The influence of bilingualism on the preference for the mouth region of dynamic faces

Alba Ayneto and Nuria Sebastian-Galles

Center for Brain and Cognition, Department of Technology, Universitat Pompeu Fabra, Roc Boronat 138, 08018, Barcelona, Spain.

Corresponding author: Alba Ayneto; alba.ayneto@upf.edu

Research Highlights

- Bilingual infants look longer at the mouth of dynamic faces without any linguistic content as compared to monolingual infants
- Language learning strategies are generalized to other dynamic faces such as emotional faces.
- The preference for the mouth region in bilingual infants opens a new line for further investigation about their sensitivity to informative cues conveyed by the eye region

Abstract

Bilingual infants show an extended period of looking at the mouth of talking faces, which provides them with additional articulatory cues that can be used to boost the challenging situation of learning two languages (Pons, Bosch, & Lewkowicz, 2015). However, the eye region also provides fundamental cues for emotion perception and recognition, as well as communication. Here, we explored whether the adaptations resulting from learning two languages are specific to linguistic contents or if they also influence the focus of attention when looking at dynamic faces. We recorded the eye gaze of bilingual and monolingual infants (8- and 12-month-olds) while watching videos of infants and adults portraying different emotional states (neutral, crying, and laughing). When looking at infant faces, bilinguals looked longer at the mouth region as compared to monolinguals regardless of the age. However, when presented with adult faces, 8 month-old bilingual infants looked longer at the mouth region and less at the eye region compared to 8 month-old monolingual infants, but no effect of language exposure was found at 12 months of age. These findings suggest that the bias to the mouth region in bilingual infants at 8 months of age can be generalized to other audiovisual dynamic faces that do not contain linguistic information. We discuss the potential implications of such bias in early social and communicative development.

Keywords: Bilingualism, attention, emotions, language development, infancy.

Introduction

Before the end of their first year of life, infants have already learned several important properties about their language, including a noticeable number of words. This is a remarkable achievement, considering the complexity and variability of the speech signal, among other things. Although the acoustic information is rich enough to allow for language learning, as successful language acquisition by congenitally blind infants shows, concurrent visual information in the form of articulatory movements provides redundant information. Infants may use this visual information to complement the auditory information, as adults do. Indeed, the acoustic information present in the signal in most cases is enough for competent adults to understand speech. However, in the case of difficult circumstances, such as noisy environments or when listening to non-native or heavily accented speakers, adults also make use of the redundant articulatory information provided by mouth movement to improve comprehension. In these circumstances, adults shift their attention from the eyes of the speaker to his/her mouth (Barenholtz, Mavica, & Lewkowicz, 2016; Navarra & Soto-Faraco, 2007; Sumby & Pollack, 1954).

Binding the auditory and visuo/articulatory information is already available to very young infants. By 5 months of age, infants are able to put into correspondence the articulation of speech sounds and the auditory stimuli (Burnham & Dodd, 2004; Kuhl & Meltzoff, 1982). Even more, there is evidence showing that in the first year of life, infants dynamically adapt their gaze patterns when listening to talking faces, presumably to take advantage of the articulatory information provided by the mouth. Lewkowicz & Hansen-Tift (2012) observed that when presented with audio-visual talking faces, 4-month-old monolingual infants spend more time looking at the eyes compared to looking at the mouth. However at 8 months of age, they shift their focus of attention from the eye to the mouth region when listening to both native and non-native languages. Such increase of attention starts declining towards the end of the first year of life for the native language, when infants' speech processing capacities have sufficiently improved. Tenenbaum, Shah, Sobel, Malle, & Morgan (2013) also provided converging evidence of an increase interest in the mouth region with age. According to some studies, the developmental changes of visual attention to the eyes and mouth reflect infants' ability to process speech. Indeed, the time spent looking at the mouth at 6 months and at 12 months predicts differences in language development later on (Tenenbaum, Sobel, Sheinkopf, Malle, & Morgan, 2014; Young, Merin, Rogers, & Ozonoff, 2009).

Converging evidence supporting the hypothesis that changes in gaze patterns in the first year of life reflects language development comes from studies exploring bilingual infants. Bilingual language learning poses a challenging learning situation because of the need to learn two separate systems and also because of the average reduced exposure to each language. (see for reviews Costa & Sebastián-Gallés, 2014; Werker, Byers-Heinlein, & Fennell, 2009). A series of studies suggests that bilingual infants may show an increase in attention to information presented in the mouth region when watching talking faces as compared to monolingual infants. Following the same procedure as Lewkowicz & Hansen-Tift (2012), Pons et al. (2015) investigated gaze patterns in monolingual and bilingual infants between 4 and 12 months of age. These authors reported an extended period of attention to the mouth in bilingual infants during their first year of life. As found by Lewkowicz & Hansen-Tift (2012), Pons et al. observed that monolinguals looked longer at the mouth than at the eyes at 8 months of age while watching talking faces that were speaking in their native language. They also found that at 12 months of age, monolingual infants started to shift their attention back to the eye region, looking equally long at the eye and the mouth regions. However, Pons et al. observed that bilingual infants looked longer at the mouth region both at 4 and at 12 months of age as compared to monolingual infants. The authors argued that bilingual infants would need to use the redundant articulatory cues present in the mouth region for a longer period of time to compensate for the more complex language-learning situation they face.

Pons et al.'s results fit well with previous findings investigating infants' capacity to discriminate languages based only on visual information. Indeed, Weikum et al. (2007) observed that French / English monolingual and bilingual infants at 4 and 6 months of age were equally able to visually discriminate these two languages. However, at 8 months of age, only bilingual infants were able to notice the differences. In a subsequent study, Sebastian-Galles, Albareda-Castellot, Weikum, & Werker (2012) showed that previous exposure to the languages being tested was not necessary for bilinguals' successful discrimination at 8 months of age. Sebastian-Galles et al. (2012) presented the same stimuli used in Weikum et al. (2007) to monolingual and bilingual Spanish/Catalan 8-month olds and also observed that bilinguals, but not monolinguals could discriminate between French and English. Although neither Weikum et al. nor Sebastian-Galles et al. collected gaze data, the developmental changes in the capacity to discriminate languages in these studies fit well with the differences between monolinguals and bilinguals concerning attention to the eye vs. attention to the mouth regions reported by Pons et al. (2015). The increased attention to articulatory information present in the mouth might help 8-month-old bilinguals to notice the differences between the two languages. The results of these studies open the research question about whether bilingual exposure can modulate infants' attention to faces in a more general way, regardless of the linguistic content of the stimuli.

Faces are one of the most common stimuli that infants encounter during the first years of life. They can convey emotional, social or linguistic information. Previous studies have shown that the scanning pattern of communicative faces varies as a function of age, the content, and the properties of the stimuli. A significant part of research exploring infants' scanning patterns of human faces in the first year of life has used static images of faces with emotional content. Taken together, these studies show that infants generally look longer at the eye region than at the mouth region (Hunnius, de Wit, Vrins, & von Hofsten, 2011; Peltola, Leppänen, Vogel-Farley, Hietanen, & Nelson, 2009). However, when presented with dynamic faces, 6 month-old infants tend to focus longer on the mouth region than when presented with static ones (Shic, Macari, & Chawarska, 2014). Other studies have reported a developmental shift from the eye to the mouth area when infants look at dynamic faces in the context of communicative interactions. Such is the case when Hunnius & Geuze (2004) presented infants from 6 to 26 weeks old with a silent video of their mother's face in which the mother was smiling, nodding, looking at the baby, and pronouncing a few words (see Frank, Vul, & Saxe (2012) for converging evidence with infants and children from 3 to 30 months). It is important to notice that although Hunnius & Geuze (2004) and Frank, Vul, & Saxe (2012) presented dynamic faces, the videos were not accompanied by speech sounds. As Hunnius & Geuze (2004) pointed out, it might be the case that the increased attention to the mouth as infants got older could reflect increased surprise to see a talking mouth without any sound. However, it could also be the case that the shift to the mouth area could be a generalization effect of the adaptations triggered by language acquisition processes as described above. Indeed, more recent studies comparing dynamic emotional and talking faces provide evidence of a modulation of the preference for the eye and mouth region depending on the emotional and linguistic content of faces.

Tenenbaum et al., (2013) presented 6-, 9-, and 12- month-old infants with dynamic faces of a speaking or smiling woman who could be either looking into the camera or to an object she was describing. The authors reported an overall decline of attention to the eye region and an increase to the mouth region with age in infants of 9 to 12 months of age. The shift to the mouth area was greater for speaking faces and the difference increased between 6 and 9 months (see also Shic et al., (2014) for converging results with 6-month-olds and Buchan, Paré, & Munhall, (2007) with adults). Summarizing, these results show a developmental path during the first year of life in infants' dynamic allocation of attention to maximize the language learning processes and to adjust to the properties of the presented faces. Such changes are likely to be the outcome of speech processing demands on the one hand, and the development of a more mature attentional system on the other.

In this context, it is interesting to analyse bilinguals' gaze behaviour, when presented with dynamic faces without any linguistic content. The results of Pons et al. showed an increase in attention to the mouth region in bilingual infants during the first year of life in this period. It might be the case that when presented with dynamic non-speaking faces, as when presented with speaking ones, bilingual infants will pay less attention to the eye region than monolinguals do, therefore showing a generalization of the gaze behaviour across faces. However, there is some evidence of increased attentional capacities in bilingual infants that might modulate their looking behaviour according to the contents of the face.

Several studies indicate that bilinguals may be more flexible in shifting their focus of attention when compared to monolinguals (see for reviews Bialystok, Craik, & Luk, 2012; Costa & Sebastián-Gallés, 2014). Using an anticipatory cueing paradigm, Kovács & Mehler (2009) showed that both monolingual and bilingual infants at 7 months of age were equally able to learn a new rule (to anticipate if a reward would appear in a right or left square after seeing a cue). However, when they had to learn a second rule that implied inhibiting the previously learned one, only bilingual infants succeeded. Poulin-Dubois, Blaye, Coutya, & Bialystok (2011) showed that at 24 months, bilingual infants outperformed monolingual infants in the Stroop task, a task involving conflict inhibition. In both studies, bilingual infants showed an increased flexibly to control their focus of attention. Taking this evidence into consideration, it could be the case that the pattern observed by Pons et al. (2015) would be restricted to speech stimuli and that bilingual infants can modulate their focus of attention depending on the nature of the situation (as monolinguals do at 12 months when presented with their native language). If this were the case then bilingual and monolingual infants would show equivalent patterns of visual attention when seeing emotional faces.

Here we investigate patterns of selective attention of monolingual and bilingual infants at 8 and 12 months of age when viewing non-linguistic emotional dynamic faces portraying different emotional states (happiness, distress, or neutral). In the first experiment we exposed infants to other infants' faces. We chose these stimuli because they are an ecological display that infants are exposed to from very early in their life, either in daycares or in playgrounds. Furthermore, it is unlikely that children will expect linguistic information to come from baby faces. Infants as young as 6-month-old have shown sensitivity to other infants' emotions (Geangu, Hauf, Bhardwaj, & Bentz, 2011) and they are able to detect, discriminate, and match the facial and vocal affective displays of other infants (Vaillant-Molina, Bahrick, & Flom, 2013). In the second experiment, infants were tested with a parallel set of stimuli with adult faces, allowing for more direct comparison with previous studies.

Experiment 1

Methods

The research reported in this manuscript has been conducted in accordance with the principles expressed in the Declaration of Helsinki and approved by the local ethical committee. All parents signed an informed consent for their infants to participate in this study.

Participants

Eighty-eight infants were retained for the analysis: Twenty-two monolingual 8-month-olds (11 boys, *M* = 233, *SD* = 6.78 days), twenty-two bilingual 8-month-olds (13 boys, *M* = 233, *SD* = 7.80 days), twenty-two monolingual 12-month-olds (11 boys, *M* = 360, *SD* = 8.55 days) and twenty-two bilingual 12-month-olds (11 boys, *M* = 356, *SD* = 6.69 days). Thirty-two additional infants were tested but not included in the sample due to: improper calibration (n = 9: lack of infants' cooperation or the eye tracker could not find infants' pupil), technical error (n = 2), experimental error (n = 2), fussiness or crying (n = 2) or less than 50% of data could be obtained in at least one of the three conditions in the analysis time window (n = 17; 8-month-olds: 6 monolinguals, 6 bilinguals; 12-month-olds: 4 monolinguals, 1 bilingual). The sample size was determined according to previous studies (Pons et al. 2015).

Participants were recruited by visiting maternity rooms at the Hospital Quirón and the Clínica Sagrada Família in Barcelona, Spain. All participants were healthy, full-term infants (> 37 GW) and exposed to Catalan or Spanish, or both. A questionnaire (adapted from Bosch & Sebastián-Gallés, 2001) was administered to determine infants' language background and familiarity. Bilingual infants were exposed to their main language (either Spanish or Catalan) up to 75% of the time (3 bilingual infants were exposed to a third language: two were exposed 5% of the time and 1 bilingual infant was exposed 10% of the time).

Stimuli

The stimuli were three 30-second long video recordings of three male infants displaying neutral, positive (laughing) or negative (crying) emotions, adapted from Geangu et al. (2011). We used these videos because they did not have any linguistic content, had high ecological value, and produced an emotional response. Videos were shortened with respect to the original ones (from 50 to 30 seconds) based on the previous study showing significant data loss during the second half of the stimuli presentation due to a decrease in visual attention. The stimuli were also edited to equalize the size of the faces. The two videos depicting the infant crying and the infant laughing were more dynamic and included highly salient acoustic stimuli (especially in the crying one). The infant in the positive emotion video (laughing) looked less at the camera and the infant in the negative emotion video (crying) did not look at the camera. The video depicting the neutral emotion was more static and quiet, and the infant was looking at the camera most of the time. Figure 1a shows representative screenshots of each video.

Apparatus and procedure

Infants were tested in a sound-attenuated room and seated in an infant seat ≈65 cm from the screen and monitored through a camera during the session. The caregiver sat behind the infant. The eye tracker was calibrated using the infant nine-point calibration from Tobii Studio 2.3.0 before each recording. Participants started either with the presentation of the neutral or the positive condition, followed by the other condition, with the order being counterbalanced. The negative condition was always presented last, following the procedure by Geangu et al. (2011), in order to avoid carry-over effects related to potential distress (see Geangu et al. 2011 for a more detailed description). Each video was presented once, preceded by a 14 second long attractor video of a moving rattle shape associated with a sound. Stimuli were presented using Tobii Studio 2.3.0 on a 24” screen and gaze was measured using a Tobii 60XL near infrared eye tracker, recording at a frequency of 60 Hz. The videos and the attractor were 720 x 580 pixels in size that were presented on a grey background. The total duration of the study was 2.2 minutes. An extra attractor and an extra video of a laughing baby (30 seconds) were always presented at the end of the study to improve the mood of the baby after seeing the video of the crying baby. This video was irrelevant to the study and never analyzed.

Data analysis

For the analysis, we defined three areas of interest (AOIs) with a rectangle/square shape for each video separately that covered the eye region, the mouth region, and the whole face. We computed the proportion of total looking time (PTLT) as done by Pons et al. (2015), by dividing the time infants spent looking in the AOI (eyes or mouth) by the total time looking at the face. The time window of the analysis was the first 8 seconds of the video. This time window was chosen according to previous studies analyzing gaze behavior of faces. Frank et al. (2012) and Tenenbaum et al. (2013) used videos of the same length and even shorter in length. The complete video was not analyzed due to an increase in movement of the infants in the video. The statistical analyses were performed following the same design implemented by Pons et al. (2015).

Results and discussion

A mixed repeated measure ANOVA on the proportion of time spent looking at each region was performed with four factors: Age (8m.o, 12 m.o) and Language Background (Monolingual, Bilingual) as between factors, and AOI (Eyes, Mouth) and Condition (Neutral, Cry, Laugh) as within factors. The ANOVA showed a main effect of Condition ($F(2,168) = 134.26$, $p < 0.001$, $\eta^2 = 0.62$: Neutral = 0.30 ± 0.11, Cry = 0.46 ± 0.05, Laugh = 0.43 ± 0.05) and a main effect of AOI ($F(1,84) = 36.42$, $p < 0.001$, $\eta^2 = 0.30$, Mouth = 0.53 ± 0.22, Eyes = 0.26 ± 0.22). The ANOVA also showed an interaction between Condition and AOI ($F(2,168) = 109.8$, $p < 0.001$, $\eta^2 = 0.57$). This interaction reflected that infants looked longer at the mouth region in the Cry ($t(87)=-10.89$, $p<0.001$) and in the Laugh condition ($t(87)= -6.26$, $p <0.001$), but not in the Neutral condition ($t(87)= 0.99$, $p>0.3$). Indeed the difference between the PTLT for the mouth and eyes was larger for the Cry (ΔPTLT mouth-eyes = 0.53) than for the Laugh condition (ΔPTLT mouth-eyes = 0.34) (see Table 1).

Critically, a significant interaction between AOI and Language Background ($F(1,84) = 5.18$, $p = 0.025$, $\eta^2 = 0.058$) was also found. Planned comparisons were run to characterize the PTLT in each AOI for each language group. The analyses showed that both language groups looked longer at the mouth compared to the eyes (Monolinguals: $t(43) = -2.51$, $p = 0.016$; Bilinguals: $t(43) = -6.42$, $p < 0.01$). The comparison of monolingual and bilingual infants showed that bilinguals looked less at the eye region ($t(86) = -2.42$, $p = 0.043$) and more at the mouth region ($t(86) = 2.048$, $p = 0.017$) (see Figure 2). No other effects or interactions reached significance.

The main result of this experiment is that bilingual infants looked more at the mouth region than bilingual infants, regardless of the type of stimuli. The data thus support the existence of a bias to the mouth region in this population, likely reflecting a generalization of the attentional bias that bilingual language learning induces. However, in contrast with previous research we did not find an effect of age. There is evidence showing that familiarity with stimuli modulates gaze patterns. Wheeler et al. (2011) observed a change in face scanning in infants between 6 and 10 months of age. In this study, infants were presented with faces from their own race or from a different race. The results showed that older infants increasingly looked longer at the eyes and less at the mouth than younger infants. The authors related the shift to experience with own-race faces. It might be that the relatively reduced experience with infant faces may have attracted infants' attention to the mouth area. In experiment 2 we presented infants with recordings of adult faces.

Experiment 2

Methods

Methods were the same as in the previous experiment except for the following.

Participants

44 infants were retained for the analysis: 12 monolingual 8-month-olds (3 boys, *M* =241, *SD* = 15.9 days), 12 bilingual 8-month-olds (7 boys, *M* = 247, *SD* = 15.21 days), 12 monolingual 12-month-olds (5 boys, *M* = 359, *SD* = 19.16 days) and 12 bilingual 12-month-olds (7 boys, *M* = 352, *SD* = 17.67 days). Six additional infants were tested but not included in the sample due to: improper calibration (n = 1: the eye tracker could not find the infants' pupil), fussiness (n = 1) or less than 50% of data could be obtained in at least one of the three conditions in the analysis time window (n

= 4; 8-month-olds: 1 monolingual, 1 bilingual; 12-month-olds: 2 bilinguals). None of the infants participated in the previous study.

Stimuli

The stimuli were three 8-second long audiovisual video recordings of three female actresses displaying neutral, positive (laughing) or negative (crying) emotions. The stimuli were edited to equalize the size of the faces. A large effort was made to ensure that the adult videos were equivalent to the infant ones. In the neutral video, the girl was moving her eye gaze, and producing the sound “hum!” while chewing a piece of chewing gum (in order to incorporate movement in the mouth area). In the laughing and crying conditions, the actresses were laughing and crying and producing emotional vocalizations.

The video depicting the neutral emotion was more static and quiet compared to the crying and laughing one, and the female was looking at the camera most of the time (to be as similar as possible to the infant video). Figure 1b shows representative screenshots of each video.

Apparatus and procedure

Infants were tested in a sound-attenuated room and seated on a caregiver’s lap, ≈65 cm from the screen and monitored through a camera during the session. Caregiver wore black glasses. The eye tracker was calibrated using the infant five-point calibration before each recording. Participants started either with the presentation of the neutral or the positive condition, followed by the other condition, the order was counterbalanced (according to experiment 1). Each video was presented once. Before each video started, an attractor (the same as in experiment 1) was presented for at least 7 seconds. When the 7 seconds were over, if the infant was looking at the screen, the next video started. If not, the attractor continued either until the infant fixated on the screen or until 14 seconds passed. This change was introduced because in experiment 1 some infants got bored of the attractor before it finished.

Stimuli were presented using Matlab 7.11 the Psychtoolbox and Tobii Analytics Software Devlopment Kit (Tobii Analytics SDK) on a 23" screen (1080 x 1920 pixels) and gaze was measured using a Tobii 300TXL near infrared eye tracker, recording at a frequency of 60 Hz. The videos and the attractor were 720 x 580 pixels in size that were presented on a grey background. The total duration of the study was between 1 and 1.5 minutes in length. An extra attractor and an extra video of a laughing character (a Teletubby) were always presented at the end of the study to improve the mood of the baby after seeing the video of the crying adult. As in the first experiment, this video was irrelevant to the study and never analyzed.

Data analysis

We followed the same procedure as in experiment 1. In this case we analyzed the whole video because it was 8 seconds long (the same time window as in experiment 1).

Results

A mixed repeated measure ANOVA on the proportion of time spent looking at each region was performed with four factors: Age (8m.o, 12 m.o) and Language Background (Monolingual, Bilingual) as between factors, and AOI (Eyes, Mouth) and Condition (Neutral, Cry, Laugh) as within factors. The ANOVA showed a main effect of Condition ($F(2,88) = 67.75$, $p < 0.001$, $\eta^2 = 0.60$: Neutral = 0.34 ± 0.09; Laugh = 0.46 ± 0.03; Cry = 0.42 ± 0.05) and a main effect of AOI ($F(1,44) = 7.616$, $p < 0.01$, $\eta^2 = 0.15$, Mouth = 0.34 ± 0.22, Eyes = 0.48 ± 0.20). The ANOVA also showed an interaction between Condition and AOI ($F(2,88) = 8.845$, $p < 0.001$, $\eta^2 = 0.167$). Infants looked longer to the eye region in the Cry ($t(47)=3.05$, $p=0.003$) and in the Neutral condition ($t(47)= 3.89$, $p <0.001$), but not in the Laugh condition ($t(47)= 0.14$, $p=0.88$).

A significant interaction between Age and AOI was found ($F(1,44)=6.70$ $p=0.013$, $\eta^2 = 0.13$). This interaction reflected that 12-month-old compared to 8-month-old infants looked longer at the mouth region ($t(46) = 2.37$, $p=0.022$) and less at the eye region ($t(46) =- 2.226$, $p=0.028$) (See Table 2).

Critically, a significant interaction between AOI, Age, and Language Background ($F(1,44) = 7.40$, $p = 0.009$, $\eta^2 = 0.14$) was also found. Planned comparisons were run separately for both age groups to compare the PTLT in each AOI for each language group. At 8 months of age, bilingual infants looked longer at the mouth region ($t(22) = 3.16$, $p=0.004$) and less at the eye region ($t(22) = -3.60$, $p=0.002$). However, at 12 months of age, monolinguals and bilinguals did not differ either in the time they looked at the mouth region ($t(22) = -0.74$ $p=0.489$) or at the time they looked at the eye region ($t(22) = 0.92$ $p=0.367$) (See Figure 3). No other effects or interactions reached significance.

The results show a different developmental path between monolinguals and bilinguals when looking at adult audiovisual dynamic faces. At 8 months of age, bilingual infants looked longer at the mouth region compared to monolingual infants in all three conditions. However at 12 months, both groups behaved in the same way. As discussed below, these data support the existence of a generalization of the attentional bias that bilingual language learning induces, and parallels the developmental path found by Pons et al (2015).

Discussion

In this research, we investigated whether bilingualism impacts the way infants pay attention to infant and adult non-linguistic faces. In two different studies we tested 8- and 12-month-old infants' gaze patterns when watching audiovisual dynamic faces without any communicative content. In experiment 1, the faces corresponded to infants and in experiment 2 the faces were of young adult women. The results of experiment 1 showed that 8- and 12 month-old bilingual infants, regardless of the age, looked longer at the mouth region and less at the eye region, compared to age-matched monolingual infants. In experiment 2, when presented with adult non-linguistic faces, a developmental pattern emerged. Bilingual infants at 8 months looked longer at the mouth region and less to the eye region compared to monolingual infants, but no effect of language exposure was found at 12 months of age. At this age, both monolingual and bilingual infants looked longer at the mouth and less at the eye region, compared to younger infants. There are two relevant results in our research deserving comment. The first one refers to the differences between experiments 1 and 2 and the second one is the developmental trajectory observed in experiment 2. Let us consider each one in turn.

The only methodological difference between experiments 1 and 2 was the stimuli used. Although we tried to match the videos in both studies as much as possible, we cannot exclude the possibility that the reported differences might be driven by some stimulus-specific features, such as, more saliency of the mouth region in the infant faces. The amount of movement or the percentage of the eye and mouth areas in both sets of stimuli may have been different enough to trigger some differences, in particular, the overall preference for the mouth region in experiment 1 when compared to experiment 2. However, it is likely that the use of infant faces vs. adult faces in experiment 1 may have been at the origin of the observed differences, as infants have less exposure to infant faces than to adult faces. Infants visiting our lab are used to playing in kindergartens and see other infants from very early on. However they may encounter them less often and interact differently with them as compared to adults (with whom they interact more often). The lack of experience with infant faces could trigger them to explore more the audiovisual redundant cues present in the mouth region. In fact, this could explain why infants are able to match vocal and facial emotional expressions coming from infant faces before they can do it with adult faces (Vaillant-Molina et al., 2013). Previous research has suggested that the uncertainty of an audiovisual signal triggers the use of audiovisual redundant cues both in infants (Lewkowicz & Hansen-Tift, 2012; Pons et al., 2015) and in adults (Barenholtz et al., 2016). The mouth area could therefore act as a perceptual magnet making it difficult for infants to disengage from it, also causing the lack of a developmental trajectory in shifting gaze when scanning infant faces between 8 and 12 months of age in experiment 1.

In contrast with experiment 1, the results of experiment 2 uncover differences in the developmental trajectories between monolinguals and bilinguals. As reviewed in the introduction, previous research has reported a change in the way infants scan faces in the second half of the first year of life (it is worth highlighting again that most of the reviewed research has been performed using adult faces in the stimuli). Several studies have shown a decrease in attention to the eye area accompanied by a more distributed scanning of faces, including an increase in attention to the mouth area between 6 and 12 months of age (Frank et al., 2012; Hunnius & Geuze, 2004b; Oakes & Ellis, 2013; Tenenbaum et al., 2013; Wheeler et al., 2011); although Lewkowicz & Hansen-Tift (2012) and Pons et al., (2015) found a recovery of attention to the eye area by 12 months of age).

In our study, bilinguals and monolinguals showed equivalent responses at 12 months of age. This shift has been attributed to increased experience with human faces (Oakes & Ellis, 2013). Given that our monolingual and bilingual infants were tested at the same age and live in culturally equivalent environments, it is unlikely that the differences we have observed with adult faces are due to increased experience in bilinguals. However, the increased interest for the mouth region has been linked as well, to language acquisition processes. As discussed, further exploration of the mouth region provides redundant information from articulatory gestures when learning speech. The interpretation of the results that we favor is consistent with the pattern described by Pons et al. (2015). As described in the introduction, these authors compared 4-, 8- and 12- month-old monolingual and bilingual infants when looking at talking adult faces. They observed that bilinguals looked longer at the mouth than monolinguals at 4 months of age, but that at 8 months of age both groups behaved in an equivalent way, that is, they looking longer at the mouth. At 12 months, both groups differed again, as monolinguals started to shift their gaze back to the eye area, looking equally as long at the mouth and the eyes, while bilinguals still looked longer at the mouth area. Our results are consistent with this pattern, although our results are shifted by four months. As in Pons et al. (2015), we observed a difference between monolinguals and bilinguals at a young age (4 months in their study and 8 months in ours) in that bilinguals look less at the eyes and more at the mouth than monolinguals. Four months later, the differences between the two groups disappear, due to the shift to the mouth area by monolingual infants. The temporal difference between Pons et al's and our study is consistent with studies showing that the linguistic content of the stimuli attracts the attention of the infants to the mouth region (Shic et al., 2014; Tenenbaum et al., 2013).

Summarizing, although the different properties of the stimuli between the studies have modulated gazing patterns, our results point out a different developmental path when looking at non-linguistic dynamic faces between bilingual and monolingual infants. A particular language experience such as bilingualism can trigger strategies that are generalized to other domains (outside the linguistic domain). Our results also show that previous experience with the stimuli modulates patterns of face scanning: infants look longer at the mouth region of unfamiliar stimuli.

One interesting question that our conclusions raise refers to the potential cost of such increased focus on the mouth for bilinguals. It might be the case that bilingual infants´ increased attention to the mouth area would hinder their sensitivity to informative cues conveyed by the eyes, including communicative as well as emotional information. Facial information, and in particular information in the eye region, provides fundamental cues for social interaction, emotion perception, and emotion recognition (Amso, Fitzgerald, Davidow, Gilhooly, & Tottenham, 2010; Vanderwert et al., 2014). Research in our laboratory points in the direction that bilingual toddlers are less sensitive to informative cues present in the eye region. Fort, Escrichs, Ayneto, & Sebastian-Galles (2015) presented 15 month-old monolingual and bilingual infants with videos of a female speaker who at the end of each sentence raised her eyebrows or protruded her lips. Both populations did not differ in the lip protrusion condition, but significant differences emerged in the raising eye brow condition. Monolingual infants were able to disengage their attention from the mouth region and direct their gaze to the eyes of the speaker before the end of the sentence, thus anticipating the speaker's movement. Bilingual infants only shifted their attention to the eyes once the sentence was over and the speaker had started to raise her eyebrows. These results would be compatible with the possibility of a temporary advantage in monolingual infants to process emotional and communicative information conveyed by the eyes. Current research in our laboratory is exploring this hypothesis.

Author Contributions

A. Ayneto (AA) and N. Sebastian-Galles (NSG) developed the concept for the study. Testing, data collection, and data analysis were performed by AA. The manuscript was written by AA and NSG. All authors approved the final version of the manuscript for submission.

Acknowledgments

We want to thank Elena Geangu for sending us the stimuli and for her useful suggestions, Xavier Mayoral and Silvia Blanch for the technical support, and Mathilde Fort and Loreto Nácar for their helpful comments. We also want to thank all the families and infants for their participation in the studies. This research was supported by grants from the European Community's Seventh Framework Programme (FP7/2007-2013): ERG grant agreement number 323961 (UNDER CONTROL); the Spanish Ministerio de Economía y Competitividad (PSI2015-66918-P) and the Catalan Government (SGR 2014–1210). N.S.G. received the prize ‘‘ICREA Acadèmia’’ for excellence in research, funded by the Generalitat de Catalunya.

References

Amso, D., Fitzgerald, M., Davidow, J., Gilhooly, T., & Tottenham, N. (2010). Visual exploration strategies and the development of infants' facial emotion discrimination. *Frontiers in Psychology, 1*. doi:10.3389/fpsyg.2010.00180

Barenholtz, E., Mavica, L., & Lewkowicz, D. J. (2016). Language familiarity modulates relative attention to the eyes and mouth of a talker. *Cognition*, *147*, 100–105. doi:http://dx.doi.org/10.1016/j.cognition.2015.11.013

Bialystok, E., Craik, F. I. M., & Luk, G. (2012). Bilingualism: Consequences for mind and brain. *Trends in Cognitive Sciences*, *16*, 240–249. doi:10.1016/j.tics.2012.03.001

Bosch, L., & Sebastián-Gallés, N. (2001). Evidence of early language discrimination abilities in infants from bilingual environments. *Infancy*, *2*, 29–49. doi:10.1207/S15327078IN0201_3

Buchan, J. N., Paré, M., & Munhall, K. G. (2007). Spatial statistics of gaze fixations during dynamic face processing. *Social Neuroscience, 2*, 1–13. doi:10.1080/17470910601043644

Burnham, D., & Dodd, B. (2004). Auditory-visual speech integration by prelinguistic infants: Perception of an emergent consonant in the McGurk effect. *Developmental Psychobiology*, *45*, 204–220. doi:10.1002/dev.20032

Costa, A., & Sebastián-Gallés, N. (2014). How does the bilingual experience sculpt the brain? *Nature Reviews. Neuroscience*, *15*, 336–45. doi:10.1038/nrn3709

Fort, M., Escrichs, A., Ayneto, A., & Sebastian-Galles, N. (2015). You can raise your eyebrows , I don’t mind : are monolingual and bilingual infants equally good at learning from the eyes region of a talking face ? In *1st Joint Conference on Facial Analysis, Animation and Audio-Visual Speech Processing (FAAVSP-2015),* 7-11.

Frank, M. C., Vul, E., & Saxe, R. (2012). Measuring the Development of Social Attention Using Free-Viewing. *Infancy*, *17,* 355–375. doi:10.1111/j.1532-7078.2011.00086.x

Geangu, E., Hauf, P., Bhardwaj, R., & Bentz, W. (2011). Infant pupil diameter changes in response to others’ positive and negative emotions. *PLoS ONE, 6*.

doi:10.1371/journal.pone.0027132

Hunnius, S., de Wit, T. C. J., Vrins, S., & von Hofsten, C. (2011). Facing threat: infants’ and adults' visual scanning of faces with neutral, happy, sad, angry, and fearful emotional expressions. *Cognition & Emotion, 25*, 193–205. doi:10.1080/15298861003771189

Hunnius, S., & Geuze, R. H. (2004). Developmental changes in visual scanning of dynamic faces and abstract stimuli in infants: A longitudinal study. *Infancy, 6*, 231–255. doi:10.1207/s15327078in0602_5

Kovács, A. M., & Mehler, J. (2009). Cognitive gains in 7-month-old bilingual infants. *Proceedings of the National Academy of Sciences of the United States of America*, *106*, 6556–6560. doi:10.1073/pnas.0811323106

Kuhl, P. K., & Meltzoff, A. N. (1982). The bimodal perception of speech in infancy. *Science*, *218*(4577), 1138–1141. doi:10.1126/science.7146899

Lewkowicz, D. D. J., & Hansen-Tift, A. M. (2012). Infants deploy selective attention to the mouth of a talking face when learning speech. *Proceedings of the National Academy of Sciences of the United States of America*, *109*, 1431–1436. doi:10.1073/pnas.1114783109

Navarra, J., & Soto-Faraco, S. (2007). Hearing lips in a second language: Visual articulatory information enables the perception of second language sounds. *Psychological Research*, *71*, 4–12. doi:10.1007/s00426-005-0031-5

Oakes, L. M., & Ellis, A. E. (2013). An Eye-Tracking Investigation of Developmental Changes in Infants’ Exploration of Upright and Inverted Human Faces. *Infancy*, *18*, 134–148. doi:10.1111/j.1532-7078.2011.00107.x

Peltola, M. J., Leppänen, J. M., Vogel-Farley, V. K., Hietanen, J. K., & Nelson, C. a. (2009). Fearful faces but not fearful eyes alone delay attention disengagement in 7-

month-old infants. *Emotion, 9,* 560–565. doi:10.1037/a0015806

Pons, F., Bosch, L., & Lewkowicz, D. J. (2015). Bilingualism Modulates Infants' Selective Attention to the Mouth of a Talking Face,. *Psychological Science, 26,* 490-498. doi:10.1177/0956797614568320

Poulin-Dubois, D., Blaye, A., Coutya, J., & Bialystok, E. (2011). The effects of bilingualism on toddlers' executive functioning. *Journal of Experimental Child Psychology, 108,* 567–579. doi:10.1016/j.jecp.2010.10.009

Sebastian-Galles, N., Albareda-Castellot, B., Weikum, W. M., & Werker, J. F. (2012). A Bilingual Advantage in Visual Language Discrimination in Infancy. *Psychological Science, 23,* 994-999. doi:10.1177/0956797612436817

Shic, F., Macari, S., & Chawarska, K. (2014). Speech disturbs face scanning in 6-month-old infants who develop autism spectrum disorder. *Biological Psychiatry, 75,* 231–237. doi:10.1016/j.biopsych.2013.07.009

Sumby, W. H., & Pollack, I. (1954). Visual Contribution to Speech Intelligibility in Noise. *The Journal of the Acoustical Society of America, 26,* 212. doi:10.1121/1.1907309

Tenenbaum, E. J., Shah, R. J., Sobel, D. M., Malle, B. F., & Morgan, J. L. (2013). Increased Focus on the Mouth Among Infants in the First Year of Life: A Longitudinal Eye-Tracking Study. *Infancy, 18,* 534–553. doi:10.1111/j.1532-7078.2012.00135.x

Tenenbaum, E. J., Sobel, D. M., Sheinkopf, S. J., Malle, B. F., & Morgan, J. L. (2014). Attention to the mouth and gaze following in infancy predict language development. *Journal of Child Language, 42,* 1173-1190. doi:10.1017/S0305000914000725

Vaillant-Molina, M., Bahrick, L. E., & Flom, R. (2013). Young infants match facial and vocal emotional expressions of other infants. *Infancy*, *18*, E97-E111. doi:10.1111/infa.12017

Vanderwert, R. E., Westerlund, A., Montoya, L., Mccormick, S. A., Miguel, H. O., & Nelson, C. A. (2014). Looking to the eyes influences the processing of emotion on face-sensitive event-related potentials in 7-month-old infants. *Developmental Neurobiology*, *75*, 1154-1163. doi:10.1002/dneu.22204

Weikum, W. M., Vouloumanos, A., Navarra, J., Soto-Faraco, S., Sebastián-Gallés, N., & Werker, J. F. (2007). Visual language discrimination in infancy. *Science*, *316*, 1159. doi:10.1126/science.1137686

Werker, J. F., Byers-Heinlein, K., & Fennell, C. T. (2009). Bilingual beginnings to learning words. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *364*, 3649–3663. doi:10.1098/rstb.2009.0105

Wheeler, A., Anzures, G., Quinn, P. C., Pascalis, O., Omrin, D. S., & Lee, K. (2011). Caucasian infants scan own- and other-race faces differently. *PLoS ONE*, *6*. doi:10.1371/journal.pone.0018621

Young, G. S., Merin, N., Rogers, S. J., & Ozonoff, S. (2009). Gaze behavior and affect at 6 months: Predicting clinical outcomes and language development in typically developing infants and infants at risk for autism. *Developmental Science*, *12*, 798–814. doi:10.1111/j.1467-7687.2009.00833.x

Figures

Figure 1. Representative screen shots for each of the three videos and the attractor. (a) Experiment 1 with infant faces (b) Experiment 2 with adult faces.

Table 1. Mean and standard deviation (in parenthesis) of the PTLT to the mouth and eye regions for each condition and language background for experiment 1.

	Neutral		**Cry**		**Laugh**	
	Eyes	**Mouth**	**Eyes**	**Mouth**	**Eyes**	**Mouth**
Monolinguals	0.37 (0.29)	0.24(0.25)	0.25(0.23)	0.65(0.26)	0.30(0.29)	0.54 (0.27)
Bilinguals	0.29 (0.29)	0.31(0.26)	0.14(0.23)	0.80(0.27)	0.21(0.30)	0.65 (0.28)
Mean	0.33(0.28)	0.28(0.25)	0.20(0.22)	0.73(0.25)	0.26 (0.26)	0.60(0.26)

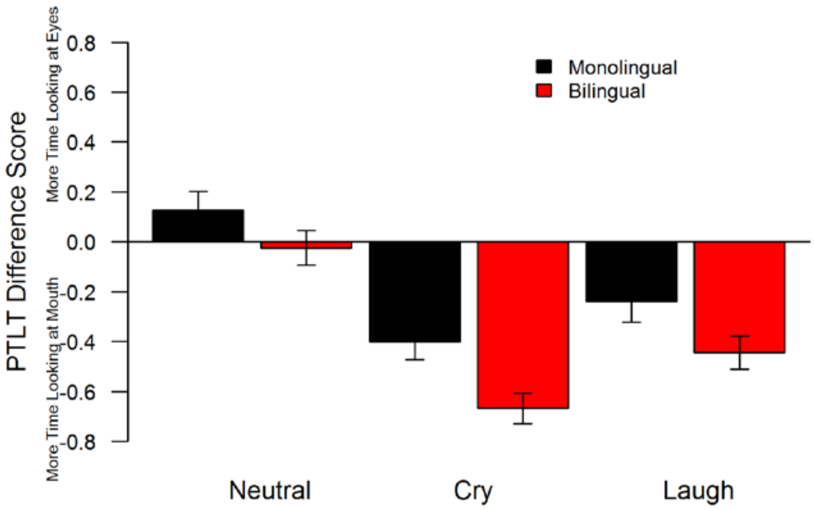

Figure 2. Experiment 1. PTLT Difference Score calculated by PTLT for the eyes minus PTLT for the mouth for both monolingual and bilingual infants in the three emotion conditions. Bilingual infants looked longer at the mouth and shorter to the mouth region as compared to monolingual infants. The error bars indicate the standard error of the mean.

Table 2. Mean and standard deviation (in parenthesis) of the PTLT to the mouth and eye regions for each condition and language background for experiment 2.

		Neutral		**Cry**		**Laugh**	
		Eyes	**Mouth**	**Eyes**	**Mouth**	**Eyes**	**Mouth**
8 months	**Monolinguals**	0.57 (0.19)	0.12(0.17)	0.76(0.18)	0.07(0.10)	0.65(0.19)	0.28 (0.21)
	Bilinguals	0.43 (0.19)	0.26(0.22)	0.43(0.18)	0.38(0.21)	0.45(0.20)	0.46 (0.25)
	Mean	0.50 (0.20)	0.19(0.21)	0.59(0.24)	0.23(0.23)	0.55(0.22)	0.37(0.24)
12 months	**Monolinguals**	0.34 (0.14)	0.33(0.23)	0.45(0.26)	0.43(0.29)	0.36(0.25)	0.57 (0.31)
	Bilinguals	0.47 (0.21)	0.24(0.21)	0.49(0.22)	0.36(0.20)	0.42(0.29)	0.53 (0.29)
	Mean	0.40(0.19)	0.29(0.22)	0.47(0.24)	0.40(0.25)	0.39(0.27)	0.55 (0.29)

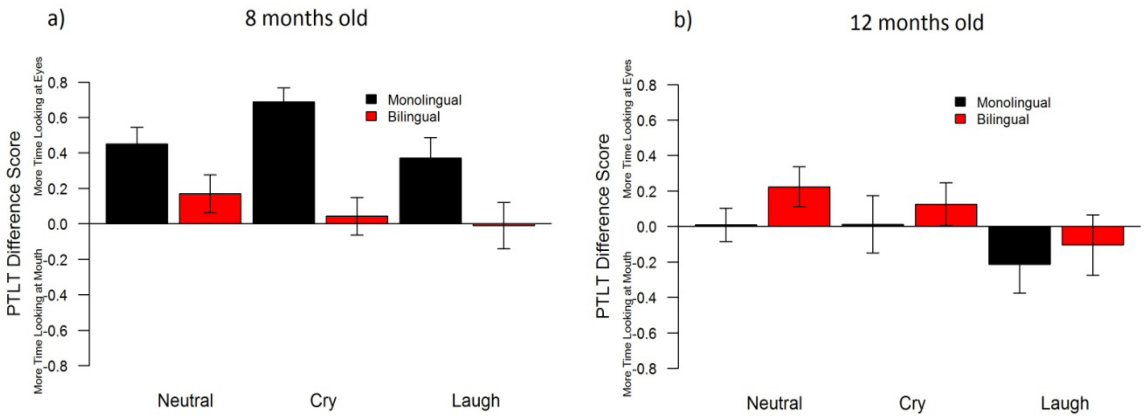

Figure 3. Experiment 2. PTLT Difference Score calculated by PTLT for the eyes minus PTLT for the mouth for both monolingual and bilingual infants in the three emotion conditions. The error bars indicate the standard error of the mean. a) 8-month-old bilingual infants looked longer at the mouth region and shorter to the eye region as compared to monolingual infants. b) 12-month-old group. Bilingual and monolingual infants did not differ in their looking behavior.