Infants’ First Words are not Phonetically Specified:

Own Name Recognition in British English-learning 5-Month-Olds

Claire Delle Luche (1)

Caroline Floccia (2)

Lionel Granjon (3)

Thierry Nazzi (3)

(1) Centre for Research in Language Development throughout the Lifespan;

Department of Languages and Linguistics, University of Essex, UK

(2) School of Psychology, Plymouth University, UK

(3) Laboratoire Psychologie de la Perception UMR 8242

CNRS - Université Paris Descartes, France

Address for correspondence: Claire Delle Luche, Department of Language and Linguistics, Room 4.342, University of Essex, Wivenhoe Park, Colchester CO4 3SQ, United Kingdom; c.delleluche@essex.ac.uk; tel +44 1206 872113

Acknowledgements

This research was supported by a British-French ESRC/Agence Nationale de la Recherche grant awarded to Floccia (ESRC ES/H040927/1) and Nazzi (ANR RECONVO). The authors would like to thank Camillia Bouchon for providing the original data set and Thibaut Fux for his help in data processing.

Keywords: Word learning – consonants – vowels – speech perception – French and English.

Report Narrative

- Little is known about how very young infants represent or process familiar words at the onset of lexical acquisition.
- A previous study has found that French-learning 5-month-olds could detect a vowel change in their own name, but not a consonant change (Bouchon et al., 2015), and that this ability depends on the acoustic distance between vowels.
- Because English is acoustically very different from French, here we test British English-learning infants for the recognition of their name across vowel and consonant mispronunciations.
- British English 5-month-olds fail to systematically detect mispronunciations, but show a reliance on intensity contrasts.
- These results indicate that infants process and represent early words in a language-specific way, and pay attention to different acoustic dimensions.

Abstract

By the end of their first year of life, infants' representations of familiar words contain phonetic detail; yet little is known about the nature of these representations at the very beginning of word learning. Bouchon et al. (2015) showed that French-learning 5-month-olds could detect a vowel change in their own name and not a consonant change, but also that infants reacted to the acoustic distance between vowels. Here we tested British English-learning 5-month-olds in a similar study to examine whether the acoustic/phonological characteristics of the native language shape the nature of the acoustic/phonetic cues that infants pay attention to. In the first experiment, British English-learning infants failed to recognise their own name compared to a mispronunciation of initial consonant (e.g., Molly vs Nolly) or vowel (e.g., April vs Ipril). Yet in the second experiment they did so when the contrasted name was phonetically dissimilar (e.g., Sophie vs Amber). Differences in phoneme category (stops vs continuants) between the correct consonant versus the incorrect one significantly predicted infants' own name recognition in the first experiment. Altogether, these data suggest that infants might enter into a phonetic mode of processing through different paths depending on the acoustic characteristics of their native language.

Infants’ First Words are not Phonetically Specified:
Own Name Recognition in British English-learning 5-Month-Olds

Around their first birthday, infants’ representations of familiar words contain specific phonetic and phonological information (e.g., Hallé & de Boysson-Bardies, 1996; Mani & Plunkett, 2010; Poltrock & Nazzi, 2015; Swingley, 2005; Vihman, Nakai, DePaolis & Hallé, 2004; Yoshida, Fennell, Swingley & Werker, 2009). This ability to process fine-grained detail in speech scaffolds the building of their expanding lexicon, which will grow rapidly over the second year of life (e.g., Nazzi & Bertoncini, 2003). To achieve such level of phonetic and phonological knowledge by the end of their first year, infants not only rely on their learning of native consonant categories by the age of 10-12 months (e.g., Werker & Tees, 1984) and vowel categories by the age of 6 months (e.g., Kuhl, 1991), but also on their acquisition of perceptual constancy for speech sounds, that is, the ability to identify the same phoneme across different surrounding contexts (e.g., Hochmann & Papeo, 2014; Houston & Jusczyk, 2000). Prior to that stage, the relative weight of acoustic versus phonetic detail[1] in early speech sound representations is not fully understood (Benavides-Varela, Hochmann, Macagno, Nespor & Mehler, 2012; Eimas, 1975; Bertoncini, Bijeljac-Babic, Jusczyk, Kennedy & Mehler, 1988; Jusczyk, Pisoni & Mullenix, 1992; Jusczyk, Pisoni, Reed, Fernald & Meyers, 1983).

The large body of literature assessing young infants’ ability to discriminate speech sounds (for a review, see Kuhl et al., 2008) suggests that the gradual construction of fine-grained phonetic representations towards the end of the first year might rely on their initial

---

[1] Acoustic detail refers to durational, spectral or energy information characterising speech sounds; phonetic detail refers to an abstract representation of these sounds based on a combination of acoustic, articulatory and perceptual properties.

sensitivity to acoustic information (e.g., Narayan, Werker & Beddor, 2009). For example, spectral information such as pitch, which is central to identifying vowel quality (and tonal information), can be used by newborns or older infants to discriminate speech sequences (Bull, Eilers & Oller, 1985; Nazzi, Floccia & Bertoncini, 1998); it is also largely emphasized in infant-directed speech through the use of exaggerated pitch contour and vowel hyperarticulation, which turn out to be very effective cues for vowel discrimination at 6 months (Burnham, Kitamura & Vollmer-Conna, 2002; Kuhl et al., 1997; but see Song, Demuth & Morgan, 2010, for evidence at a later age).

Temporal information, which is critical for the characterisation of consonant contrasts such as voicing or place, is also found to be perceptually salient in infancy, with 5-to-11-month-old infants discriminating fine timing cues such as vowel duration (Eilers, Bull, Oller & Lewis, 1983) and voice onset time (VOT). Infant-directed speech usually exaggerates VOT (Englund, 2005), probably supporting infants' discrimination of VOT differences across phoneme categories (Eimas, Siqueland, Jusczyk & Vigorito, 1971) and even within (Aslin, Pisoni, Hennessy & Perey, 1981).

Finally, energy or amplitude, which is one cue to distinguish vowels from consonants, stops from continuants (e.g., Stevens & Blumstein, 1981) and stress patterns (Jusczyk, Cutler & Redanz, 1993), can be used by 5-to-11-month-olds to discriminate multisyllabic sequences (e.g., Bull et al., 1985), and has been found to modulate infants' abilities to extract and discriminate speech sounds. For example, Polka, Colantonio and Sundara (2001) attributed the surprising non-discrimination of the English /d/-/ð/ contrast at 10-12 months to the low energy of these phonemes.

In sum, attention to acoustic information might support the learning of language-specific phonetic categories. Given that the earliest reports of infants' learning of words in their native language are around 6 months of age (e.g., Bergelson & Swingley, 2012; Tincoff & Jusczyk, 1999, 2012), long before phonetic categories are fully established, it is an open question as to which degree of phonetic specificity is found in these first words.

Besides evidence of early word comprehension at 6 months, Mandel, Jusczyk and Pisoni (1995) also showed auditory name recognition by 5 months. They presented 5-month-olds in a head turn procedure with their own name (e.g., Joshua), a stress matched name (Agatha), and two stress-mismatch names (Maria, Eliza). Infants preferred their name over any other substitute, especially over the stress matched foil, suggesting that their own name representation entails more detail than a broad prosodic pattern.

If 5-month-olds have encoded more than a global prosodic contour of one of their earliest words, their own name, what kind of information do they retain in this representation? Bouchon, Floccia, Fux, Adda-Decker and Nazzi (2015) showed that French-learning 5-month-old infants recognise their own name when the initial consonant (e.g., Zictor for Victor) is changed, but not when the initial vowel is changed (e.g., Elix for Alix) , suggesting some level of phonetic specificity at that early age, and reinforcing the idea of higher perceptual saliency of vowels in early infancy (Benavides-Varela et al., 2012; Bertoncini et al., 1988). However, the authors also reported a significant correlation between infants' behaviour and some acoustic features related to vowels: infants tested on a vowel change in their name were more likely to prefer their correctly produced name (e.g., Alix) over a mispronunciation (e.g., Elix) if the acoustic distance between the two vowels was large (here, between [a] and [e]), irrespective of their phonetic distance (which was kept constant, using only 1-feature changes). The acoustic distance between vowels was estimated by comparing MFCC coefficients. This suggests that 5-month-olds would not necessarily represent phonetic information like adults in terms of features (e.g., Lahiri & Reetz, 2010) or prototypes (e.g., Iverson et al., 2003; Kuhl, 1991), but could rather use acoustic-based metrics (see Curtin, Fennell & Escudero, 2009, for such a proposal at 15 months).

In the present study, we embarked on an extension of Bouchon et al.'s study (2015) to British English, for a number of reasons. First, English and French are contrasted on a number of acoustic and phonological parameters that could impact differently task complexity when infants are presented with minimally different stimuli (Pater, Stager & Werker, 2004). Second, Bouchon et al. (2015) were searching for an early asymmetry in the processing of vowels and consonants, based on the proposal by Nespor, Peña and Mehler (2003) that these two phonological categories serve different functions in language processing (with consonants providing lexical information and vowels syntactic and prosodic information). While such an asymmetry was found in French-learning infants (e.g., Havy & Nazzi, 2009; Nazzi, 2005), recent findings on young British English and Danish learners suggest that its development follows a language-specific path (e.g., Floccia, Nazzi, Delle Luche, Poltrock & Goslin, 2014; Højen & Nazzi, 2016; Mani & Plunkett, 2007, 2008). Therefore, it was necessary to examine the behaviour of English-learning infants to provide a cross-linguistic evaluation of the findings. Finally, Bouchon et al.'s unexpected finding that own name recognition was somehow reliant on the acoustic distance between vowels called for further investigation into the weight of acoustic versus phonetic information at the onset of lexical acquisition.

English provides an interesting comparison with French first because of lexical stress, which can fall in variable syllable positions (despite being predominantly trochaic) within words and is linked to vowel reduction in unstressed syllables. Acoustic correlates of stress bear on energy, duration and spectral characteristics in a continuous and complex way (Fear, Cutler & Butterfield, 1995). In contrast, there is no contrastive lexical stress in French (Dell, 1984; Hirst, DiCristo & Nishinuma, 2001), although the final syllables of content words are usually lengthened in phrase-final positions, hence a phrasal rather than lexical assignment (Christophe, Dupoux, Bertoncini & Mehler, 1994; Delattre, 1966). Moreover, there is far less vowel reduction in French than in British English (White, Mattys & Wiget, 2012; for French vs American English, see Delattre, 1969). At 9/10 months of age, English-learning infants have been shown to distinguish stress-initial from stress-final words (Jusczyk et al., 1993), whereas French-learning infants only succeed in easier discrimination tasks, that is when tested without phonetic variability and when given long enough familiarisations (Bijeljac-Babic, Serres, Höhle & Nazzi, 2012; Skoruppa et al., 2009).

In addition, although the two languages have a comparable ratio of consonants versus vowels (e.g., Floccia et al., 2014), the vowel systems in French and English are also highly dissimilar. Delattre (1964) summarised the differences between (American) English and French as follows: ‘Comparatively, English vowels are predominantly low, back, unrounded, with a strong tendency to center the short and unstressed [...]. Duration contributes to vowel distinctions. All English vowels are more or less diphthongized. Comparatively, French vowels are predominantly high, fronted, rounded, and extreme [...]. Duration is negligible in vowel distinction. There is no diphthongization.’ (p. 82). Although this comparison was formulated for American English, description of British English tends to confirm these differences with French (Giegrich, 1992, p. 17: ‘many English vowels are at least slightly diphthongised in some or even most accents’). In addition, even though it is not quantified, the study by Dodane and Al-Tamimi (2007) shows a reduced vocalic triangle dispersion for British English compared to French. Regarding consonants, English also tends to have heavier syllables than French in terms of number of consonants, which results in greater variability in consonant clusters and in their duration (Ramus, Nespor & Mehler, 1999; White et al., 2012). Altogether, this suggests that English-learning infants might face a more difficult task when learning their phonetic or phonemic categories in comparison to French-learning infants: not only is stress information more variable in English, but the acoustic distance between vowels of English appears to be less informative than in French, and consonant clusters are heavier. Based on these comparisons, one might expect English-learning 5-month-olds to have more difficulties than French-learning ones in perceiving phonetic - and in particular vocalic - changes in their own name.

On the other hand, it could be argued that the target phonemes and their mispronunciations could be more salient in the English stimuli than the French ones. In the name recognition task, English-learning infants presented with a dissyllabic name will hear a mispronunciation on the initial, stressed, syllable. In contrast, French-learning infants, most of them presented with disyllabic names as in Bouchon et al. (2015), were presented with a mispronunciation on the initial, unaccented syllable (as French words are lengthened on their final syllable).

The second rationale for this study related to a possible cross-linguistic difference in the emergence of the functional difference between vowels and consonants (Nespor et al., 2003) during language acquisition. Although English and French adult listeners seem to display an equal ‘consonant bias’ in auditory lexical processing (Delle Luche et al., 2014), this consonant bias does not seem to emerge simultaneously in the course of language acquisition. Whereas French-learning infants display a clear-cut consonant bias from the age of 11 months onwards (e.g., Poltrock & Nazzi, 2015), such a bias has not been found robustly until the age of 30 months in British English-learning infants (Nazzi, Floccia, Moquet & Butler, 2009). Before that age, British English-learning infants show an equal sensitivity to vowels and consonants (Floccia et al., 2014; Mani & Plunkett, 2007, 2008). Therefore, the developmental trajectory found in French-learning infants of a vowel bias at 5 months (Bouchon et al., 2015) followed by a consonant bias at 11 months (Poltrock & Nazzi, 2015) would not necessarily extend to British English.

Taken together, these elements point to possible differences in the way British English- and French-learning 5-month-olds process vocalic and consonantal information in familiar words. Whereas French-learning infants showed sensitivity to vocalic contrasts – but not consonant contrasts - in a lexical context (Bouchon et al., 2015), British-English-learning infants might show a less clear-cut asymmetry to these two phonological categories. In addition, because of the highly contrastive stress system of English combined with the greater variability of vowels and consonant cluster durations, they might pay more attention than French-learning infants to acoustic dimensions such as energy or duration rather than to phonetic cues.

In Experiment 1, following Bouchon et al. (2015), we used a head-turn procedure to test the preference of 5-month-olds for their own name versus a mispronounced version. In one test group, infants’ names started with a consonant and the mispronunciation bore on this first consonant (e.g., Molly vs Nolly); in the second test group, infants’ names started with a vowel and the mispronunciation bore on this vowel (e.g., April vs Ipril). To control that observed performance of the test infants is related to lexical representations rather than uncontrolled properties of the recordings, we also tested two control groups: Consonant Change and Vowel Change control groups were presented with the names from the test group in their correct and incorrect versions, after ensuring that control infants had had no exposure to that particular name (for example, a child named Robin would be presented with Molly vs Nolly, after enquiring that there was no one named ‘Molly’ in her environment). For those two groups, no preference for the correct version over the incorrect one was expected; if any difference emerged it would be due to discrimination at the pre-lexical level, and not at the lexical level as for the test groups.

If English-learning infants behave like their French-learning peers, infants in the Vowel Change test group should prefer their own name to its mispronunciation; no preference for correct or incorrect names should be found in the Consonant Change test group and the two control groups. Additional analyses of the relation between infants’ behaviour in response to correct and incorrect pronunciations and the acoustic distance between these pronunciations will be undertaken. If such links were found, similar to those found in the French study, we could reasonably conclude that by 5 months, across two highly phonologically distinct languages, infants’ sensitivity to acoustic and phonetic information follows a similar developmental path.

On the contrary, if English-learning infants face a more difficult task than their French peers because of the aforementioned differences between these two languages (see White et al., 2012), they may have more difficulties accessing fine grained phonetic information. In particular, because of the less contrasted acoustic and perceptual cues in the

English vowel system, infants may face a harder task detecting a vowel change in their name. That is, phonetic changes could be less salient compared to French, resulting in smaller spectral distance (MFCC) and a more difficult detection. Rather, we could expect a range of correlations between infants' behaviour and the acoustic characteristics of the stimuli such as duration, intensity, pitch, and formants, reflecting their early sensitivity to the correlates of stress information. If such results were found, it would signal that the early speech perception system develops in a language-specific way in terms of phonetic processing during the first months of life.

Experiment 1

Two groups of British English-learning infants were tested in a head turn preference task for their own name (correct pronunciation or CP) versus a mispronunciation (MP) of their name on its first consonant (Consonant Change test group) or vowel (Vowel Change test group). In two additional groups, infants were presented with the CP and MP of another child (consonant change control group and vowel change control group). The procedure is identical to that used in Bouchon et al. (2015), except for a few minor changes that will be signalled below.

Participants. All 120 participants were healthy British-English-learning monolingual 5-month-old infants (see Table 1). When parents were invited to participate, no mention was made of the precise condition in which the child would be tested (we only mentioned word recognition, not name recognition). A long questionnaire was sent to the parents with irrelevant questions alongside the crucial question 'what is(are) the name(s) you usually employ to address your child?' Based on the parents' response, we only assigned to the test condition infants who were almost exclusively called by their own name or by nicknames (as

‘Monkey’ and ‘Pops’). Only infants with a monosyllabic or a trochaic disyllabic name were included in the test group. Infants in the control conditions were chosen so that they would not know anyone in their environment with the name used in the experiment (this was checked on the day of testing, as in Bouchon et al., 2015). As much as possible, control children with names starting with a single consonant were assigned to the Consonant Change control group (true for 28 infants) and those with names starting with a vowel to the Vowel Change control group (true for 18 infants). The data of 29 additional infants were excluded due to fussiness (17), being inattentive (6), experimenter error (2), being an outlier (difference in looking times between CP and MP above or below 2 standard deviation of the group mean; 4).

Insert Table 1

Stimuli. Each of the 60 test infants heard repetitions of stimuli corresponding to their CP and MP names. Due to a few infants having the same names, we used 27 pairs in the Consonant Change condition and 22 pairs in the Vowel Change condition. The same stimuli were presented to the 60 yoked control infants. The MP of the names always consisted of a 1-feature change (Table 2). In the Consonant Change condition, 10 infants were tested on a place-of-articulation change, 10 on a voicing change, and 10 on a manner-of-articulation change. In the Vowel Change condition, 15 infants were tested on a place change and 15 on a height change.

A female native English speaker recorded 15 tokens each of CP and MP names, in isolation, and in a child friendly, affirmative tone. To achieve this, she produced a series of approximately 30 CPs in close succession, followed by 30 MPs. The 15 best tokens were selected in each list so that MPs were comparable to CPs in terms of intonation patterns and durations. For both CPs and MPs, two files including the 15 tokens were created, the second file presenting the same tokens in reversed order. As the experiment spanned over a long period, three different speakers were used (adding this factor in the main ANOVA on looking times did not impact the results). Each CP-MP pair, however, was produced by the same speaker. All sound files lasted 24 s, and the tokens were normalized for amplitude with Praat (Boermsa & Weenink, 2010).

Insert Table 2

Procedure. After informed consent was obtained from the parent, the child sat on the parent's lap in the experimental head-turn booth. The parent wore headphones playing loud music to mask the auditory stimuli. The experimenter sat outside but could observe the infant via a video camera (without sound capture). At the beginning of each trial, a green light flashed directly in front of the infant to get her attention When the experimenter judged that the infant was looking at the central light, she started the trial, causing the green light to stop and a red light to start flashing either to the left or the right side (randomised and counterbalanced) of the infant. When the infant turned towards the side where the light was flashing, a sound file was played via a speaker located just below the flashing light, until the end of the file, or until the child looked away for more than 2 s. Any trial during which the infant looked away within 1.5 s was aborted and the word list was repeated. If two consecutive trials were aborted, or if three trials were aborted during the experiment, the participant's data was excluded for inattentiveness. The session consisted of 12 trials divided in 3 blocks; in each block, the two CP and the two MP lists were presented (Bouchon et al., 2015, used 8 trials instead of 12). Order of the different lists within each block was randomized. Half of the children in each group heard a CP in the very first trial, whereas the other half heard an MP.

Acoustic analyses of the stimuli

To allow a comparison with the analyses presented in Bouchon et al. (2015), three acoustic dimensions were measured to characterise the contrasted phonemes of CPs and MPs: duration, intensity, and Mel Frequency Cepstral Coefficients as a measure of spectral distance (MFCCs; see Bouchon et al., 2015, for detailed explanations). This distance is based on spectral information regardless of intensity, and collapsed across durations by adjusting a Dynamic Time Warping. Departing from Bouchon et al., pitch and individual formant values were also measured on vowel segments. While pitch, formants, duration and intensity are indicators of acoustic saliency, MFCCs provide a perceptually relevant measure of phonemic distance (Davis & Mermelstein, 1980) that has been widely used in automatic speaker and speech processing (e.g., Patel & Rao, 2011).

For each CP/MP pair, duration, intensity and, for vowels, pitch and formants were measured for the 15 tokens of the contrasted phonemes using PRAAT (for intensity, the minimum pitch was constrained by Praat at 1000 Hz, creating windows of analyses of 3.2 ms). This was first used to calculate mean duration, intensity, pitch and formant values of the contrasted phonemes, in order to compare the relative salience and discriminability of the contrasted consonants and vowels. Second, we computed normalised duration, intensity, pitch and formant differences (*Diff.duration*: duration difference between the contrasted phonemes of CPs and MPs divided by their mean; *Diff.intensity* and *Diff.pitch* are defined similarly), in order to test their relation with individual performance. For formant analysis, the F1, F2 and F3 values of each token were measured at 50% of the vowel duration and converted in mel; for each CP/MP pair the median of each formant for the 15 MP tokens minus the median for the 15 CP tokens was calculated. This will be referred to as *Diff.medianF1*, *Diff.medianF2*, and *Diff.medianF3.*

MFCCs were calculated in exactly the same way as in Bouchon et al. (2015) with the French stimuli. The subset of MFCCs employed in the classification to measure MFCC distances included 12 coefficients, c1 to c12, in order to best represent the envelope of the mel-spectrum. MFCC distances correspond to the Euclidian distance between two tokens calculated for the 12 coefficients (i.e., the square root of the summed squared differences between the two MFCC sets). The relation between individual performance and the normalised MFCC distance between CPs and MPs was tested using *Diff.spectral*, defined for each CP/MP pair as the ratio of the mean cross-category distance between the 15 CPs and the 15 MPs of a given pair and the mean internal variability within the 15 CPs and the 15 MPs of that pair.

Results

Overall analysis. Mean listening times (LTs) to the CP and MP names were calculated for each infant[2]. Group averages are presented in Figure 1. A three-way ANOVA was conducted on LTs with the within-subject factor of pronunciation (CP vs MP) and the between-subject factors of group (Test vs Control) and condition (Consonant Change vs Vowel Change). Neither the effect of condition ($F(1, 116) < 1$), group ($F(1, 116) = 1.21$; $p = .27$), pronunciation ($F(1, 116) < 1$) nor the condition × pronunciation interaction ($F(1, 116) < 1$) reached significance. In addition, neither the pronunciation x group interaction ($F(1, 116) < 1$), the group x condition interaction ($F(1, 116) = 1.14$, $p = .27$) nor the 3-way interaction between pronunciation x group x condition ($F(1, 116) < 1$) reached significance, establishing

---

[2] Analyses conducted with 8 trials instead of 12 trials to compare with Bouchon et al.'s study yielded similar results, both for behavioural responses and relations between these and the acoustic characteristics of the stimuli.

that infants were behaving in the same way in all 4 sub-groups. Finally, when feature (place, voicing or manner; all n = 10) was introduced in the ANOVA conducted on the two Consonant Change groups, no main effect of feature or interaction with other factors was found. Similarly, when the factor feature (place or height; all n = 15) was included in the two Vowel Change groups, again, no effect of feature and no interaction involving feature was found. Overall, these results do not show that British-English-learning 5-month-olds are sensitive to a consonant or a vowel mispronunciation in their own name.

Insert Figure 1

Comparison with the French data. We compared these data to those from Bouchon et al. (2015), using *LT.diff* as a dependent variable, which is the difference between CP and MP, with positive values indicating a preference for the correctly produced name. We ran a repeated measure ANOVA on *LT.diff* with group (Test vs Control), condition (Consonant Change vs Vowel Change) and language (French vs English). There was first a main effect of group ($F(1, 232) = 4.02$, $p = .046$, $\eta^2_p = .02$), due to a larger *LT.diff* for test infants than control infants (.73 vs -.01 for control infants), suggesting that overall, infants tended to prefer their correctly pronounced name over the mispronounced version, across all phoneme conditions and languages. However, the triple interaction between language, condition and group was significant ($F(1, 232) = 5.11$, $p = .025$, $\eta^2_p = .02$). In addition, the interaction between group and language was marginally significant ($F(1, 232) = 3.15$, $p = .08$, $\eta^2_p = .01$), due to a marginal difference between the test groups (mean *LT.diff* : French 1.18; English 0.28; $F(1, 116) = 2.98$, $p = .09$, $\eta^2_p = .03$), but no significant difference between the two control groups (French -.21; English .19; $F(1, 116) < 1$). The difference between the French and English vowel change test groups approached significance (mean *LT.diff*: French 1.86, English 0.44, $t(58) = 1.81$, $p = .075$), and so did the difference between the French and

English vowel change control groups (mean *LT.diff*: French -1.06, English 0.50, $t(58) = 1.92$, $p = .060$). However, no difference was found between the French and English consonant change test groups ($t(58) < 1$) or control groups ($t(58) = 1.15$). Therefore, although infants tended to show preference for their own name over its mispronunciation when languages and phoneme conditions are pooled together, this effect was mainly driven by French-learning infants presented with vowel changes.

In what follows, we conducted a thorough investigation of the links between acoustic dimensions in the stimuli and infants' listening behaviour, providing similar analyses as in Bouchon et al. (2015), with the addition of pitch and formant measures for vowels.

Acoustic measures

Acoustic measures were performed on the stimuli (see Stimuli section and Figure 2 for details), which consisted of 27 different CP/MP pairs in the Consonant Change condition and 22 CP/MP pairs in the Vowel Change condition.

Duration, Intensity, Pitch and Formants. Regarding acoustic salience, the initial consonants lasted 81.5 ms (SD 46.3) and were 61.6 dB (SD 7.8) loud on average, while the initial vowels lasted 112.8 ms (SD 45.6) and were 78.7 dB (SD 2.1) loud. Consonants were significantly shorter (duration: $F(1, 47) = 8.44$, $p = .006$, $\eta^2_p = .15$) and softer (intensity: $F(1, 47) = 172.7$, $p < .0001$, $\eta^2_p = .79$) than vowels, hence establishing, as expected, that consonants were less acoustically salient than vowels.

Regarding discriminability, within each pair of contrasted phoneme (e.g., the /m/ in Molly vs the /n/ in Nolly), consonant CPs were on average 16.4 ms shorter (95% *CI* = [-29.6, -3.2]) and 0.10 dB louder (95% *CI* = [-2.8, 2.6]) than consonant MPs; vowel CPs were on average 3.6 ms longer (95% *CI* = [-3.9, 11.2]) and .09 dB softer (95% *CI* = [-.24, 0.41]) than vowel MPs. With consonants and vowels collapsed, there was no effect of pronunciation on duration (*F*(1, 47) = 2.31, *p* = .14) or on intensity (*F*(1, 47) < 1). However, there was a pronunciation × condition interaction for duration (*F*(1, 47) = 5.70, *p* = .021, η^2_p = .11) but not for intensity (*F*(1, 47) < 1). This was due to consonant CPs being significantly shorter than consonant MPs (paired t-test *t*(26) = 2.38, *p* = .025).

Finally, adding to Bouchon et al.'s analyses, we also compared mean pitch values in vowels and found no significant difference between CPs (mean 263.5 Hz; SD 41.5) and MPs (mean 261.8 Hz; SD 44.3; paired t-test *t*(21) < 1). Regarding formant values, vowels in CPs had a mean F1 of 853 Hz (mel scale; SD 132) and 771 Hz in MPs (SD 114; *t*(21) = 2.17, *p* = .041). No difference was found for F2 (CP: 1398 Hz, SD 148; MP: 1362 Hz, SD 193) and F3 (CP: 1839 Hz, SD 107; MP: 1859 Hz, SD 95). The difference on the first formant is not surprising given that F1 changes are often associated with height contrasts, which was manipulated in half our stimuli. To summarise these measures on vowels, CPs and MPs could not be reliably distinguished based on intensity, duration and pitch. Vowel CPs and MPs, however, could be distinguished using F1 – as far as numerical distance would necessarily translate into perceptual distance. For consonants, CPs and MPs could not be reliably distinguished based on intensity, but there was a significant 16.4 ms duration difference, the only perceptual cue that may have been exploited by infants.

Regarding spectral measures, the acoustic/phonetic distance (*Diff.spectral,* based on MFCCs) was used to further assess discriminability (as mentioned above, MFCCs are not meaningful with respect to salience). On average, *Diff.spectral* was 1.18 (SE = .03) for consonant contrasts, which was not significantly different than the same index for vowels (1.25, SE = .05; *t*(47) = 1.32, *p* = .19). This establishes that consonant contrasts were not acoustically more distinct than vowel contrasts, once normalised for intensity and duration.

For comparison purposes with Bouchon et al. (2015), Table 3 provides the present values and those found in French. Interestingly, vowels in English were about 16 dB louder than consonants, whereas vowels in French were only 8 dB louder than consonants. In addition, although English phonemes were slightly longer than French ones, MFCC distances were smaller in English than in French, especially for consonant contrasts which might suggest that they are less distinct in English than French, at least in this sample.

Insert Table 3

Acoustic predictors of preference measures

Similarly to Bouchon et al. (2015), we explored whether listening preferences (*LT.Diff*) were driven by acoustic distance between CPs and MPs such as measured by *Diff.duration, Diff.intensity* and *Diff.spectral*. We also evaluated the pitch and formant differences for the analysis of vowel stimuli (*Diff.pitch, Diff.medianF1, Diff.medianF2* and *Diff.medianF3*).

First, a multiple linear regression was run on all 60 test infants (30 Consonant Change, 30 Vowel Change) with *LT.diff* as the dependent variable and the 3 acoustic distances as predictors (for which there was no colinearity, all *VIFs* < 1.14). The model significantly explained 16.5 % of the variance in *LT.diff* ($R^2_{adjusted} = .12$; $F(3, 56) = 3.68$, $p = .017$; SEE = 2.79). This was due mainly to *Diff.intensity* which significantly predicted the difference in LTs between CPs and MPs, but inversely ($\beta_{intensity} = -.34$, $p = .011$; see below). The two other predictors, *Diff.spectral* and *Diff.duration*, did not contribute significantly to the model. The same regression analysis conducted on the 60 control infants yielded a non-significant model explaining 12.1 % of the variance ($R^2_{adjusted} = .07$; $F(3, 56) = 2.57$, $p = .063$; SEE = 2.63).

Second, since it is possible that infants process consonants and vowels differently, leading to different effects of acoustic distance within each category, we re-ran the same regression separately for Consonant and Vowel Changes. In the Consonant Change condition, for the 30 test infants (all *VIFs* < 1.5), the 3-predictor model explained significantly 26.1 % of the variance in *LT.diff* ($R^2_{adjusted}$ = .17; *F*(3, 29) = 3.05, *p* = .046; SEE = 2.64). In this model, *Diff.intensity* was the only significant predictor ($\beta_{intensity}$ = -.47, *p* = .028). For the 30 control infants (all *VIFs* < 1.5), the 3-predictor model did not explain a significant part of the variance in *LT.diff* (R^2 = .16, $R^2_{adjusted}$ = .06; *F*(3, 29) = 1.59, *p* = .22; SEE = 1.73).

In the Vowel Change condition, for the 30 test infants (all *VIFs* < 1.1), the 3-predictor model did not explain any significant part of the variance in *LT.diff* (R^2 = .11, $R^2_{adjusted}$ = .01; *F*(3, 29) = 1.1, SEE = 3.06). The addition of *Diff.pitch* as a predictor in this model did not significantly increase the portion of explained variance (R^2 = .12, $R^2_{adjusted}$ = -.02; *F*(4, 25) < 1; SEE = 3.11). A similar result was found for the 30 vowel control infants, without *Diff.Pitch* (R^2 = .15, $R^2_{adjusted}$ = .06; *F*(3, 29) = 1.58, SEE = 3.34), or with *Diff.Pitch* (R^2 = .17, $R^2_{adjusted}$ = .04; *F*(4, 29) = 1.32, SEE = 3.37). Finally, when the three formant measures were included, the model was still not significant for both test (R^2 = .16, *F*(7, 29) < 1) and control infants (R^2 = .18, *F*(7, 29) < 1).

A first interpretation of the effect of *Diff.intensity* is that infants pay relatively more attention to their correct name than its mispronunciation when the initial MP phoneme has more energy than its CP counterpart. The reason why this effect is found in consonants but not in vowels is possibly due to the distribution of *Diff.intensity* between the two conditions: although the mean *Diff.intensity* values are very similar in consonants (+0.10) and in vowels (-0.09), the values for consonants are much more widespread than for vowels (SE for consonants: 1.32; for vowels: 0.16). This is of course explained by the fact that a one-feature change in consonants can result in qualitatively different phonemes (e.g., Leo/Zeo or Rory/Jory) whereas a one-feature change in vowels still result in another, similar shaped, vowel. Infants are presented with larger intensity differences between CPs and MPs in consonants than vowels, which confirms intensity as an explanatory factor for their discrimination score (this is illustrated in Figure 2, which uses the same scales for the consonant and vowel groups). This is particularly true for manner changes (filled round markers on top, Figure 2) which are widely spread along the intensity difference axis. This distribution is inherent to the type of change: voicing and place do not affect intensity of the consonants as much, but changing manner pairs up quiet and loud consonants.

Insert Figure 2

Another way to explain the effect of *Diff.intensity* is that infants would be differently sensitive to initial consonants with a short amplitude-rise-time such as stops, over continuants that have a long amplitude-rise-time. As reported by Nittrouer and Studdert-Kennedy (1986, p. 214, adapted from Mack & Blumstein, 1983), stops and continuants can be distinguished by their ratio of the “rms energy of a brief acoustic segment immediately following release offset to the rms energy of a brief acoustic segment at release onset”. It must be noted however that children from the age of 4 years, just like adults, do not seem to rely on this acoustic property to distinguish for example stops from glides, but rather use formant change information (Nittrouer, Lowenstein & Tarr, 2013).

We included the factor of CP consonant type (continuants vs stops; classification following Skandera & Burleigh, 2011: plosives and affricates are stops and all other consonants are continuants) in an ANOVA on LTs with pronunciation (CP vs MP) and group (test vs control). Out of the 30 Consonant Change test infants, 17 had a name starting with a continuant and 13 with a stop. The only significant effect was the triple interaction between consonant type, pronunciation and group ($F(1, 56) = 6.32$, $p = .015$, $\eta^2 =.10$). This was due to an interaction between pronunciation and group for those infants hearing their name starting with a stop (*F*(1, 24) = 5.35, *p* = .030, η^2 =.18). Test infants in this group tended to listen longer to their own name than its mispronunciation (10.0 s vs 8.8 s, *t*(12) = 1.75, *p* = .11) while the reverse tendency was found for control infants (7.5 s vs 8.2 s, *t*(12) = -1.52, *p* = .16). No interaction between pronunciation and group was found for those infants hearing names starting with a continuant (*F*(1, 32) = 1.70, *p* = .20, η^2 =.05). An identical analysis taking MP consonant type as a factor did not lead to any significant effect (out of the 30 consonant change test infants, 11 heard an MP starting with a stop and 19 with a continuant). A first interpretation of these results is that infants whose name starts with a stop are better at identifying a consonant-initial change; a second, more plausible, interpretation is that infants are better at detecting a stop-to-stop change than a continuant-to-continuant change. Indeed the distribution of changes in our stimuli is such that nearly all stop-initial names changed into a stop-initial mispronunciation (N= 10 out of 13), while nearly all continuant-initial names changed into continuant-initial mispronunciations (N= 16 out of 17).

To relate these findings back to the impact of *Diff.intensity* in the regression analyses, we have ranked the consonant CP/MP pairs according to the value of *Diff.intensity* in Table 4, including information about feature change and phoneme category of the CP and the MP (stop or continuant; Stevens & Blumstein, 1981). First, as was seen in Figure 2, manner changes are mainly found on both sides of the *Diff.intensity* continuum. In addition, it is noteworthy that names starting with a stop are more likely to be grouped together (top lines of Table 4), while continuant-initial names are mainly located at the bottom half the table. This suggests that the correlation between *Diff.intensity* and infants' discrimination scores might be spurious and reflect instead the fact that infants tend to better discriminate stop-to-stop changes than continuant-to-continuant changes. Indeed, pairs of CP/MP leading to the strongest correct name preference are not only showing larger intensity differences, but also have stops as initial phonemes.

Insert Table 4

Discussion

In this first experiment we tested British English-learning 5-month-olds for their preference of their own name versus a mispronunciation involving its initial phoneme, consonant or vowel. Contrary to Bouchon et al. (2015) who conducted a very similar study in French, and found sensitivity to mispronunciation (marked by an own name preference) in the vowel change condition but not in the consonant change condition, we failed to find overall evidence of name preference in either condition. Yet, in terms of task difficulty, English-learning infants were presented with a potentially easier task as compared to the French-learning infants, thanks to stress location: for all infants the mispronunciation bore on the initial, stressed, syllable (54 out of 60 test infants had a trochaic disyllabic name and 6 had a monosyllabic name). In contrast, for the 56 French-learning infants with a disyllabic name (out of 60), the mispronunciation bore on the initial, unaccented, syllable (as French words are lengthened on their final syllable). Therefore, the target phonemes and their mispronunciations were presumably more salient for the English stimuli than for the French ones, yet English-learning infants failed to prefer their name consistently.

However, English test infants' behaviour in the consonant change condition appeared to be predicted by the type of consonant found at the onset of their name: those presented with a CP involving a stop consonant were more likely to prefer their correctly produced name over its mispronunciation, as opposed to those whose name started with a continuant. Importantly, the type of phoneme found in the mispronunciation, stop or continuant, did not predict infants' discrimination pattern, suggesting that infants' representation of their own name was more consistent, or more readily accessed, with a stop in initial position. However, a caveat to this interpretation is that most consonant changes retained category: stops were changed into stops and continuants into continuants (even in most cases when manipulating manner). Therefore our results could be seen as infants being generally more sensitive to stop-to-stop changes rather than continuant-to-continuant changes.

It must be noted that these findings argue against the claim that the predominance of complex consonant clusters in English as compared to French would have, comparatively, enhanced British English learners' sensitivity to vowels. Indeed British English-learning infants showed, if anything, greater sensitivity to consonant changes than to vowel changes, which is the opposite of what was found for French (Bouchon et al., 2015).

Before further discussing these results, and due to the lack of an overall name recognition effect when tested against a one-feature mispronunciation, we decided to validate our general protocol by testing a new group of 5-month-olds on the recognition of their own name versus an entirely different name (for example Victor would hear an alternation of 'Victor' and 'Jacob'; see Mandel et al., 1995). We were also aware of the findings – as yet unexplained - that British English learners' vocabulary scores are significantly smaller than those of their American English counterparts, when measured with parental reports throughout their first three years (Fenson et al., 1994; Hamilton, Plunkett & Schafer, 2000). Therefore, the possibility remained that 5 months was too early to observe own name recognition in British English infants, contrary to what had been reported by Mandel et al. (1995) with American English-learning infants.

Experiment 2

Experiment 2 tested whether British English-learning infants would display a preference for their own name over a phonetically different one (Mandel et al., 1995). Foils were chosen so that they would share the number of syllables and stress pattern of the target names, so as to ensure that infants would not discriminate on supra-segmental features but on segmental information.

Participants. Eighteen full term 5-month-old infants (8 females, 10 males) aged 5 months and 8 days on average (range [4;27 – 6;6]) were tested. Eleven additional participants were excluded: fussiness (8), outlier (1) and technical issues and experimenter error (2).

Stimuli. The same questionnaire as in Exp1 was sent to parents before their visit and only those infants who were addressed consistently by their own name or a nickname were included in the study. As in Exp1, the name had to be either monosyllabic or a trochaic disyllabic one. Again, fifteen tokens of the names and their foils were produced in a child directed friendly voice, by the same speaker. For each name, the paired foil had a different initial phoneme, no phonological overlap overall, the same number of syllables and the same stress pattern (apart from the pair Thomas/Nate, due to experimental error). Half the infants heard their own name in the first trial, and the other half heard the foil in the first trial.

Insert Table 5

Procedure. Identical to that of Exp1.

Results and Discussion

Infants listened longer to their own name (M = 10.66 s, SE = .69) than the foil (M = 9.36 s, SE = .69; *t*(18) = 2.50, *p* = .023). Out of the 18 infants, 13 showed this pattern of results. Given the findings in Exp1 that infants whose name started with a stop were more likely to prefer its correct version to its incorrect version than infants whose name started with a continuant consonant, the same analysis was conducted in Exp2. Out of the 15 names starting with a consonant, 7 started with a stop and 8 with a continuant. An ANOVA with Name (name vs foil) and Initial phoneme (stop vs continuant) did not provide any significant difference, with both types of consonants resulting in a similar name preference (interaction Name x Initial phoneme: $F(1, 13) < 1$). Contrary to Exp1, most changes here involved a category change (stop-to-continuant: N = 6; continuant-to-stop: N = 4; stop-to-stop: N = 2; continuant-to-continuant: N = 0; stop-or-continuant-to-vowel or vice-versa: N = 6). The absence of post-hoc initial phoneme category effect (stop vs continuant) in this second experiment is either due to the small number of observations per cell, or to the quasi-absence of stop-stop or continuant-continuant changes. It could be also due to the task infants face in this experiment, since being presented with their name versus a very different speech sequence might not require infants to pay attention to fine-grained details as in Exp1.

Experiment 2 establishes that when presented with their own name versus a phonetically different unknown name, British English-learning 5-month-olds prefer their own name, as found by Mandel et al. (1995) for American English-learning infants of the same age. This suggests that the failure to find a name preference in Exp1 was not caused by methodological issues, or by British English-learning infants being late in learning their name as compared to American English- and French-learning infants. Rather, it suggests that early word representation in British English-learning infants does not include phonetic details at 5 months of age.

General Discussion

To examine speech perception development between two languages where infants' lexical processing seems to regularly differ (e.g., perception of vowels and consonants: Mani & Plunkett, 2008; Nazzi, 2005; perception of stress: Höhle, Bijeljac-Babic, Herold, Weissenborn & Nazzi, 2009; Jusczyk et al., 1993; Skoruppa et al., 2009), we tested British English-learning infants' recognition of their own name against an initial phoneme mispronunciation that could be either a consonant or a vowel, in a direct replication of a study examining French-learning infants by Bouchon et al. (2015). The contribution of phonetic versus acoustic information to behavioural patterns was also evaluated. To summarise, French-learning 5-month-olds were found to discriminate their name from a mispronunciation on the first vowel, but not on the first consonant. In addition, their listening behaviour in the vowel condition was significantly predicted by the MFCC distance between the paired vowels: that is, the larger the acoustic distance between the initial vowel in the infant's name (e.g., /a/ in Alix) and its mispronunciation (e.g., /e/ in Elix), the more likely they were to prefer their own name. In contrast, name preference did not occur for a consonant change, although consonants were found to be further apart from one another than vowels in terms of MFCC distance. The developmental scenario that was proposed in light of these findings was as follows: infants start lexical processing with a vowel bias, possibly because vowels are more salient than consonants (see also Benavides-Varela et al., 2012; Bertoncini et al., 1988), and to initially distinguish between vowels, they exploit acoustic distance as extracted for example through MFCCs. With further language exposure and maturation, their attention turns towards consonants, which, although less salient in the signal, are more distinct from one another in terms of acoustic/phonetic distance. This switch from a vowel to a consonant bias could be achieved either through lexical development (Bergelson & Swingley, 2012; Tincoff & Jusczyk, 1999, 2012), the construction of phonemic (Werker & Curtin, 2005) and phonetic (Feldman, Griffiths, Goldwater & Morgan, 2013) categories, and/or thanks to their refining temporal resolution abilities (Werner, 1996).

The current results complicate this picture slightly. British English-learning 5-month-olds behave very differently from French-learning 5-month-olds: overall, they do not discriminate an initial phoneme change in their name, neither vowel nor consonant. Moreover, whereas French-learning infants were reliably found to use acoustic/phonetic distance in vowels to discriminate changes, here weaker evidence suggests that English-learning infants exploit instead energy information in consonants. This was observed with a higher propensity to discriminate CP from MP when their name starts with a short amplitude-rise-time consonant (stops) rather than a longer amplitude-rise-time consonant (continuant). Given the distribution of category changes found in our stimuli, another, more plausible, interpretation is that infants were actually more sensitive to stop-to-stop changes than continuant-to-continuant changes. This suggests, first, that the information that infants use at 5 months to process and represent words is not yet in a phonetic format, as they seem to be more sensitive to acoustic cues or broadly defined phonetic features (stop vs continuant, that are distinguished in terms of their rms energy profiles) than to detailed phonetic features (such as place, aperture, etc.). Second, the English and French data together also suggest that 5-month-olds use different sets of acoustic cues depending on the language they are learning. It must be noted that we reanalysed the French data to examine whether the 30 consonant change test infants would also be differently sensitive to consonant changes when their name starts with a stop (N = 9) rather than a continuant (N = 21). Contrary to the English findings, the interaction between consonant type (stop vs continuant) and group (test vs control) was not significant (*F*(1, 56) < 1), indicating that French-learning infants were not particularly sensitive to changes when their name start with stop consonant. Similarly, they were equally likely to (not) discriminate a stop-to-stop change (N=7) or a continuant-to-continuant change (N=17; *F*(1, 22) < 1). As pointed out before, disyllabic French names have final lengthening, as opposed to having a trochaic pattern like most English disyllabic words: a word-initial short amplitude-rise-time stop would therefore be more perceptually salient in English than French, perhaps explaining the effect found in English infants and not in French.

As suggested by a Reviewer, our results could be partially accounted for by English and French differing on glottalisation in vowel-initial words. This is a frequent phenomenon in English (Dilley, Shattuck-Hufnagel & Ostendorf, 1996), but it occurs to a lesser extent in French (Fougeron, 2001; Malécot, 1975). We evaluated the occurrence of glottal stops in vowel-initial names, CP and MP, with the visualisation of spectrograms (5 randomly selected tokens for each type of vowel, adding up to 45 tokens in English and 40 in French): 10 tokens in English contained a clearly glottalised initial vowel (22 %), and 9 other contained some glottalisation (20 %). In French, we measured fewer glottalised vowels (5, 13 %), with some glottalisation in 14 tokens (35%). Therefore, names in both studies did not differ greatly in that respect, although the tendency was as expected from the literature, with clearer glottalisation in English than French. Beyond these two studies, it could be that exposure to more frequent glottalisation in English than French prevented the extraction of a fine-grained representation of the following vowels in English-learning infants.

The remaining possibility that British English-learning infants might be slightly late in retrieving segmental information as compared to those learning French (or even other languages) finds some echo in the comparison between vocabulary learning curves as measured with the adaptations of the MacArthur Communicative Development Inventories (Fenson et al., 1994) in 13 languages. Bleses et al. (2008) report that the two slowest curves of word comprehension between the ages of 8 and 15 months are for British English and Danish, the latter being a language in which the vowel system is particularly complex. The word comprehension curve in French (French CDI: Hilaire et al., 2001) is comparable to that observed in the majority of languages in that study. American English infants, however, perform better than their British counterparts, a finding which could either be related, according to Bleses et al. (2008), to the presence of rhoticity in American English which provides an additional cue to segmentation, or to the tendency for American speakers to produce secondary stress where British speakers produce unstressed syllables. Note that in our experiments, infants are brought up in a rhotic British English environment, although rhoticity did not impact phoneme manipulation. Whatever the explanation might be for British English infants to have been reported as knowing less words than American English infants (see Hamilton et al., 2000), the possibility that British English learners are genuinely late as compared to many other languages must not be discarded, and would relate to our finding. A few weeks or months might be necessary to catch up with French-learning infants, in which case one would expect to observe the same 'vowel bias' stage in young British English infants as was found in French-learning infants by Bouchon et al. (2015), before developing a sensitivity for consonants. Note however that even in older British English children, the emergence of the consonant bias as seen in French (e.g., Nazzi, 2005; Poltrock & Nazzi, 2015) or Italian (Hochmann, Benavides-Varela, Nespor & Mehler, 2011) is not found until the age of 30 months (Nazzi et al., 2009), and is preceded by a stage during which infants are equally sensitive to consonants and vowels (e.g., Floccia et al., 2014; Mani & Plunkett, 2007). So even if, at 5 months, British English infants are delayed in the processing of phonetic information in vowels as compared to French learners, their learning path may rely on different cues, as can be concluded from the different pathways infants take regarding the development of the consonant bias.

Rather than British English infants being late as compared to French ones, another possibility is that these two populations of infants enter lexical processing with different tools, because of the fundamentally different properties of their two native languages. Faced with a language like French with phrase-final lengthening, a syllable-timed rhythm, and mostly steady-state vowels, infants might focus on vowels in continuity with their initial bias (e.g., Benavides-Varela et al., 2012; Bertoncini et al., 1988). From these vowels, they start to extract acoustic information that will later coalesce into phonetic categories, with the help of a growing lexicon (e.g., Feldman, Myers, White, Griffiths & Morgan, 2013). With the refinement of their temporal resolution abilities (Werner, 1996) and lexicon increase (or their protolexicon of word forms), they will become more and more able to process differences between consonants, ultimately developing a consonant bias as early as 11 months (e.g., Poltrock & Nazzi, 2015). In contrast, British English-learning infants, exposed to a language with variable lexical stress, reduced vowels and frequent diphthongisation, have to work with a different weighting of acoustic cues: perhaps they become particularly sensitive to energy information (as was seen in Exp1 with names starting with short amplitude-rise-time stop consonants over larger amplitude-rise-time continuants), as one of the acoustic correlates of stress (Fear et al., 1995). This sensitivity would apply not necessarily to the highly variable class of vowels but would generalise to all kinds of phonemes, which would explain why British English-learning infants are later found to be equally sensitive to consonants and vowels (e.g., Floccia et al., 2014; Mani & Plunkett, 2007) before developing a consonant bias by the age of 30 months (Nazzi et al., 2009). In this approach, the format of representations or the processing of familiar words at the age of 5 months would not only be acoustically based – as opposed to phonetically specified - but also highly language-specific. Indeed, the current results compared to those of Bouchon et al. (2015) suggest that the acoustic characteristics of the native language shape the focus of infants' perceptual lens (see also Floccia et al., 2014; Højen & Nazzi, 2016).

How would infants' language-specific perceptual sensitivity for acoustic cues converge towards an adult-like phonetic and phonemic mode of perception? One possible way to enter into a phonemic mode of perception would be the acquisition of a lexicon (or at least a proto-lexicon). Indeed, the counter-intuitive idea that learning of phonemic categories emerges from the encoding of early word forms – as opposed to the more traditional, reverse, view (e.g., Lilienfeld et al., 2014; Martin & Fabes, 2008) - has recently abounded in the literature (e.g., Feldman et al., 2013a; Swingley, 2009). In the influential PRIMIR model for word learning and speech perception development, Werker and Curtin (2005) proposed for example that "as the vocabulary expands and more words with overlapping features are added, higher order regularities emerge from the multidimensional clusters. These higher order regularities gradually coalesce into a system of contrastive phonemes [..]" (p. 217). Feldman et al. (2013b) showed that 8-month-olds, just like adults, would use word-level information to process vowel contrasts, suggesting a strong role of top-down lexical information in the building of phonemic and phonetic categories. This learning process was further illustrated in a non-parametric Bayesian model (Feldman et al., 2013a), showing that phonetic (vocalic) ambiguity could be solved by using information about which sounds occur together in words (see also Elsner, Goldwater & Eisenstein, 2012). Moreover, according to Martin, Peperkamp and Dupoux (2013), it might not be necessary to have computed a lexicon per se to derive phonological categories, as infants could use top-down information found in a proto-lexicon, that is, a repertoire of all most frequent n-grams in the input (Fourtassi & Dupoux, 2014; Ngon et al., 2013; Poltrock & Nazzi, 2015). In favour of this hypothesis and regarding the current finding, a study by Mani and Plunkett (2010) shows that 12-month-old British English infants can detect various types of vowel changes in familiar words, irrespective of their acoustic characteristics. Interestingly, infants with larger vocabulary were more likely to detect these mispronunciations than infants with smaller vocabulary, suggesting that attunement of sensitivity to phonetic changes is partially driven by lexical growth between 5 (the current study) and 12 months (Mani & Plunkett, 2010). Moreover, a few studies have recently shown that learning words increases infants' sensitivity to phonetic (Yeung & Werker, 2009) or prosodic (Yeung & Nazzi, 2014) differences around 9-10 months of age.

In conclusion, we have shown that by the age of 5 months, British English learning infants represent and process early words with a different perceptual lens than their French-learning peers (Bouchon et al., 2015). These results call for further cross-linguistic investigations of the nature of the information that infants encode or process when presented with early familiar words, at the onset of lexical acquisition. The specificity of word representations and the lexical processing biases reported in older infants might result from a combination of language-specific acoustic biases reflecting the phonology of the native language, and a lexically-driven learning process that would originate with the language-specific format of the very first words being stored in the (proto)lexicon.

References

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*, *52*, 1135-1145.

Benavides-Varela, S., Hochmann, J. R., Macagno, F., Nespor, M., & Mehler, J. (2012). Newborn's brain activity signals the origin of word memories. *Proceedings of the National Academy of Sciences*, *109*(44), 17908-17913.

Bergelson, E., & Swingley, D. (2012). At 6–9 months, human infants know the meanings of many common nouns. *Proceedings of the National Academy of Sciences*, *109*(9), 3253-3258.

Bertoncini, J., Bijeljac-Babic, R., Jusczyk, P. W., Kennedy, L. J., & Mehler, J. (1988). An investigation of young infants' perceptual representations of speech sounds. *Journal of Experimental Psychology: General*, *117*(1), 21.

Bijeljac-Babic, R., Serres, J., Höhle, B., & Nazzi, T. (2012). Effect of bilingualism on lexical stress pattern discrimination in French-learning infants. *PLoS One*, *7*(2), e30843.

Bleses, D., Vach, W., Slott, M., Wehberg, S., Thomsen, P., Madsen, T. O., & Basbøll, H. (2008). Early vocabulary development in Danish and other languages: A CDI-based comparison. *Journal of child language*, *35*(03), 619-650.

Bouchon, C., Floccia, C., Fux, T., Adda ‐ Decker, M., & Nazzi, T. (2015). Call me Alix, not Elix: Vowels are more important than consonants in own-name recognition at 5 months. *Developmental Science, 18*(4), 587-598. DOI: 10.1111/desc.12242

Bull, D., Eilers, R. E., & Oller, D. K. (1985). Infants' discrimination of intensity variation in multisyllabic stimuli. *Journal of Acoustical Society of America*, *76*, 13-17.

Burnham, D., Kitamura, C., & Vollmer-Conna, U. (2002). What's new, Pussycat? On talking to babies and animals. *Science*, *296*, 1435. doi:10.1126/science.1069587

Christophe, A., Dupoux, E., Bertoncini, J., & Mehler, J. (1994). Do infants perceive word boundaries? An empirical study of the bootstrapping of lexical acquisition. *The Journal of the Acoustical Society of America*, *95*(3), 1570-1580.

Curtin, S., Fennell, C., & Escudero, P. (2009). Weighting of vowel cues explains patterns of word–object associative learning. *Developmental Science*, *12*(5), 725-731.

Davis, S. B., & Mermelstein, P. (1980). Comparison of parametric representations for monosyllabic word recognition in continuously spoken sentences. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, *28*(*4*), 357-366.

Delattre, P. (1964). Comparing the vocalic features of English, German, Spanish and French. *IRAL-International Review of Applied Linguistics in Language Teaching*, *2*(1), 71-98.

Delattre, P. (1966). A comparison of syllable length conditioning among languages. *IRAL-International Review of Applied Linguistics in Language Teaching*, *4*(1-4), 183-198.

Delattre, P. (1969). An acoustic and articulatory study of vowel reduction in four languages. *International Review of Applied Linguistics*, *12*, 295-325.

Dell, F. (1984). L'accentuation dans les phrases en français. In *Forme sonore du langage : Structure des représentations en phonologie*, F. Dell, D. J. Hirst & J. R. Vergnaud (Eds.), 65-122. Paris: Hermann.

Delle Luche, C., Poltrock, S., Goslin, J., New, B., Floccia, C., & Nazzi, T. (2014). Differential processing of consonants and vowels in the auditory modality: A cross-linguistic study. *Journal of Memory and Language*, *72*, 1-15.

Dilley, L., Shattuck-Hufnagel, S., & Ostendorf, M. (1996). Glottalization of word-initial vowels as a function of prosodic structure. *Journal of Phonetics*, *24*, 423-444.

Dodane, C., & Al-Tamimi, J. (2007). *An acoustic comparison of vowel systems in adult-directed-speech and child-directed speech: Evidence from French, English & Japanese*. Paper presented at the ICPhS XVI, Saarbrücken.

Eilers, R. E., Bull, D. H., Oller, D. K., & Lewis, D. C. (1983). The discrimination of vowel duration by infants. *Journal of Acoustical Society of America*, *75*(*4*), 1213-1218.

Eimas, P. D. (1975). Auditory and phonetic coding of the cues for speech: Discrimination of the [r-l] distinction by young infants. *Perception & Psychophysics*, *18*(*5*), 341-347.

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

Elsner, M., Goldwater, S., & Eisenstein, J. (2012, July). Bootstrapping a unified model of lexical and phonetic acquisition. In *Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics: Long Papers-Volume 1* (pp. 184-193). Association for Computational Linguistics.

Englund, K. T. (2005). Voice onset time in infant directed speech over the first six months. *First Language*, *25*(*2*), 219-234.

Fear, B. D., Cutler, A., & Butterfield, S. (1995). The strong/weak syllable distinction in English. *The Journal of the Acoustical Society of America*, *97*(3), 1893-1904.

Feldman, N. H., Griffiths, T. L., Goldwater, S., & Morgan, J. L. (2013a). A role for the developing lexicon in phonetic category acquisition. *Psychological Review*, *120*(4), 751.

Feldman, N. H., Myers, E. B., White, K. S., Griffiths, T. L., & Morgan, J. L. (2013b). Word-level information influences phonetic learning in adults and infants. *Cognition*, *127*(3), 427-438.

Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D. J., Pethick, S. J., & Stiles, J. (1994). Variability in early communicative development. *Monographs of the Society for Research in Child Development*, i-185.

Floccia, C., Nazzi, T., Delle Luche, C., Poltrock, S., & Goslin, J. (2014). English-learning one-to two-year-olds do not show a consonant bias in word learning. *Journal of Child Language*, *41*(05), 1085-1114.

Fougeron, C. (2001). Articulatory properties of initial segments in several prosodic constituents in French. *Journal of Phonetics*, *29*, 109-135. doi:10.006/jpho.2000.0114

Fourtassi, A., & Dupoux, E. (2014). A rudimentary lexicon and semantics help bootstrap phoneme acquisition. *CoNLL-2014*, 191.

Giegrich, H. J. (1992). *English phonology. An introduction*. Cambridge: Cambridge University Press.

Hallé, P. A., & de Boysson-Bardies, B. (1996). The format of representation of recognized words in infants' early receptive lexicon. *Infant Behavior and Development*, *19*(4), 463-481.

Hamilton, A., Plunkett, K., & Schafer, G. (2000). Infant vocabulary development assessed with a British communicative development inventory. *Journal of Child Language*, *27*(03), 689-705.

Havy, M., & Nazzi, T. (2009). Better processing of consonantal over vocalic information in word learning at 16 months of age. *Infancy*, *14*(4), 439-456.

Hilaire, G., Kern, S., Viguié, A., Dudognon, P., Langue, J., & Romieu, J. (2001). Le développement communicatif des enfants français de 8 à 30 mois. *Le Pédiatre*, XXXVI(182), 7–13.

Hirst, D., Di Cristo, A., & Nishinuma, Y. (2001) Prosodic parameters of French: A cross-language approach. In *Contrastive Studies of Japanese and Other Languages Series*, National Institute for Japanese Language, Tokyo, 7-20.

Hochmann, J. R., Benavides ‐ Varela, S., Nespor, M., & Mehler, J. (2011). Consonants and vowels: Different roles in early language acquisition. *Developmental Science*, *14*(6), 1445-1458.

Hochmann, J. R., & Papeo, L. (2014). The invariance problem in infancy: A pupillometry study. *Psychological Science*, *25*(11), 2038-2046.

Höhle, B., Bijeljac-Babic, R., Herold, B., Weissenborn, J., & Nazzi, T. (2009). Language specific prosodic preferences during the first half year of life: Evidence from German and French infants. *Infant Behavior and Development*, *32*(3), 262-274.

Højen, A., & Nazzi, T. (2016). Vowel bias in Danish word-learning: Processing biases are language-specific. *Developmental Science, 19*(1)*,* 41-49.

Houston, D. M., & Jusczyk, P. W. (2000). The role of talker-specific information in word segmentation by infants. *Journal of Experimental Psychology: Human Perception and Performance*, *26*(5), 1570.

Iverson P., Kuhl P. K., Akahane-Yamada R., Diesch E., Tohkura Y. I., Kettermann A., et al. (2003). A perceptual interference account of acquisition difficulties for non-native phonemes. *Cognition, 87*, B47–B57.

Jusczyk, P., Cutler, A., & Redanz, N. J. (1993). Infants' preference for the predominant stress patterns of English words. *Child Development*, *64*, 675-687.

Jusczyk, P. W., Pisoni, D. B., & Mullennix, J. (1992). Some consequences of stimulus variability on speech processing by 2-month-old infants. *Cognition*, *43*(3), 253-291.

Jusczyk, P., Pisoni, D. B., Reed, M. A., Fernald, A., & Myers, M. (1983). Infants' discrimination of the duration of a rapid spectrum change in nonspeech signals. *Science*, *222*(*4620*), 175-177. doi:10.1126/science.6623067

Kuhl, P. K. (1991). Human adults and human infants show a “perceptual magnet effect” for the prototypes of speech categories, monkeys do not. *Perception & Psychophysics*, *50*(2), 93-107.

Kuhl, P. K., Andruski, J. E., Chistovich, I. A., Chistovich, L. A., Kozhevnikova, E. V., Ryskina, V. L., et al. (1997). Cross-language analysis of phonetic units in language addressed to infants. *Science*, *277*, 684-686. doi:10.1126/science.277.5326.684

Kuhl, P. K., Conboy, B. T., Coffey-Corina, S., Padden, D., Rivera-Gaxiola, M., & Nelson, T. (2008). Phonetic learning as a pathway to language: new data and native language magnet theory expanded (NLM-e). *Philosophical Transactions of the Royal Society*, *363*, 979-1000. doi:10.1098/rstb.2007.2154

Lahiri, A., & Reetz, H. (2010). Distinctive features: Phonological underspecification in representation and processing. *Journal of Phonetics*, *38*(1), 44-59.

Lilienfeld, S. O., Lynn, S. J., Namy, L. L., Woolf, N. J., Jamieson, G., Marks, A., & Slaughter, V. (2014). *Psychology: From Inquiry to Understanding*. Pearson Higher Education.

Mack, M., & Blumstein, S. E. (1983). Further evidence of acoustic invariance in speech production: the stop–glide contrast. *The Journal of the Acoustical Society of America*, *73*(5), 1739-1750.

Malécot, A. (1975). The glottal stop in French. *Phonetica*, *31*(1), 51-63.

Mandel, D. R., Jusczyk, P. W., & Pisoni, D. B. (1995). Infants’recognition of the sound patterns of their own names. *Psychological Science*, *6*(5), 314.

Mani, N., & Plunkett, K. (2007). Phonological specificity of vowels and consonants in early lexical representations. *Journal of Memory and Language*, *57*(2), 252-272.

Mani, N., & Plunkett, K. (2008). Fourteen-month-olds pay attention to vowels in novel words. *Developmental Science*, *11*(1), 53-59.

Mani, N., & Plunkett, K. (2010). Twelve-month-olds know their cups from their keps and tups. *Infancy*, *15*(5), 445-470.

Martin, A., Peperkamp, S., & Dupoux, E. (2013). Learning phonemes with a proto-lexicon. *Cognitive Science*, *37*(1), 103-124.

Martin, C. L., & Fabes, R. (2008). *Discovering Child Development*. Cengage Learning.

Narayan, C. R., Werker, J. F., & Beddor, P. S. (2010). The interaction between acoustic salience and language experience in developmental speech perception: evidence from nasal place discrimination. *Developmental Science*, *13*, 407-420.

Nazzi, T. (2005). Use of phonetic specificity during the acquisition of new words: Differences between consonants and vowels. *Cognition*, *98*(1), 13-30.

Nazzi, T., & Bertoncini, J. (2003). Before and after the vocabulary spurt: two modes of word acquisition? *Developmental Science*, *6*(2), 136-142.

Nazzi, T., Floccia, C., & Bertoncini, J. (1998). Discrimination of pitch contours by neonates. *Infant Behavior & Development*, *21*(*4*), 779-784.

Nazzi, T., Floccia, C., Moquet, B., & Butler, J. (2009). Bias for consonantal information over vocalic information in 30-month-olds: Cross-linguistic evidence from French and English. *Journal of Experimental Child Psychology*, *102*(4), 522-537.

Nespor, M., Peña, M., & Mehler, J. (2003). On the different roles of vowels and consonants in speech processing and language acquisition. *Lingue e Linguaggio*, *2*(2), 203-230.

Ngon, C., Martin, A., Dupoux, E., Cabrol, D., Dutat, M., & Peperkamp, S. (2013). (Non) words, (non) words, (non) words: Evidence for a protolexicon during the first year of life. *Developmental Science*, *16*(1), 24-34.

Nittrouer, S., Lowenstein, J. H., & Tarr, E. (2013). Amplitude rise time does not cue the/bɑ/–/wɑ/contrast for adults or children. *Journal of Speech, Language, and Hearing Research*, *56*(2), 427-440.

Nittrouer, S., & Studdert-Kennedy, M. (1986). The stop–glide distinction: Acoustic analysis and perceptual effect of variation in syllable amplitude envelope for initial/b/and/w. *The Journal of the Acoustical Society of America*, *80*(4), 1026-1029.

Patel, I., & Rao, Y. S. (2011). Speech recognition rsing HMM with MFCC-An Analysis Using Frequency Specral Decomposition Technique. Signal & Image Processing.

Pater, J., Stager, C., & Werker, J. F. (2004). The perceptual acquisition of phonological contrasts. *Language*, *80*, 384-402.

Polka, L., Colantonio, C., & Sundara, M. (2001). A cross-language comparison of/d/–/ð/perception: evidence for a new developmental pattern. *The Journal of the Acoustical Society of America*, *109*(5), 2190-2201.

Poltrock, S., & Nazzi, T. (2015). Consonant/vowel asymmetry in early word form recognition. *Journal of Experimental Child Psychology*, *131*, 135-148.

Ramus, F., Nespor, M., & Mehler, J. (1999). Correlates of linguistic rhythm in the speech signal. *Cognition*, *73*(3), 265-292.

Skandera, P., & Burleigh, P. (2011). *A manual of English phonetics and phonology: Twelve lessons with an integrated course in phonetic transcription*. Gunter Narr Verlag.

Skoruppa, K., Pons, F., Christophe, A., Bosch, L., Dupoux, E., Sebastián-Gallés, N., Limissuri, R. A., & Peperkamp, S. (2009). Language-specific stress perception by 9-month-old French and Spanish infants. *Developmental Science*, *12*(6), 914-919.

Song, J. Y., Demuth, K., & Morgan, J. L. (2010). Effects of the acoustic properties of infant-directed speech on infant word recognition. *Journal of Acoustical Society of America*, *128*(*1*), 380-400. doi:10.1121/1.3419786

Stevens, K. N., & Blumstein, S. E. (1981). The search for invariant acoustic correlates of phonetic features. In P. D. Eimas & J. L. Miller (Eds.), *Perspectives on the study of speech* (pp. 1-38). Hillsdale, NJ: Lawrence Erlbaum.

Swingley, D. (2005). 11-month-olds' knowledge of how familiar words sound. *Developmental Science*, *8*(5), 432-443.

Swingley, D. (2009). Contributions of infant word learning to language development. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1536), 3617-3632.

Tincoff, R., & Jusczyk, P. W. (1999). Some beginnings of word comprehension in 6-month-olds. *Psychological Science*, *10*(2), 172-175.

Tincoff, R., & Jusczyk, P. W. (2012). Six-month-olds comprehend words that refer to parts of the body. *Infancy*, *17*(4), 432-444.

Vihman, M. M., Nakai, S., DePaolis, R. A., & Hallé, P. (2004). The role of accentual pattern in early lexical representation. *Journal of Memory and Language*, *50*(3), 336-353.

Werker, J. F., & Curtin, S. (2005). PRIMIR: A developmental framework of infant speech processing. *Language Learning and Development*, *1*(2), 197-234.

Werker, J. F., & Tees, R. C. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development*, *7*(1), 49-63.

Werner, L. A. (1996). The development of auditory behavior (or what the anatomists and physiologists have to explain). *Ear and Hearing*, *17*(5), 438-446.

White, L., Mattys, S. L., & Wiget, L. (2012). Language categorization by adults is based on sensitivity to durational cues, not rhythm class. *Journal of Memory and Language*, *66*, 665-679.

Yeung, H.H., & Nazzi, T. (2014). Object labeling influences infant phonetic learning and generalization. *Cognition, 132*(2), 151-163.

Yeung, H.H., & Werker, J.F. (2009). Learning words' sounds before learning how words sound: 9-month-old infants use distinct objects as cues to categorize speech information. *Cognition, 113*(2), 234-243.

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

Table 1. Participant information (age and gender), illustration of the 4 experimental conditions and stimuli examples.

Groups (all n = 30)	Stimuli (example)	Infant's name (example)	Age in days (SD)	Girls/boys
Consonant change condition				
Test	e.g., Molly vs Nolly	Molly	161 (14)	11/19
Control	e.g., Molly vs Nolly	Jessica	167 (15)	15/15
Vowel change condition				
Test	e.g., April vs Ipril	April	161 (14)	14/16
Control	e.g., April vs Ipril	Elsa	163 (12)	15/15

Table 2. List of names (CP) and their mispronunciation (MP) used in the Consonant change condition (left) and the Vowel change condition (right), with information about the feature change. For each name, the name of the associated control infant is provided. Changes involve one feature apart from /i:/-/e/ which involves aperture and tenseness (in bold), following the system proposed by Giegrich (1992; note that in case of feature redundancy, only the more informative one is retained, e.g., /æ/-/ʌ/).

Consonant change condition						Vowel change condition					
	CP	MP	Phonemes	Feature change	Control infant's name		CP	MP	Phonemes	Feature change	Control infant's name
1	Ben	Wen	/b/-/w/	Manner	Cian	1	Alfie	Elfie	/æ/-/e/	Aperture	Oliver
2	Joelly	Roelly	/dʒ/-/r/	Manner	Lucy	2	Anwynn	Enwynn	/æ/-/e/	Aperture	Hayley
3	Jorga	Shorga	/dʒ/-/ʒ/	Manner	Chloe	3	April	Ipril	/eɪ/-/aɪ/	Aperture	Elsa
4	Leo	Zeo	/l/-/z/	Manner	Seth	4	Eddie	Addie	/e/-/æ/	Aperture	Dexter
5	Logan	Zogan	/l/-/z/	Manner	Sophie	5	Elsie	Alsie	/e/-/æ/	Aperture	Amelia
6	Miya	Wiya	/m/-/w/	Manner	Reggie	6	Emma	Imma	/e/-/ɪ/	Aperture	Jolie
7	Miya	Wiya	/m/-/w/	Manner	Freddie	7	Esther	Asther	/e/-/æ/	Aperture	Anais
8	Monkey	Wonkey	/m/-/w/	Manner	Louie	8	Ethan	Eathan	**/i:/-/e/**	Aperture	Oliver
9	Nonny	Zonny	/n/-/z/	Manner	Annabelle	9	Ethan	Eathan	**/i:/-/e/**	Aperture	Luke
10	Rory	Jory	/r/-/dʒ/	Manner	Keely	10	Ethan	Eathan	**/i:/-/e/**	Aperture	Zachery
11	Caitlin	Taitlin	/k/-/t/	Place	James	11	Isaac	Eisaac	/aɪ/-/eɪ/	Aperture	Osian
12	Finley	Thinley	/f/-/θ/	Place	George	12	Isaac	Eisaac	/aɪ/-/eɪ/	Aperture	Felix
13	Finn	Thinn	/f/-/θ/	Place	Gray	13	Isaac	Eisaac	/aɪ/-/eɪ/	Aperture	Laila
14	Molly	Nolly	/m/-/n/	Place	Jessica	14	Isaac	Eisaac	/aɪ/-/eɪ/	Aperture	Aubrey
15	Poppy	Toppy	/p/-/t/	Place	Zachary	15	Isla	Eisla	/aɪ/-/eɪ/	Aperture	Emily
16	Reuben	Weuben	/r/-/w/	Place	Isabelle	16	Adam	Udam	/æ/-/ʌ/	Place	Lucas
17	Sam	Tham	/s/-/θ/	Place	Luka	17	Adam	Udam	/æ/-/ʌ/	Place	Willow
18	Sam	Sham	/s/-/ʃ/	Place	Noah	18	Alex	Ulex	/æ/-/ʌ/	Place	Henry
19	Thomas	Pomas	/t/-/p/	Place	Freya	19	Anya	Unya	/æ/-/ʌ/	Place	Isabelle
20	Warwick	Rarwick	/w/-/r/	Place	Beatrice	20	Archie	Orchie	/a:/-/ɔ:/	Place	Jacob
21	Daniel	Taniel	/d/-/t/	Voicing	Jack	21	Arthur	Orthur	/a:/-/ɔ:/	Place	George
22	Dory	Tory	/d/-/t/	Voicing	Florrie-May	22	Edmund	Udmund	/e/-/ʌ/	Place	Olivia
23	Finlay	Vinlay	/f/-/v/	Voicing	Theo	23	Edward	Udward	/e/-/ʌ/	Place	Iona
24	Finlay	Vinlay	/f/-/v/	Voicing	Phoebe	24	Effie	Uffie	/e/-/ʌ/	Place	Sonny
25	Jack	Chack	/dʒ/-/tʃ/	Voicing	Fearne	25	Ella	Ulla	/e/-/ʌ/	Place	Isabel
26	Jacob	Chacob	/dʒ/-/tʃ/	Voicing	Roberta	26	Erin	Urin	/e/-/ʌ/	Place	Isaac
27	Jacob	Chacob	/dʒ/-/tʃ/	Voicing	Kiera	27	Erin	Urin	/e/-/ʌ/	Place	Oscar
28	Kalli	Galli	/k/-/g/	Voicing	Hugo	28	Esme	Usme	/e/-/ʌ/	Place	Amaya
29	Pops	Bops	/p/-/b/	Voicing	Freya	29	Esme	Usme	/e/-/ʌ/	Place	Theodore
30	Sophie	Zophie	/s/-/z/	Voicing	Harrison	30	Oryn	Aryn	/ɒ/-/æ/	Place	Ebonie

Table 3. Summary of characteristics of stimuli used in the current experiment (English) and in the Bouchon et al. (2015) study (French). 'Diff.' corresponds to the mean difference between, for example, the duration of the consonant CPs minus that of the consonant MPs. Standard errors are in brackets.

		English				French			
		Duration	Intensity	Pitch	MFCC	Duration	Intensity	Pitch	MFCC
Consonant	CP	73.4 (6.4)	61.7 (1.2)			68.3 (5.3)	72.0 (1.3)		
	MP	89.7 (8.2)	61.6 (1.4)			78.1 (5.9)	70.9 (1.6)		
	Diff.	-16.4* (6.9)	0.10 (1.3)		1.18 (.03)	-9.8 (6.1)	1.2 (2.1)		1.54 (.06)
Vowel	CP	114.7 (8.8)	78.6 (0.3)	261.8 (9.4)		105.4 (3.2)	78.6 (0.3)		
	MP	111.0 (9.3)	78.7 (0.3)	263.5 (8.8)		108.3 (3.6)	79.1 (0.5)		
	Diff.	3.6 (3.8)	-.09 (0.2)	-1.7 (5.8)	1.25 (.05)	-2.8 (2.4)	-0.5 (0.6)		1.36 (.03)

Table 4. Individual results for the 60 children in the consonant change group (LTdiff test and LTdiff cont), with the associated pair of CP/MP and the corresponding Diff.intensity value. The feature change associated with the CP/MP pair is also provided, together with the phoneme category of the CP and the MP (stop or continuant).

	Diff.intensity	Ltdiff test	Ltdiff cont	Feature change	Phoneme category
Jorga-Shorga	-14.41	4.14	-3.64	manner	stop-cont
Joelly-Roelly	-13.22	3.79	0.96	manner	stop-cont
Ben-Wen	-12.88	0.62	0.93	manner	stop-cont
Jack-Chack	-3.52	2.17	-2.46	voice	**stop-stop**
Warwick-Rarwick	-3.19	-0.19	1.31	place	*cont-cont*
Poppy-Toppy	-3.16	-2.66	-0.08	place	**stop-stop**
Jacob-Chacob	-2.97	0.31	0.69	voice	**stop-stop**
Jacob-Chacob	-2.97	2.63	-3.77	voice	**stop-stop**
Daniel-Taniel	-2.48	1.60	-1.12	voice	**stop-stop**
Sam-Sham	-2.15	-0.09	-0.88	place	*cont-cont*
Caitlin-Taitilin	-1.43	3.92	-2.69	place	**stop-stop**
Thomas-Pomas	-1.19	-2.98	-0.35	place	**stop-stop**
Reuben-Weuben	-1.09	0.11	-0.16	place	*cont-cont*
Finlay-Vinlay	-1.04	0.86	1.38	voice	*cont-cont*
Finlay-Vinlay	-1.04	-1.45	-0.66	voice	*cont-cont*
Miya-Wiya	-0.74	5.57	1.35	manner	*cont-cont*
Miya-Wiya	-0.74	-1.24	1.94	manner	*cont-cont*
Finn-Thinn	0.66	-5.62	-0.42	place	*cont-cont*
Kalli-Galli	1.04	3.11	1.44	voice	**stop-stop**
Monkey-Wonkey	1.22	4.98	1.02	manner	*cont-cont*
Sophie-Zophie	1.32	-0.13	3.36	voice	*cont-cont*
Dory-Tory	2.03	-1.30	-0.97	voice	**stop-stop**
Molly-Nolly	2.29	1.57	-1.34	place	*cont-cont*
Sam-Tham	3.25	-3.41	0.40	place	*cont-cont*
Pops-Bops	4.37	0.00	1.00	voice	**stop-stop**
Finley-Thinley	5.49	-2.85	-2.81	place	*cont-cont*
Logan-Zogan	7.30	-4.06	2.11	manner	*cont-cont*
Leo-Zeo	8.87	0.98	1.77	manner	*cont-cont*
Nonny-Zonny	13.39	-4.82	-1.91	manner	*cont-cont*
Rory-Jory	15.05	-1.91	0.38	manner	cont-stop

Table 5: List of names and their foils used in Exp2

Name	Foil
Isla	Harry
Joel	Ewan
Cleo	Finlay
Ewan	Joel
Jacob	Reuben
Ella	Robyn
Freya	Callan
Hugo	Skyler
Charlie	Ella
Lily	Hugo
Holly	Alex
Boris	Freya
Daisy	Boris
Faith	Ben
Alfie	Robyn
Thomas	Nate
Xavier	Thomas
Sophie	Amber

Figure Captions

Figure 1. Mean looking times (s) towards the Correct Pronunciation of the names (filled bar) versus the Mispronunciation (hatched bar), in the Consonant change condition (left) and the Vowel change condition (right). Within each condition, results are broken down between the test group who heard their own name versus a mispronunciation (left) and the control group who heard a name and its mispronunciation but never encountered that name before (right). Brackets represent +/- 1 standard error.

Figure 2. *LT.Diff* as a function of *Diff.Intensity* for consonant changes on top, and for vowel changes below. LT.Diff = difference in looking times between the correct pronunciation of a name (CP) and its incorrect pronunciation (MP); positive values indicate a preference for the correct pronunciation. The same scale is used for both distributions to illustrate the wider range of values for consonants than for vowels. For the consonants, manner, place and voicing changes are depicted with a different marker (see key).

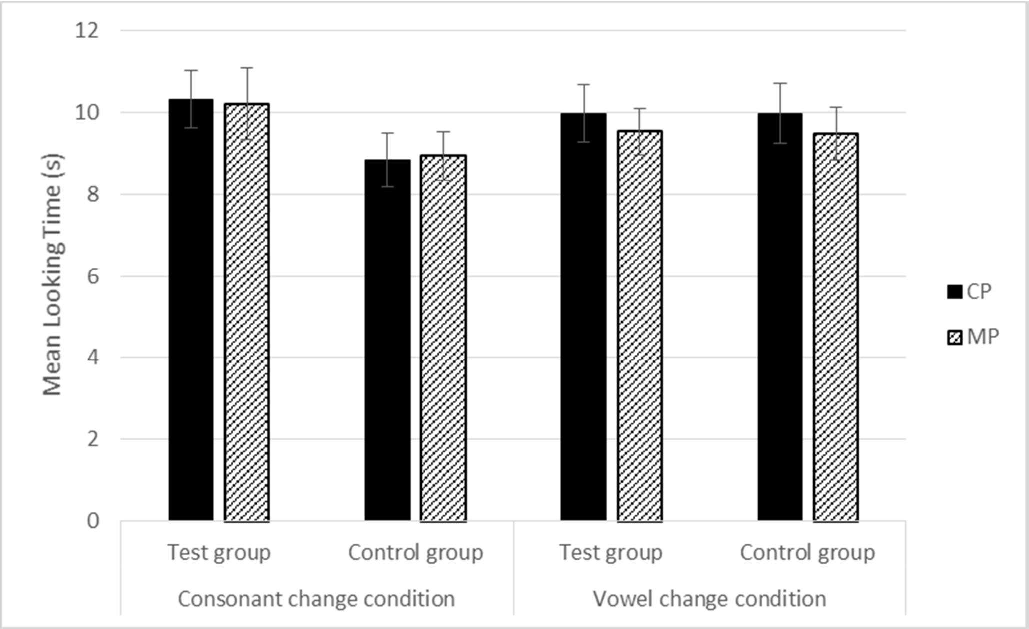
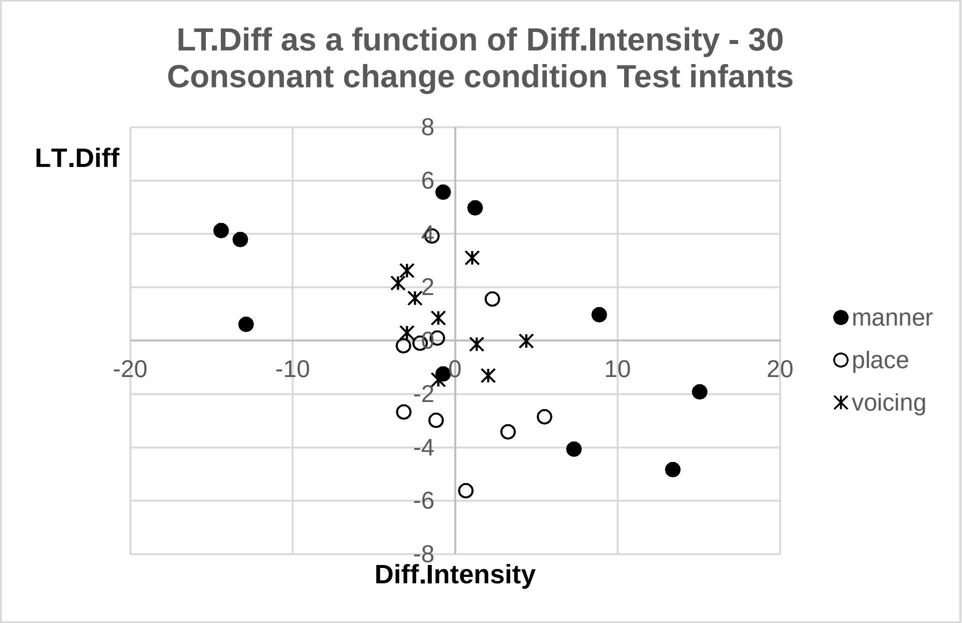

12
10
8
6
4
2
0
Mean Looking Time (s)
■ CP
▨ MP
Test group
Control group
Test group
Control group
Consonant change condition
Vowel change condition

Figure 1.


LT.Diff as a function of Diff.Intensity - 30 Consonant change condition Test infants
LT.Diff
Diff.Intensity
8
6
4
2
0
-2
-4
-6
-8
-20
-10
0
10
20
manner
place
voicing

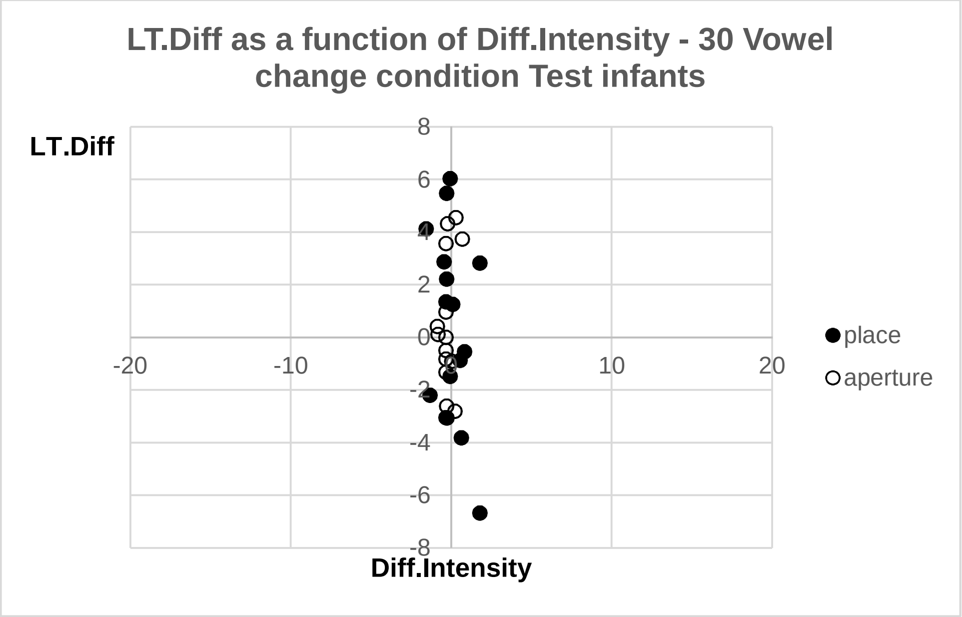

LT.Diff as a function of Diff.Intensity - 30 Vowel change condition Test infants
LT.Diff
Diff.Intensity
8
6
4
2
0
-2
-4
-6
-8
-20
-10
0
10
20
place
aperture

Figure 2.