Narrowing of intersensory speech perception in infancy

Ferran Pons[a], David J. Lewkowicz[b,1], Salvador Soto-Faraco[c], and Núria Sebastián-Gallés[c]

[a]Departament de Psicologia Bàsica, Universitat de Barcelona, Pg. Vall d'Hebrón 171, 08035 Barcelona, Spain; [b]Department of Psychology, Florida Atlantic University, 777 Glades Road, Boca Raton, FL 33431; and [c]Institució Catalana de Recerca i Estudis Avançats (ICREA) and Departament de Tecnologies de la Informació i les Comunicacions, Universitat Pompeu Fabra, c/ Tanger 122-140, 08018 Barcelona, Spain

Edited by Dale Purves, Duke University Medical Center, Durham, NC, and approved May 20, 2009 (received for review April 15, 2009)

The conventional view is that perceptual/cognitive development is an incremental process of acquisition. Several striking findings have revealed, however, that the sensitivity to non-native languages, faces, vocalizations, and music that is present early in life declines as infants acquire experience with native perceptual inputs. In the language domain, the decline in sensitivity is reflected in a process of perceptual narrowing that is thought to play a critical role during the acquisition of a native-language phonological system. Here, we provide evidence that such a decline also occurs in infant response to multisensory speech. We found that infant intersensory response to a non-native phonetic contrast narrows between 6 and 11 months of age, suggesting that the perceptual system becomes increasingly more tuned to key native-language audiovisual correspondences. Our findings lend support to the notion that perceptual narrowing is a domain-general as well as a pan-sensory developmental process.

audiovisual speech | infants | perceptual narrowing

Conventional developmental theory holds that perceptual and cognitive functions broaden and improve as humans grow and mature (1). Mounting evidence has shown, however, that younger infants' perceptual sensitivity to speech, music, and faces is broader than older infants' sensitivity. This is evident in findings showing that younger infants respond to native as well as non-native perceptual inputs but that older infants respond only to native inputs (2–6). These findings are contrary to the conventional view and suggest that perceptual narrowing plays a key role in the emergence of adult-like perceptual expertise.

In the speech perception domain, evidence shows that an initially broad perceptual response profile to phonetic contrasts in native and non-native languages narrows across the first year of life. This narrowing leads to a decline in responsiveness to phonetic contrasts that are not phonemically relevant in the infant's native language (3, 7, 8). For example, young English-learning infants can perceive particular phonetic contrasts in English as well as in Hindi. In contrast, older English-learning infants no longer respond to the Hindi contrasts (3). Moreover, the response decline to non-native contrasts that is observed during the first year of life is accompanied by an increase in responsiveness to native contrasts (7).

Studies of infant responsiveness in other domains of perceptual processing have provided additional evidence of narrowing. For example, studies of infant response to visually specified speech have found that 4- and 6-month-old English and French monolingual infants can discriminate between silent French and English visual articulations but that 8-month-old infants no longer do (4). Studies of infant response to music have found that 6-month-old American infants can perceive violations in both Western and non-Western musical rhythms but that 12-month-old infants do not respond to violations of non-Western rhythms (8). Finally, studies of face perception have provided two types of evidence of narrowing. First, these studies have shown that whereas 6-month-old infants can discriminate between different human faces and between different monkey faces, 9-month-old infants discriminate only between different human faces (2). Similarly, these studies have found that whereas 3- and 6-month-old infants can discriminate between the faces of different ethnic groups, 9-month-old infants can discriminate only between faces in their native ethnic group (5). Together, all of these findings suggest that perceptual narrowing is a domain-general developmental process.

It is important to emphasize that perceptual narrowing does not reflect a complete loss of perceptual sensitivity to non-native sensory inputs; rather, it reflects a reorganization of perceptual mechanisms that then leads to decreased sensitivity to non-native sensory inputs (9). The reorganization is driven by perceptual experience that, in the vast majority of cases, is restricted to one's native auditory and visual environment. The effects of such selective experience can be seen in the results of studies showing that additional exposure to non-native stimulus materials can maintain responsiveness to these materials past the age of narrowing. For example, infants who grow up in bilingual homes continue to respond to visible articulation differences in a non-native language at an age when monolingual infants no longer respond to them (6). Similarly, Western infants' response to non-Western musical rhythms can be maintained by additional exposure to non-Western rhythms (6). Finally, infant response to the faces of other species can be maintained by additional experience with such faces during the period of decline (10).

Although the foregoing evidence clearly shows that perceptual narrowing occurs within several domains of perception and suggests that this is a domain-general feature of early perceptual development, this evidence comes from studies of unisensory processing. If, however, perceptual narrowing is a domain-general process, then multisensory processing also should undergo narrowing. Establishing that this is the case is critical for gaining further insights into the general aspects of perceptual development because our world and, thus, most of our perceptual experiences are multisensory in nature. This means that we must constantly integrate information across sensory modalities to achieve coherent and meaningful representations (11, 12). The tendency of humans to engage in cross-modal integration during perception is most dramatically illustrated by the well-known McGurk (13) and ventriloquism (14) effects. In both cases, when observers are presented with conflicting auditory and visual inputs – conflicting phonetic information in the case of the McGurk effect and spatially discrepant cues in the case of the ventriloquism effect—perceptual resolution of the discrepancy leads to illusory experiences. These kinds of experiences demonstrate that multisensory integration is a relatively automatic and, thus, essential, feature of perceptual functioning.

Author contributions: F.P., D.J.L., S.S.-F., and N.S.-G. designed research; F.P. and D.J.L. performed research; F.P. and D.J.L. contributed new reagents/analytic tools; F.P. and D.J.L. analyzed data; and F.P., D.J.L., S.S.-F., and N.S.-G. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

[1]To whom correspondence should be addressed. E-mail: lewkowic@fau.edu.

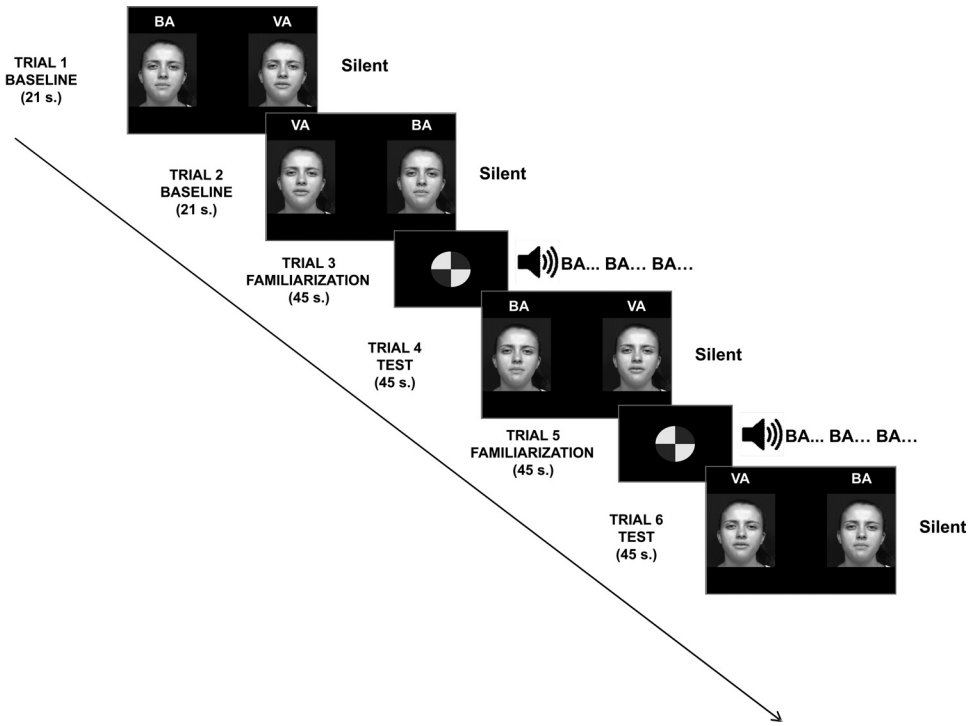

Fig. 1. Schematic representation of the procedure used in the infant portion of the current study. Only one of the two auditory conditions is shown.

There is little doubt that intersensory integration is critical for adaptive functioning and it is likely that it also narrows in early development. To test this possibility, Lewkowicz and Ghazanfar (15) investigated cross-species intersensory integration in 4-, 6-, 8- and 10-month-old infants and predicted that younger but not older infants would exhibit evidence of integration. Consistent with this prediction, Lewkowicz and Ghazanfar found that the two younger groups of infants successfully matched visible and audible monkey calls but that the two older groups did not. Subsequent studies (16) have indicated that (i) younger infants' successful intersensory matching is based on onset and offset synchrony of corresponding visible and audible calls, (ii) older infants' failure to match is not due to unisensory processing deficits, and (iii) decreased responsiveness to cross-species intersensory relations persists into later development.

The findings of Lewkowicz and Ghazanfar (15) demonstrate that intersensory perception is initially broad and then narrows by the second half of the first year of life. When these findings are considered together with the findings of perceptual narrowing in the speech, music, and face perception domains, they lead to the conclusion that experience-dependent perceptual narrowing is not only a domain-general but a pan-sensory process as well. If that is the case, then this raises the possibility that intersensory perception of non-native speech also narrows and that this results in a decline of infants' ability to match non-native visible and audible phonemes. The purpose of this study was to put this hypothesis to experimental test.

We used a variant of the intersensory matching procedure (17) to investigate whether intersensory perception of non-native speech narrows in early development. This procedure involves presenting side-by-side visual stimuli that differ on some dimension and measuring infants' visual preference after they have been familiarized to an auditory stimulus that corresponds to one of the two visual stimuli. If it is found that infants look at the corresponding visual stimulus more than at the non-corresponding stimulus, then it can be concluded that these infants must have perceived the relationship between the auditory stimulus and the corresponding visual stimulus and, thus, that they must have performed intersensory matching. The fact that, in this procedure, the sound was not presented at the same time as the visual stimuli ensured that audiovisual synchrony did not mediate intersensory matching and that infants were basing their response on the extraction of the amodal invariance of the specific phoneme. Thus, in the present study, we first presented side-by-side silent videos of two different syllables to 6- and 11-month-old, Spanish-learning and English-learning infants and measured their looking preferences. We then familiarized the infants with one of the corresponding auditory syllables and followed this up with a second silent test of visual preferences.

Prior studies of infant intersensory perception of talking faces have shown that infants as young as 2 months of age and as old as 18 months can perceive the amodal invariance of human facial gestures and vocalizations (18–20). Critical from the standpoint of the current study is the fact that all of these studies have investigated infants' intersensory perception of native speech only (21–23). Thus, finding that younger infants can match the audible and visible attributes of non-native speech but that older infants do not would provide evidence that infant sensitivity to intersensory correspondence in non-native speech does indeed narrow during the first year of life.

We tested separate groups of 24 6- and 11-month-old Spanish-learning infants, along with 24 6-month-old and 16 11-month-old English-learning infants, to determine whether these infants could match the visual and auditory phonetic information specifying the English contrast /ba/ - /va/. The /b/ and /v/ are homophones in Spanish and, thus, fall under a single, tightly-tuned, /b/ phonemic category. This means that the phonetic distinction between a /ba/ and /va/ does not exist in Spanish. Consequently, if narrowing of intersensory responsiveness to non-native speech occurs, then younger, but not older, Spanish-learning infants should exhibit successful intersensory matching. In contrast, English-learning infants should exhibit successful intersensory matching at both ages.

The experiment consisted of six trials (Fig. 1). The first two

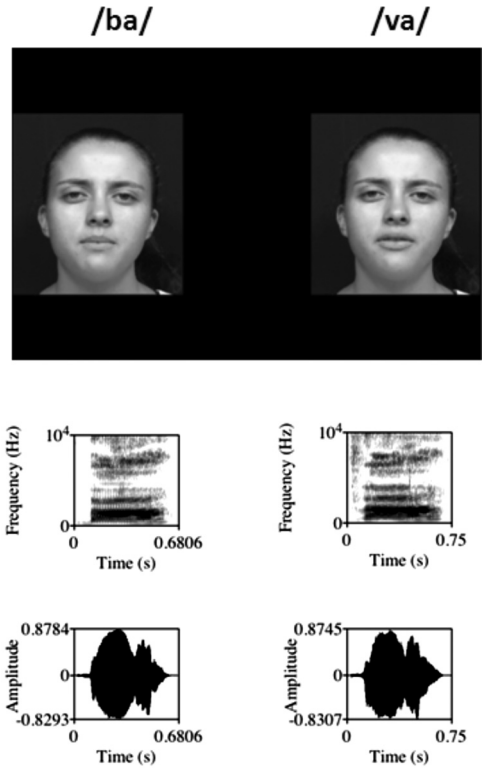

Fig. 2. Single video frames of the facial gestures representing the /ba/ and /va/ syllables. Depicted below the visual gestures are the spectrograms (*Top*) and oscillograms (*Bottom*) of the corresponding syllables.

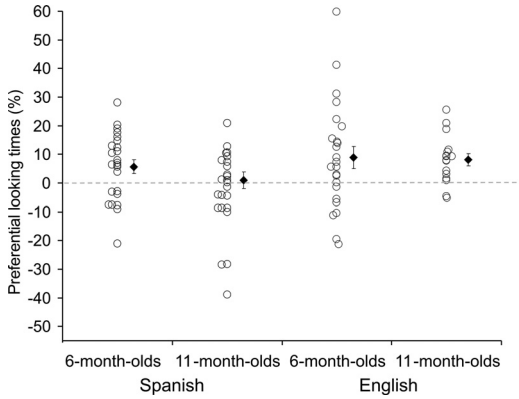

Fig. 3. Distribution of looking time difference scores to the matching face (percentage of total time that infants looked at the matching face during the test trials minus the percentage of total time that they looked at the matching face during baseline trials) for each language and age group. Open circles represent each infant's difference score; filled circles with error bars represent mean difference score and standard error of the mean for each group.

trials constituted the baseline condition and consisted of the presentation of side-by-side silent video clips of a female bilingual Spanish–English speaker repeatedly producing a /ba/ syllable on one side and a /va/ syllable on the other side (counterbalanced for side across these two trials). The remaining four trials constituted the test condition and investigated whether hearing one of the syllables during auditory-only familiarization trials would lead to longer looking at the matching visual syllable in subsequent visual preference test trials. This was done by interspersing two auditory familiarization trials (i.e., the third and the fifth trials) with two test trials (the fourth and sixth trials). Half of the infants heard the /ba/ syllable and the other half heard the /va/ syllable during the two familiarization trials while a colorful ball rotated on the computer screen (Fig. 2 shows the spectrograms and oscillograms of each syllable and the corresponding facial gestures). During each of the two test trials that followed each familiarization trial, infants viewed the two visible syllables presented side-by-side and in silence once again (the side on which the syllables were presented was counterbalanced across these two test trials).

Results

Infant Matching of Audible and Visible Speech. Examination of the data from the baseline trials indicated that the duration of looking directed at each of the two (silently presented) visible syllables did not differ in any of the four groups of infants; Spanish 6-month-olds [t (23) = 1.32, n.s.], Spanish 11-month-olds [t (23) = −0.15, n.s.]; English 6-month-olds [t (23) = −0.56, n.s.], English 11-month-olds [t (15) = 1.58, n.s.]. To determine whether infants performed intersensory matching, we computed the proportion of time that infants looked at the matching visible syllable during the baseline condition and compared it with the proportion of time that infants looked at the matching visible syllable during the test trials in the test condition. Each of these two proportion scores was computed by dividing the total amount of looking directed at the matching visible syllable by the total amount of looking directed at the matching plus the non-matching syllables.

We expected that the proportion of looking time directed at the matching visible syllable would be greater during the test condition than during the baseline condition in those infants who made successful intersensory matches. Given our earlier prediction that Spanish-learning infants would exhibit perceptual narrowing but that English-learning infants would not, we expected that infants in both language groups would exhibit intersensory matching at 6 months of age but that only the English-learning infants would do so at 11 months of age. Consistent with our predictions, at 6 months of age, both language groups exhibited an increase in looking time to the matching visible syllable after auditory familiarization, but at 11 months of age only the English-learning infants did so (Fig. 3). To test the statistical reliability of these effects, we first compared the proportion of looking scores across the two language groups separately at each age. To do this, we submitted the proportion of looking scores to a repeated-measures analysis of variance (ANOVA), with Condition (baseline, test) as the within-subjects factor and Language Group (Spanish, English) as the between-subjects factor. At 6 months of age, the ANOVA yielded a main effect of Condition [F (1, 46) = 10.54, $P < 0.002$] and no significant interactions, indicating that, together, the infants in the two language groups spent a greater proportion of their looking time on the matching syllable after familiarization than before it, and that this was not affected by whether they were learning Spanish or English. In contrast, at 11 months of age, the ANOVA yielded a Condition by Language Group interaction [F (1, 38) = 7.14, $P < 0.02$], indicating that the proportion of looking at the matching visible syllable following familiarization than before it was affected by whether infants were learning Spanish or English.

To further test these effects, we analyzed the data as a function of age separately for each language group. Thus, we submitted the proportion of looking time scores for each language group, respectively, to a repeated-measures ANOVA, with Condition (baseline, test) as the within-subjects factor and Age (6 or 11 months) as the between-subjects factor. The ANOVA for the Spanish-learning infants revealed that proportion of looking directed at the matching visible syllable after familiarization rather than before it depended on age [Condition × Age interaction, F (1, 46) = 4.58, $P < 0.05$]. As can be seen in Fig.

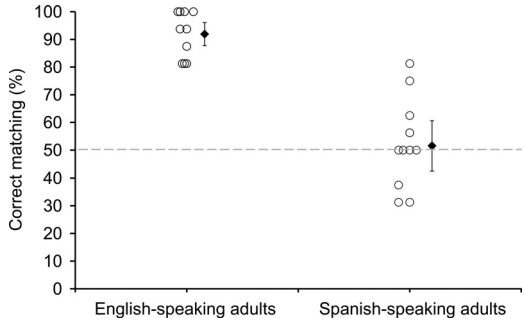

Fig. 4. Percentage of correct matches made by the adult participants in the study. Open circles represent each participant's matching score; filled circles represent group mean score and standard error of the mean.

3, the 6-month-old Spanish-learning infants looked longer at the matching visible syllable after familiarization [baseline vs. test, t (23) = 2.41; $p < .03$], whereas the 11-month-old Spanish-learning infants did not [t (23) = 0.78, n.s.]. The ANOVA for the English-learning infants revealed that their proportion of looking time differed as a function of Condition [F (1, 38) = 11.62, $P < 0.03$] and that this pattern of looking did not differ as a function of Age. Separate two-tailed t-tests indicated that both English-learning age groups exhibited a greater proportion of looking at the visible syllable after familiarization than before it: (t (23) = 2.33, $P < 0.03$ at 6 months and t (15) = 3.85, $P < 0.01$ at 11 months). Overall, the present findings confirm our prediction that infants' ability to perceive the amodal invariance of non-native speech is initially broad and that it narrows toward the end of the first year of life, whereas the ability to perceive the amodal invariance of native speech does not narrow.

Adult Matching of Audible and Visible Speech. Although the results from the Spanish-and English-learning infants unequivocally demonstrate that narrowing of intersensory integration of non-native audible and visible speech occurs in early development, these data do not indicate whether the narrowing is maintained into adulthood. To investigate this question, we tested 12 Spanish-speaking and 10 American English-speaking adults' ability to match the same audible and visible syllables that we used with the infants. If the decline in intersensory perception of non-native speech is maintained into adulthood, then the Spanish-speaking adults should not exhibit successful intersensory matching, whereas the English-speaking adults should.

To test the adults, we presented them with a total of 16 forced-choice intersensory matching trials. In each trial, participants first heard one of the two possible syllables (/va/ or /ba/) repeated twice and, 1 second later, they watched the two silent visible syllables presented side-by-side twice. On half of the trials, participants heard the /ba/ syllable and on the rest they heard the /va/ syllable (with order randomized across trials). The participants were asked to choose which of the two faces corresponded to the previously heard syllable. As can be seen in Fig. 4, the results were consistent with the maintenance of perceptual narrowing. The Spanish-speaking adults failed to match the audible and visible syllables (correct match, 51.56%; SD, 18.07%; not significantly different from chance performance [t (11) = 0.31, n.s.]), whereas the English-speaking adults matched the audible and visible syllables correctly in 91.9% of the cases, SD = 8.36% [t (9) = 15.84, $P < 0.001$]. Comparison of the percentage of correct matches across these two groups indicated that the English-speaking adults' matching scores were significantly higher than the Spanish-speaking adults' scores [t (20) = 6.72, $P < 0.001$].

Discussion

Intersensory perception is critical for the emergence of a unified and meaningful representation of the world. In the speech perception domain, this includes the ability to perceive the visual and auditory attributes of spoken language as part of a unified event. Despite the fact that it has now been clearly established that infant perception of non-native auditory (3, 24, 25) and visual (4) speech narrows and that infant intersensory response to the facial and vocal calls of another species (15) also narrows, no studies to date have investigated whether infant perception of human multisensory speech likewise narrows around native phonemes. Here, we investigated Spanish-learning and English-learning infants' and adults' response to a phonetic intersensory correspondence that is not phonemically relevant in Spanish but that is relevant in English. Our prediction was that the Spanish participants would exhibit narrowing of their intersensory response if narrowing is a general characteristic of early perceptual development, but that the English participants would not. Consistent with this prediction, results showed that Spanish-learning infants exhibited evidence of narrowing by 11 months of age. Moreover, they showed that this narrowing was maintained into adulthood in that Spanish adults, like the 11-month-old Spanish-learning infants, failed to match the audible and visible non-native phonemes. In contrast, results showed that English-learning infants did not exhibit evidence of narrowing, in that both age groups matched the audible and visible attributes of /ba/ and /va/ and, as expected, that the English adults did too. Overall, the findings from this study show that the process of narrowing leaves older infants with a perceptual system that is more finely tuned to the language of their environment but, as a consequence, one that is no longer sensitive to non-native language categories.

When the present findings of the narrowing of intersensory perception within the same species (human subjects discriminating human stimuli) are considered in the context of previous findings of the narrowing of cross-species intersensory perception (15), they further strengthen the claim that perceptual narrowing is both a domain-general and a pan-sensory developmental process. Such an all-encompassing process that is so broadly tuned in early infancy is especially adaptive because it makes it easier for young and inexperienced infants to take advantage of the redundancy that multisensory speech typically offers. In general, audiovisual redundancy increases the overall salience of the information (26) and, as a result, enhances perception, learning, and discrimination (12, 27–29). Although the present study did not directly investigate the role of cross-modal redundancy in infant speech perception, in their daily lives infants usually experience speech as a multisensory event. As a result, as long as infants can take advantage of the cross-modal redundancy inherent in speech and, critically, as long as they can do so regardless of whether the speech is native or non-native, then redundancy is certain to facilitate their acquisition of their native language categories. This is particularly important early in life when inexperience and lack of lexical knowledge might make it difficult for infants to perceive complex forms of intersensory unity in multisensory speech.

In addition to playing an important role in the acquisition of native language categories early in life, multisensory redundancy can also facilitate the acquisition of a second language later in life. For example, proficient second language (L2) speakers can discriminate a difficult non-native contrast when it is audiovisual but not when it is only heard (30). If that is the case, then it is also possible that multisensory redundancy might facilitate L2 acquisition before adulthood. This possibility remains to be explored but, in the meantime, suggests that the initial decline in auditory, visual, and audiovisual sensitivity to non-native phonemes may be reversible to some extent during later acqui-

sition of L2 by providing explicit experience with audiovisual L2 speech.

In conclusion, previous studies have shown that, as infants grow and amass highly selective experience with their native language, their perception of speech becomes reorganized and, as it does, they become more attuned to the auditory (5) and visual (6) attributes of their native language environment. This unisensory tuning takes place while infants already possess the ability to perceive the auditory and visual attributes of their native language as a perceptually unified event (18–20). What has not been known until now, however, is whether the narrowing observed in infant perception of unisensory non-native sensory inputs generalizes to infant perception of multisensory speech. Our findings demonstrate that infant perception of audiovisual speech also narrows by the end of the first year of life, resulting in perceptual tuning around native-language speech categories. In addition, our findings are consistent with earlier evidence (15) that perceptual narrowing and the resulting development of perceptual expertise is governed by a domain-general and pan-sensory process that begins at birth (31).

Materials and Methods

Participants. In all, 114 full-term infants were recruited and tested in Barcelona, Spain (63 infants) and in Boca Raton, Florida (51 infants). All of these infants were raised in a monolingual environment, hearing Spanish (Barcelona) or English (Boca Raton) at least 90% of the time according to parental report. Of these infants, 26 did not contribute any usable data because of fussiness or crying (10 infants in Spain and three in the United States), parental interference (two infants in Spain), experimental error (three infants in Spain; one in the US), and less than 90% of English heard at home (seven infants in the US). The remaining 88 infants who contributed data consisted of a group of 24 6-month-old Spanish-learning infants (mean age, 6 months; range, 5.2–6.1 months; 12 female and 12 male); a group of 24 6-month-old English-learning infants (mean age, 6 months; range, 5.1–6.1 months; 14 male and 10 female); a group of 24 11-month-old Spanish-learning infants (mean age, 11 months; range, 10.2–11.2 months; 12 female and 12 male); and a group of 16 11-month-old English-learning infants (mean age, 11 months; range, 10–11.2 months; eight female and eight male).

Apparatus and Stimuli. Infants were seated either in an infant seat or in their parent's lap in a sound-attenuated and dimly illuminated room, ≈100 cm in front of a projector screen (Barcelona) or ≈50 cm in front of two side-by-side, 17-inch (43.2 cm), liquid crystal display monitors (Boca Raton) separated by a 6.7-cm gap. The experimenter, who could not see the stimuli and was unaware of specific test conditions, was located in a separate room where s/he controlled stimulus presentation and observed and recorded the infant's looking responses via a closed-circuit camera located under the screen (Barcelona) or between the two monitors for off-line analyses.

The video clips showed a full-face image of a female bilingual Spanish–English speaker silently uttering the English syllables /ba/ and /va/ (Fig. 2). The videos, each measuring 43.2 cm across, were presented 48.3 cm apart (measured center to center; Barcelona) or on the two side-by-side monitors (Boca Raton).

The sound pressure level of the audio syllable was set at 65 ± 5 dB [A]. The trials involving the presentation of video clips (baseline and test trials) began with an image of a green expanding and retracting ball for 5 seconds to attract the infant's attention toward the screen, followed by the presentation of the video clips. The familiarization trials, during which infants heard the audio syllable, were accompanied by the presentation of a colorful rotating ball in the center of the infants' visual field to provide the infants with something to look at while they listened to the syllable.

Procedure. The experiment consisted of six trials (Fig. 1). The first two trials were the baseline trials and each had a duration of 21 seconds. During these trials, we presented two silent, side-by-side video clips of the female speaker articulating the /ba/ on one side and the /va/ on the other side (the side of syllable presentation was counterbalanced across these trials). The third and fifth trials were auditory familiarization trials in which we presented one of the two syllables (/ba/ or /va/) acoustically for 45 seconds (once every 2.3 seconds for a total of 15 repetitions). Half of the infants in each age group heard the /ba/ syllable and the other half heard the /va/ syllable during the familiarization trials. Each familiarization trial was followed by a 45-second intersensory matching test trial (trials 4 and 6), during which we presented the two visible syllables side-by-side in silence. Here, we once again measured visual preferences to determine whether listening to one of the syllables first affected subsequent looking preferences when infants were shown the two visual syllables. The side of visual syllable presentation was counterbalanced across these two intersensory matching test trials.

ACKNOWLEDGMENTS. This work was supported by the Spanish Ministerio de Educación y Ciencia (Consolider Ingenio 2010 CE-CSD2007–00121 and CDS00012, SEJ2007–60751, and SEJ 2007–64103) and by the National Science Foundation (grant BCS-0751888 to D.J.L.). We thank Janet F. Werker for facilitating the cross-Atlantic collaboration and Helena Ràfols and Kelly McMullen for their assistance in infant recruitment and data collection.

1. Piaget J (1952) *The origins of intelligence in children* (International Universities Press, New York).
2. Pascalis O, Haan MD, Nelson CA (2002) Is face processing species-specific during the first year of life? *Science* 296:1321–1323.
3. Werker JF, Tees RC (1984) Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behav Dev* 7:49–63.
4. Weikum WM, Vouloumanos A, Navarra J, Soto-Faraco S, Sebastián-Gallés N, Werker JF (2007) Visual language discrimination in infancy. *Science* 316:1159.
5. Kelly DJ, Quinn PC, Slater AM, Lee K, Ge L, Pascalis O (2007) The other-race effect develops during infancy: Evidence of perceptual narrowing. *Psychol Sci* 18:1084–1089.
6. Hannon EE, Trehub SE (2005) Tuning in to musical rhythms: Infants learn more readily than adults. *Proc Natl Acad Sci USA* 102:12639–12643.
7. Kuhl PK, Stevens E, Hayashi A, Deguchi T, Kiritani S, et al. (2006) Infants show a facilitation effect for native language phonetic perception between 6 and 12 months. *Dev Sci* 9:F13–F21.
8. Hannon EE, Trehub SE (2005) Metrical categories in infancy and adulthood. *Psychol Sci* 16:48–55.
9. Werker JF, Tees RC (2005) Speech perception as a window for understanding plasticity and commitment in language systems of the brain. *Dev Psychobiol* 46(Suppl):233–234.
10. Pascalis O, Scott LS, Kelly DJ, Shannon RW, Nicholson E, et al. (2005) Plasticity of face processing in infancy. *Proc Natl Acad Sci USA* 102:5297–300.
11. Lewkowicz DJ (2000) The development of intersensory temporal perception: An epigenetic systems/limitations view. *Psychol Bull* 126:281–308.
12. Stein BE, Meredith MA (1993) *The merging of the senses* (MIT Press, Cambridge, MA).
13. McGurk H, MacDonald J (1976) Hearing lips and seeing voices. *Nature* 264:229–239.
14. Jack CE, Thurlow WR (1973) Effects of degree of visual association and angle of displacement on the ''ventriloquism'' effect. *Percept Motor Skills* 37:967–979.
15. Lewkowicz DJ, Ghazanfar AA (2006) The decline of cross-species intersensory perception in human infants. *Proc Natl Acad Sci USA* 103:6771–6774.
16. Lewkowicz DJ, Sowinski R, Place S (2008) The decline of cross-species intersensory perception in human infants: Underlying mechanisms and its developmental persistence. *Brain Res* 1242:291–302.
17. Lewkowicz DJ, Turkewitz G (1981) Intersensory interaction in newborns: Modification of visual preferences following exposure to sound. *Child Dev* 52:827–832.
18. Kuhl PK, Meltzoff AN (1982) The bimodal perception of speech in infancy. *Science* 218:1138–1141.
19. Patterson ML, Werker JF (1999) Matching phonetic information in lips and voice is robust in 4.5-month-old infants. *Infant Behav Dev* 22:237–247.
20. Patterson ML, Werker JF (2003) Two-month-old infants match phonetic information in lips and voice. *Dev Science* 6:191–196.
21. Poulin-Dubois D, Serbin LA, Derbyshire A (1998) Toddlers' intermodal and verbal knowledge about gender. *Merrill-Palmer Q* 44:338–354.
22. Poulin-Dubois D, Serbin LA, Kenyon B, Derbyshire A (1994) Infants' intermodal knowledge about gender. *Dev Psychol* 30:436–442.
23. Walker-Andrews AS (1986) Intermodal perception of expressive behaviors: Relation of eye and voice? *Dev Psychol* 22:373–377.
24. Best CT, McRoberts GW, Sithole NM (1988) Examination of perceptual reorganization for nonnative speech contrasts: Zulu click discrimination by English-speaking adults and infants. *J Exp Psychol Hum Percept Perform* 14:345–360.
25. Kuhl PK, Williams KA, Lacerda F, Stevens KN, et al. (1992) Linguistic experience alters phonetic perception in infants by 6 months of age. *Science* 255:606–608.
26. Sumby WH, Pollack I (1954) Visual contribution to speech intelligibility in noise. *J Acoust Soc Am* 26:212–215.
27. Lewkowicz DJ, Kraebel K (2004) in *Handbook of Multisensory Processing*, eds Calvert G, Spence C, Stein B (MIT Press, Cambridge).
28. Partan S, Marler P (1999) Communication goes multimodal. *Science* 283:1272–1273.
29. Rowe C (1999) Receiver psychology and the evolution of multicomponent signals. *Anim Behav* 58:921–931.
30. Navarra J, Soto-Faraco S (2007) Hearing lips in a second language: Visual articulatory information enables the perception of second language sounds. *Psychol Res* 71:4–12.
31. Lewkowicz DJ, Leo I, Simion F (2010) Intersensory perception at birth: Newborns match non-human primate faces & voices. *Infancy*, in press.