolingual and bilingual 5- and 9-month-old infants' performance ($N = 119$).

	β	SE	t	p
(Intercept)	-0.513	0.141	-3.638	.001
Trial Type [Native]	0.554	0.140	3.957	<.001
Trial Type [Non-Native]	0.608	0.140	4.344	<.001
Age Group [9-mos]	0.324	0.183	1.766	.079
Lang Group [Bilingual]	0.551	0.213	2.585	.010
Test Order [Non-Native first]	0.079	0.112	0.709	.480
Trial Type [Native] \times Age Group [9-mos]	-0.359	0.198	-1.812	.071
Trial Type [Non-Native] \times Age Group [9-mos]	-0.429	0.198	-2.166	.031
Trial Type [Native] \times Lang Group [Bilingual]	-0.181	0.231	-0.786	.432
Trial Type [Non-Native] \times Lang Group [Bilingual]	-0.184	0.231	-0.799	.425
Age Group [9-mos] \times Lang Group [Bilingual]	-0.751	0.293	-2.569	.011
Trial Type [Native] \times Age Group [9-mos] \times Lang Group [Bilingual]	0.367	0.316	1.159	.247
Trial Type [Non-Native] \times Age Group [9-mos] \times Lang Group [Bilingual]	0.318	0.317	0.988	.324

List of Figures

Figure 1. Looking duration (sec) in response to the habituation /pa/, native /ba/, and non-native /p^ha/ stimuli in the speech perception task by five- and nine-month-old monolingual and bilingual infants (the internal line represents the Median, and the hinges extend to the first and third quartiles).

Figure 2. Vowel triangles for IDS and ADS produced by mothers of 5- and 9-month-old monolingual and bilingual infants in Spanish and Basque.



