The phonetic landscape in infant consonant perception is an uneven terrain

Youngja Nam[1,2] and Linda Polka[1,2]
[1]School of Communication Sciences and Disorders, McGill University
[2]Centre for Research on Brain, Language & Music
2001 McGill College Avenue, 8th floor, Montreal, Quebec, H3A 1G1

Youngja Nam: young.nam@mail.mcgill.ca
Linda Polka: Linda.polka@mcgill.ca

Correspondence concerning this article should be addressed to Linda Polka, School of Communication Sciences and Disorders, McGill University, 2001 McGill College Avenue, 8th floor, Montreal, Quebec, H3A 1G1, Canada.
Contact: Linda.polka@mcgill.ca

PostPrint – Nam, Y. & Polka, L. (2016) The phonetic landscape in infant consonant perception is an uneven terrain, *Cognition*, 155: 57-66.

Research highlights

- Young infants display a perceptual bias favoring stop /b/ over fricative /v/.
- This is the first perceptual bias documented in infant consonant manner perception.
- This perceptual bias may be due to a listening preference.
- Like vowels, some consonants are more perceptually salient than others in infancy.
- Both discrimination abilities and biases shape infant consonant manner perception.

PostPrint – Nam, Y. & Polka, L. (2016) The phonetic landscape in infant consonant perception is an uneven terrain, *Cognition*, 155: 57-66.

Abstract

Previous research revealing universal biases in infant vowel perception forms the basis of the Natural Referent Vowel (NRV) framework (Polka & Bohn, 2011). To explore the feasibility of extending this framework to consonant manner perception, we investigated perception of the stop vs. fricative consonant contrast /b/-/v/ to test the hypothesis that young infants will display a perceptual bias grounded in the acoustic-phonetic properties of these sounds. We examined perception of stop-initial /bas/ and fricative-initial /vas/ syllables in English-learning and French-learning 5- to 6-month-olds. The /b/ and /v/ sounds distinguish words in English and French but have different distributional patterns; in spoken English /b/ occurs more frequently than /v/ whereas in spoken French /v/ occurs more frequently than /b/. A perceptual bias favoring /b/ over /v/ emerged in two experiments. In Experiment 1, a directional asymmetry was observed in discrimination; infants noticed when /vas/ changed to /bas/ but not when /bas/ changed to /vas/. In Experiment 2, a robust listening preference favoring stop-initial /bas/ was evident in responses from the same infants. This is the first study to show a perceptual bias related to consonant manner and to directly measure a consonant perception bias within the same infants. These data encourage further efforts to extend the NRV principles to perception of consonant manner. These findings indicate that we need to reform our view of infant speech perception to accommodate the fact that both discrimination abilities and biases shape speech perception during infancy.

Keywords: infant speech perception, perceptual bias, consonants, Natural Referent Vowel framework, asymmetry

PostPrint – Nam, Y. & Polka, L. (2016) The phonetic landscape in infant consonant perception is an uneven terrain, *Cognition*, 155: 57-66.

The phonetic landscape in infant consonant perception is an uneven terrain

1. Introduction

Research over several decades shows that young infants are universal perceivers, born with the ability to discriminate virtually any phonetic contrast that languages use (Saffran,Werker & Werner, 2006). In recent years, there has been increasing evidence that young infants also display some language-universal phonetic biases. These findings enrich our understanding of the infant's initial perceptual capacities. Although infant perception affords access to a detailed phonetic landscape with all or most relevant categories, there appears to be an irregular terrain in which some phonetic categories are more perceptually prominent than others.

Most of the work fueling this enriched view has emerged from research on infant vowel perception. Many studies show that infant discrimination of a vowel change presented in one direction is significantly better compared to when the same change is presented in the reverse direction (for reviews, see Polka & Bohn, 2003, 2011). In young infants, these directional asymmetries are independent of language experience and follow a common pattern – within the traditional vowel space, discrimination is easier when the direction of change is from a central to a more peripheral vowel. Polka and Bohn proposed that these asymmetries arise because infant perception is biased favoring vowels that occupy the most extreme positions in the traditional articulatory/acoustic vowel space (defined by the first and second formant frequencies). There is currently a high level of interest in perceptual asymmetries and increasing evidence that vowel perception is biased at least in the early months of infancy (e.g., Pons, Albareda-Castellot, &

Sebastián-Gallés, 2012) and in adult perception of non-native contrasts as well (e.g., Tyler, Best, Faber, & Levitt, 2014).

To integrate and explain these vowel perception findings, Polka and Bohn (2011) proposed the Natural Referent Vowel (NRV) framework. According to NRV, vowel perception asymmetries reflect a universal perceptual bias that is grounded in acoustic-phonetic properties. When adjacent vowel formants move closer together, energy is focused into a narrower spectral region creating a spectral bump or prominence. This focalization of spectral energy occurs in tandem with changes in vocal tract constriction (in a graded fashion) as you move from the center to the corners of the traditional vowel space. Following Schwartz, Abry, Boë, Ménard and Vallée (2005), Polka & Bohn (2011) proposed that this formant convergence (or focalization) is the acoustic-phonetic pattern behind vowel perception biases. Thus, vowels that are more focal are *natural referents* due to their salient acoustic and articulatory properties which make them easier to detect and to hold in memory. These properties bias perception and give rise to directional asymmetries in discrimination tasks.

At a broad conceptual level, the NRV should also make predictions for consonants along dimensions with comparable auditory-articulatory properties. No previous work has specifically tested these NRV-based predictions, but several studies document directional asymmetries in consonant perception, which have been explained in diverse ways. For example, Kuhl, Stevens, Hayashi, Deguchi, Kiritani, and Iverson (2006) observed directional asymmetries in discrimination of /r/-/l/ by both English-learning and Japanese-learning infants. These findings point to a language-independent effect given that /r-l/ is a native contrast in English but not in Japanese. The authors suggest that discriminating the /ra/ to /la/ change may be more difficult because the /a/ vowel interferes (via forward masking) with the formant transitions for /la/ more than for /ra/.

Tsuji, Mazuka, Cristia and Fikkert (2015) reported another consonant discrimination asymmetry that also points to a language-independent effect. Using a habituation paradigm, they tested Dutch-learning and Japanese-learning 4- to 6-month-olds' discrimination of /ompa/ and /onta/; these non-word stimuli contain contrasting native consonants in both languages. Infants habituated with /ompa/ (with labial consonants) noticed the change to /onta/ but infants habituated to /onta/ (with coronal consonants) did not notice the change to /ompa/ (see also Dijkstra & Fikkert, 2011). It is well documented that infants begin to perceptually tune into native consonant categories between 6 and 12 months (e.g., Kuhl, 1992, 1994; Polka & Werker, 1994; Werker & Tees, 1984). Thus, this finding shows that perception of coronal-labial consonant place is asymmetric before attunement to native language phonetics or phonotactics is underway. Even so, early attunement would not explain the asymmetric findings reported by Tsuji et al. (2015) because coronal consonants occur more frequently than labials in Dutch whereas the reverse pattern is found in Japanese.

Finally, Segal, Hejli-Assi and Kishon-Rabin (2016) observed asymmetric discrimination of the non-native /b-p/ voicing contrast by older (10-12 months) but not younger (4-6 months) infants acquiring Arabic, which has a /b/ category but no /p/ category. The older Arabic infants discriminated a /p/ to /b/ change but not a /b/ to /p/ change; this pattern reveals greater sensitivity to a voicing change in the non-prototypic to prototypic direction compared to the reverse direction, consistent with the Native Language Magnet (NLM) model proposed by Kuhl (1991). No directional effect was evident in younger or older Hebrew-learning infants who experience /b/-/p/ as a native contrast. Thus, these data also suggest an effect of native language attunement.

Miller and Eimas (1996) also found directional asymmetries in discrimination of within-category voice onset time (VOT) variants of English stops in English-learning 3- to 4-month-olds. Although these data are also consistent with the NLM model, it is unclear whether this asymmetry reveals a language-specific or language-independent pattern because only English infants were tested.

To summarize, a few prior studies reveal directional asymmetries in infant consonant perception and support different interpretations. Crucially, no previous work has assessed the possibility that asymmetries arise from differences in the perceptual strength of different sound categories. In this study, we firstly extend documented discrimination asymmetries to a new consonant contrast and secondly demonstrate a listening preference for one of the categories. This listening preference shows both that infants can discriminate in such cases, and that the perception field is not level.

To advance this theoretical position, we bring new data to bear on perceptual asymmetries in consonant perception and explore the feasibility of extending the principles of the NRV framework to consonants. As a first step, we focus on infant perception of consonant manner of articulation, specifically the stop-fricative contrast, /b/ vs. /v/. We propose that young infants will display a perceptual bias favoring stop over fricative consonants and that this bias is a natural consequence of acoustic-phonetic patterns generated by stop and fricative consonant production. These predictions are motivated by prior research, which is reviewed below.

Stop consonants are produced by creating a rapid and complete oral closure, which causes a buildup of intra-oral air pressure, followed by an abrupt release of the closure which is accompanied by a brief burst of the maintained air (e.g., Ladefoged, 1993). This rapid closure/release gesture causes a sharp, abrupt shift in signal amplitude, which is a prominent landmark in the speech signal (e.g., Stevens & Keyser, 1989). Fricatives are produced by forcing air through a narrow constriction in the vocal tract causing a turbulent airflow that generates an audible noise. Fricative production also creates a shift in signal amplitude but the shift is smaller and slower than in stop production (e.g., Reetz & Jongman, 2011; Repp, Liberman, Eccardt, & Pesetsky, 1978). Stop consonants are found in all languages and they are strongly favored over fricatives in phonemic inventories across languages of the world (Maddieson, 1984). Stops, being easier to produce, also appear earlier than fricatives and are often substituted for fricatives in early word production, e.g., an English child might say "**b**egetable" for "**v**egetable".

Perception studies conducted with adults reveal directional asymmetries in discrimination of stop-fricative contrasts consistent with the claim that stops are more salient than fricatives. For example, Tsushima, Shiraki, Yoshida, and Sasaki (2003, 2005) observed directional asymmetries in Japanese listeners' discrimination of the non-native English /ba/-/va/ in a variety of AX discrimination tasks. In this work, the /v/ to /b/ direction of change was consistently easier compared to the reverse direction, revealing a bias favoring stop consonant /b/. A similar pattern was reported by Zhang, Imada, Kawakatsu and Kuhl (2006); their measures of brain activity using magnetoencephalography (MEG) showed that both native English and non-native Japanese listeners were more sensitive to stop /ba/ deviants amidst fricative /va/ stimuli than the other way around. Recently, Nam (2015) examined discrimination of eight stop-fricative contrasts (6 from English; 2 from Persian) by adult native speakers of English, French and Korean. In this research, directional asymmetries were observed for almost every non-native contrast and also for several native stop-fricative contrasts. As in prior studies, discrimination was easier in the fricative to stop direction, indicating a stop bias.

Directional asymmetries involving stop-fricative contrasts have not been investigated in infants. However, asymmetries are reported in research exploring early word recognition and word learning abilities in toddlers who are in the 2nd year of life. Altvater-Mackensen and Fikkert (2010) reported an asymmetry when they tested Dutch-learning 14-month-olds in the switch task using the minimal-pair non-words /paap/ and /faap/, which feature the native Dutch /p-f/ contrast. The switch task is a standard procedure for assessing associative word learning by measuring the ability to map a novel spoken label with a novel visual object (e.g., Stager & Werker, 1997). In this task, Dutch 14-month-olds noticed that the word/object label was switched when they were familiarized with /faap/ as the object label but not when they were familiarized with /paap/ as the object label. Thus, they noticed the fricative-to-stop change but not the reverse stop-to-fricative change. A second experiment confirmed that 14-month-olds could detect the acoustic change from /paap/ to /faap/ in a simpler discrimination task that did not involve learning an object/label mapping (Altvater-Mackensen & Fikkert, 2010). However, discrimination was tested in just one direction and thus did not provide data to address perceptual asymmetry. A third study showed that 25-month-old, but not 18-month-old, Dutch toddlers noticed when familiar words were mispronounced by a change in the initial consonant in both directions, i.e. from /v/ to /b/ or from /b/ to /v/ (Altvater-Mackensen, van der Feest, & Fikkert, 2014).

According to Altvater-Mackensen and colleagues lexical processing is asymmetric in early phonological development either because infants have not yet specified all of the features required to process native lexical forms or they cannot yet fully exploit their perceptual and representational abilities to process all word forms with equal efficiency. However, notably asymmetries have been reported in adult neural processing of non-words that form a stop-nasal manner contrast, /edi/-/eni/, and a coronal-velar place contrast, /edi/-/egi/ (Cornell, Lahiri, &

Eulitz, 2013). These adult findings support the Featurally Under-specified Lexicon (FUL) model, which claims that feature under-specification is an inherent part of the lexical processing architecture throughout development (Lahiri & Reetz, 2010).

To summarize, stop-fricative asymmetries are evident in prior research with adult and toddlers. Different, but not mutually exclusive, factors have been proposed to explain these findings. Asymmetries may have a perceptual basis (related to acoustic-phonetic salience) that later influences the acquisition of lexical representations. Lexical phonological representations may also promote asymmetries independently of perception. Research with young infants can advance our understanding of the mechanisms behind asymmetries. If there is a perceptual asymmetry grounded in the acoustic-phonetic properties of stops and fricatives, then it should be present in young infants who are not yet able to recognize words or produce these sounds. Here we report two perceptual experiments designed to determine whether young infants display a bias in their perception of the stop-fricative contrast /b/-/v/. In Experiment 1, we show directional asymmetries in infant discrimination of /bas/ and /vas/ syllables consistent with our prediction that young infants have a perceptual bias favoring stop consonants. In Experiment 2, we measure infant listening preferences for these same syllables and provide direct evidence for a stop consonant bias. These findings support the view that stops serve as natural referents in the perception of consonant manner.

2. Experiment 1: Discrimination

In Experiment 1, we test the hypothesis that stops are more perceptually salient than fricatives for young infants and we predict that this salience difference will result in a directional asymmetry in a discrimination task. To do so, we tested discrimination of /bas/ and /vas/ syllables in English-learning and French-learning infants. We choose the stop-fricative contrast

/b/-/v/ because this contrast is phonemic in English and in French, and allows us to build on prior research that tested this contrast in toddlers and adults. We tested 5- to 6-month-olds who are not yet babbling, cannot produce consonant sounds, and have little or no ability to recognize words. Thus, a perceptual asymmetry at this age cannot be due to production preferences or lexical representations and would strengthen the conclusion that stops are more perceptually salient than fricatives due to their acoustic-phonetic properties.

If there is a language-independent bias, then we would expect both French and English infants to show the same directional asymmetry even though /b/ and /v/ have different distributional properties in French and English. In regards to type frequency, stops occur more than twice as frequently as fricatives in infant-directed English (Lee, Davis, & MacNeilage, 2010). Also, in spoken English, /b/ occurs much more frequently than /v/ word-initially although similar rates were found for /b/ and /v/ when all word positions are considered (Mines, Hanson, & Shoup, Mines1978). In spoken Canadian French, the opposite is found with /v/ occurring almost twice as often as /b/ in analyses reported by Durand, Laks and Lyche (2002, 2009) (see also Côté, to appear). In addition, /v/ was also more frequent than /b/ word-initially (Malécot, 1974). A similar distributional pattern was evident in European French in both adult speech (Adda-Decker, 2002) and child-directed speech (Le Calvez, 2004).

We tested discrimination in a habituation paradigm using the “look-to-listen” procedure. In this task, infants are first habituated to multiple instances of one syllable (e.g., /bas/) and then are presented test trials with either new instances of the same syllable (i.e., /bas/) or instances of a novel syllable (e.g., /vas/). In this task, infants typically listen longer to the novel category when they can distinguish the categories. We expect to observe this novelty effect when infants are habituated to /vas/ and tested with /bas/ as the novel category. However, if infants have a stop bias, as we have predicted, then the novelty effect will be weaker or absent when infants are habituated to /bas/ and tested with /vas/ as the novel category.

2.1. Methods

2.1.1. Participants

The participants were 48 normally developing full-term 5-6-month-old infants (27 boys; M_{age} = 151 days; range = 130-180 days): 21 monolingual English-learning infants (M_{age} = 150 days; range = 130-170 days) and 27 monolingual French-learning infants (M_{age} = 152 days; range = 130-180 days). All infants had at least 90% exposure to either English or French, as assessed from parental reports of language experience. An additional 11 infants were excluded from analysis due to fussiness/crying (n=5), failure to habituate within the maximum of 12 trials (n=1), parental interference (n=1), equipment or program errors (n=3), experimenter error (n=1).

2.1.2. Stimuli

We chose /bas/ ("bahs") and /vas/ ("vahs") syllables because these units form nonsense syllables that conform to the phonological rules of both English and French. A native female speaker of American English produced multiple tokens of each target syllable (/bas/ and /vas/) in a carrier phrase ("I'll say [target] again."). We excised the tokens from the phrase and analyzed them using Praat (Boersma & Weenink, 2010). To include /bas/ syllables with good instances of /b/ for both languages, we selected tokens in which the initial /b/ was produced with voicing lead (pre-voiced with negative VOT values)[1]. We selected six /bas/ and six /vas/ syllables so that they have matching values with respect to fundamental frequency (f0), vowel intensity, syllable duration, /s/ duration, and formant frequencies at vowel midpoint (vowel quality). This was

[1] In English, VOT (voice onset time) for /b/ includes both pre-voiced and short lag values; in French, VOT for /b/ is consistently pre-voiced.

done to ensure that perceptual differences were not due to extraneous acoustic differences that do not reliably distinguish /b/ and /v/. In the final set the mean duration was 390 ms (SD = 32.4) for /bas/ syllables and 388 ms (SD =34.0) for /vas/ syllables. The selected tokens were identified and rated as good instances of the target syllables by three monolingual English adults and three monolingual French adults.

We conducted additional measures with the selected tokens to examine cues known to differentiate the target consonants, /b/ and /v/. These acoustic measures are reported in Table 1. The duration of initial /b/ and /v/ aperiodic noise was measured by simultaneously referring to waveform and spectrographic displays. As expected, the noise duration was shorter and higher in amplitude for /b/ than for /v/. The amplitude rise time and the amplitude slope were also measured during the initial noise segment. The amplitude rise time is the time interval between the onset of the noise segment and the point where the peak noise intensity is reached. The amplitude slope measure is the rate of change in intensity during the amplitude rise time interval; a higher slope value corresponds to a more rapid rate of amplitude increase. As expected, given the differences in the articulatory dynamics that differentiate these sounds, the stop /b/ had a shorter amplitude rise time and higher amplitude slope compared to the fricative /v/ consonants. We also measured the frequency of the first (F1) and second (F2) formant at the onset of the consonant and at the end of the initial formant transition into the following /a/ vowel. As expected, the F1 frequency at onset was higher for /b/. There was some overlap in F1 change values but the mean value was significantly smaller for /b/ than /v/. No differences in F2 onset or F2 change were observed.

Table 1. Acoustics of the initial consonants for the /bas/-/vas/ contrast. The ** indicates a measure with non-overlapping distribution; *indicates that there is some overlap in the distributions but the mean values are significantly different ($p < .001$). F1 = first formant frequency; F2 = second formant frequency. F1 change is the amount of increase in F1 frequency in the initial 50 ms of the syllable. F2 change is the amount of increase in F2 in the initial 50 ms of the syllable.

	/bas/		/vas/			
	Mean	Range	Mean	Range	Cue status	Difference between means
Noise duration (ms)	9.4	8.3-10.7	37.9	28.6-58.5	**	28.5
Noise amplitude (dB)	73.3	71-75.7	67	64-69.8	**	6.3
Amplitude rise time (ms)	3.8	2.8-6.6	21.2	15.3-30.1	**	17.4
Amplitude rise slope (dB/ms)	0.7	0.6-0.8	0.04	0.01-0.11	**	0.66
F1 at onset (Hz)	762	718-828	613	549-712	**	149
F1 change (Hz)	182	134-220	333	193-406	*	151
F2 at onset (Hz)	1224	1163-1322	1277	1244-1342		53
F2 change (Hz)	67	3-121	58	26-94		9

Using the six tokens of each syllable type (/bas/; vas/), 4 tokens were used to create habituation trial sequences, and 2 tokens were used to create test trial sequences. Each sequence presented /bas/ or /vas/ syllables in a random order constrained so that no token was presented twice in succession. For each syllable type, four different sound files were constructed to present on habituation trials, and two different sound files were created to present on test trials. Each file was 25 seconds long and contained 12 syllables with a 1500ms inter-stimulus-interval. Throughout the task, file presented on each trial was varied to avoid having the same file sequence presented twice in a row.

2.1.3. Procedure

The "look-to-listen" procedure was conducted using *Habit 2000* software program (Cohen, Atkinson, & Chaput, 2000); this procedure has been used extensively to investigate infant speech contrast discrimination (see Polka, Jusczyk, & Rvachew, 1995 for a detailed description). In this task, the infant was seated on the parent's lap facing a TV monitor surrounded by a blue curtain that covered the front and side walls of the test room to remove any visual distractions. Stimuli were played over a loudspeaker positioned below the monitor and behind the curtain, and a video-camera lens protruded through the curtain just below the monitor. The experimenter was located outside the room and viewed the infant over a monitor connected to the video camera. Prior to testing the experimenter was trained extensively with video recordings of similar test sessions until coding of looking behavior was highly reliable (> 95% accuracy).

Each trial started when the infant looked at a visual "attention getter" (a flashing red disk) presented on the TV monitor; once the infant was attentive, a black-and-white checkerboard pattern appeared on the monitor and an audio file began to play. The audio file continued as long as the infant looked at the checkerboard. The experimenter recorded the infant's looking time online by pressing a button when the infant fixated on the checkerboard pattern. The infant's fixation time served as an index of his/her listening time. The experimenter did not know which syllable category was being presented to the infant on each trial. Each trial terminated either when the infant looked away from the checkerboard for more than two consecutive seconds or when a complete 25-second file had played. When either event occurred, the auditory stimulus stopped and the 'attention getter' reappeared to start the next trial. If the infant's looking time failed to reach 1 second on a trial, the program did not count the infant's fixation as an actual look. In this case, the computer reset the trial and replayed the sound file. The Habit program calculated the average looking time of the infant over a sliding three-trial window and tracked the maximum average (3-trial) looking time. To reach the habituation criterion the infant's looking time had to drop below 65% of the maximum 3-trial average (like Polka, Rvachew, & Molnar, 2008) or complete 12 habituation trials.

The test session consisted of four consecutive stages: pre-test, habituation, test, and post-test. During the pre-test and post-test stage (each was a single trial), the audio file was a piece of lively instrumental music. Pre-test and post-test trials were not analyzed. The pre-test trial was a warm-up to show the infant what would happen in the task. The post-test trial was presented to ensure that infants remained engaged throughout the task; infants who failed to listen longer to the post-test trial compared to the last test trial were removed from the analysis.

During the habituation and test stages, the audio file was one of the syllable sequence files (repeating /bas/ or /vas/ syllables, as described above). During habituation, half of the infants heard /bas/ syllables and the other half heard /vas/ syllables. Habituation trials continued until the habituation criterion was met or to a maximum of 12 trials, and then the test phase began immediately. There were four test trials, which consisted of *new-old-new-old* trials, always in this order. The syllables presented on *old* trials were acoustically different but phonetically identical to those presented during the habituation phase; on *new* trials, the syllables belonging to the new phonetic category were presented. In this paradigm, a novelty preference (listening longer to *new* than to *old* trials) is evidence that the infant discriminated the syllable categories. Thus, if discrimination of /b-v/ is asymmetric as predicted, we expected a novelty preference to be present in the /vas/-habituated condition and to be either absent or weaker in the /bas/-habituated condition.

2.2. Results

Three analyses were conducted. First, infant age (in days) was submitted to a Language (English vs. French) × Habituation Group (/bas/-habituated vs. /vas/-habituated) ANOVA. There was no effect of Language [$F(1, 44) = .01$, $p = .92$] or Habituation Group [$F(1, 44) = .002$, $p = .96$] or Language and Habituation Group interaction [$F(1, 44) = .59$, $p = .45$]. These results confirm that there were no age differences across language groups or habituation conditions.

Second, listening time during habituation was analyzed by summing the total listening time across all habituation trials for each infant. These scores were also submitted to a Language (English vs. French) × Habituation Group (/bas/-habituated vs. /vas/-habituated) ANOVA. The main effect of Habituation Group was not significant [$F(1, 44) = 3.51$, $p = .07$], although there was a marginal trend towards longer habituation times for /bas/-habituated compared to /vas/-habituated infants. There was also no main effect of Language [$F(1, 44) = 1.44$, $p = .24$] or Language X Habituation Group interaction [$F(1, 44) = .62$, $p = .44$]. These results show that habituation listening times were comparable across language groups and habituation conditions.

The third analysis focused on listening time during test trials to assess whether the novelty effect was influenced by the habituated consonant. For this analysis, we computed average listening time scores separately for *old* and for *new* test trials for each infant. These data were submitted to an ANOVA with Language (English vs. French) and Habituation Group (/bas/-habituated vs. /vas/-habituated) as between-subject factors and Trial Type (new trials vs. old trials) as a within-subject factor.

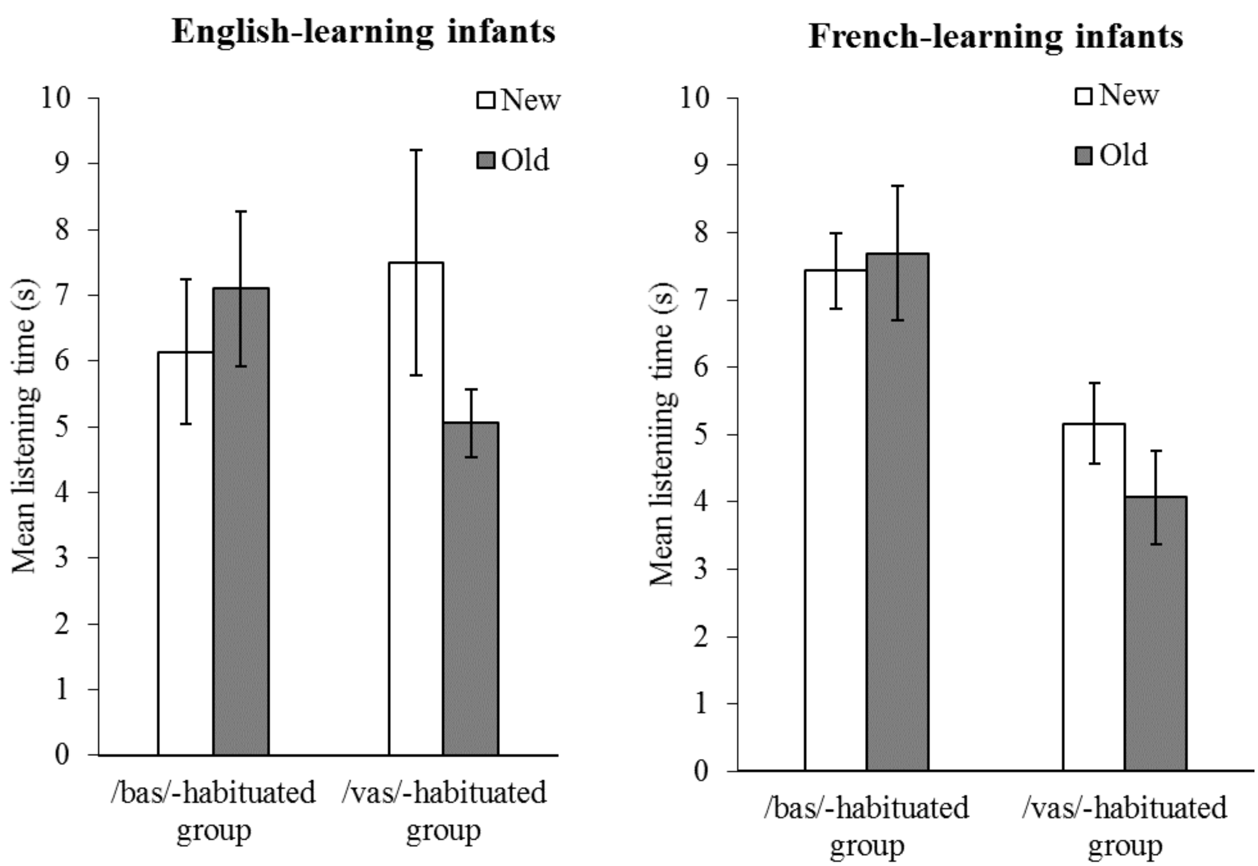

Figure 1. Mean listening time to new vs. old trials for the phonemic /bas/-/vas/ contrast within each language group during the test phase. Error bars represent standard errors.

There was no main effect of Language [$F(1, 44) = .21$, $p = .65$] or Trial Type [$F(1, 44) = 1.16$, $p = .29$]. The main effect of Habituation Group was significant [F(1, 44 = 4.5, $p = .04$, partial η2 = .09], showing that, during the test trials, /bas/-habituated infants ($M = 7.13$ s, $SD = 2.64$) listened more than /vas/-habituated infants ($M = 5.31$s, $SD = 2.76$). As expected, the interaction between Habituation Group and Trial Type was significant [$F(1, 44) = 4.87$, $p = .03$, partial η2 = .10]. This interaction was probed with simple effect analyses. As expected, simple effects of Trial Type revealed a robust novelty preference in the /vas/-habituated group [$t(23) = 2.43$, p (two-tailed) = .023], but no difference in the /bas/-habituated group [$t(23) = -0.72$, p (two-tailed) = .48]. As shown in Figure 2, this pattern is evident in English infants (left panel) and French infants (right panel). No other interactions reached significance: Language × Habituation Group, [$F(1, 44) = 2.81$, $p = .10$]; Language × Trial Type, [$F(1, 44) = .09$, $p = .76$]; Language × Habituation Group × Trial Type, [$F(1, 44) = .89$, $p = .35$].

Asymmetrical performance was also evident at the individual level. In the /vas/ habituated group, 18 out of 24 infants showed a novelty preference (*new* > *old)* on test trials; a binomial test indicates that this proportion is significantly higher than expected by chance ($p = .023$). In the /bas/-habituated group, 13 out of 24 infants showed a novelty preference, a binomial test indicates that this proportion is not significantly different from chance ($p = .839$).

2.3. Discussion

As expected, there was a robust effect of habituation condition on infant performance. Infants habituated to the fricative-initial syllable /vas/ displayed a novelty response to the stop-initial syllable /bas/, whereas infants habituated to the stop-initial syllable /vas/ did not show a novelty response to the fricative-initial syllable /bas/. This pattern was observed in each language group. Further, there appeared to be perceptual clinging to the /bas/ syllable throughout the task.

Infants in the /bas/-habituation group tended to listen longer to their habituating consonant compared to infants in the /vas/-habituation group; and during the test trials, they also maintained a high level of engagement with their habituated consonant compared to the /vas/-habituated infants. Thus, asymmetrical performance and other listening patterns point to perceptual bias favoring the stop consonant /b/. We investigated this bias directly in Experiment 2.

3. Experiment 2: Listening Preference

The task implemented in Experiment 1 provides a way to index infant discrimination indirectly through habituation and listening preference. During habituation, the infant engages with one stimulus until their interest declines; at this point, infants typically show increased interest when they encounter a stimulus that they perceive to be new. Thus, a reliable novelty response provides evidence that the old (habituated) and new stimuli are discriminated. However, and importantly, failure to display the expected novelty response is equivocal; this outcome could mean that infants cannot discriminate the old and new stimuli, or that old and new stimuli are discriminated but infants are not motivated to explore the new stimulus more than the old one. A third possibility is that infants discriminate the old and new stimuli but they have a preference for one of the stimuli over the other which continues to influence their listening behavior during the test trials. In the latter case, an asymmetric pattern may emerge (e.g., Shi & Werker, 2001). Specifically, if there is a preference for stimulus A over stimulus B, infants may show a novelty response when habituated to B but it may be weaker or absent when habituated to A. If such a bias underlies the asymmetric performance in Experiment 1, then we should observe it in a simple listening preference task. The purpose of Experiment 2 was to test directly whether young infants prefer to listen to stop-initial syllables over the fricative-initial syllables. English- and French-learning 5-6-month-olds were tested in a sequential preferential looking paradigm using the /bas/ and /vas/ stimuli from Experiment 1. We expected infants to show a stop bias, i.e., to listen longer to /bas/ than /vas/ syllables.

3.1. Methods

3.1.1. Participants

The participants were twenty normally developing full-term 5- to 6-month-olds (9 boys; M_{age} = 154 days; range = 133-182 days) including 8 monolingual English-learning infants and 12 monolingual French-learning infants. An additional 8 infants were tested and then excluded from data analysis due to crying/ fussiness (n=4), equipment failure (n=2) or experimenter error (n=2). All the infants were exposed to either English or French at least 90% of the time according to parental reports.

3.1.2. Stimuli

The stimuli were the same six tokens of each syllable type used for Experiment 1. On each trial, a sequence of either /bas/ or /vas/ syllables was presented. All six tokens were used to create these sequences. In each sequence, the order of tokens was random with the constraint that no token occurred twice in a row. Each trial sequence was 25 seconds long; syllables were spaced with a 1500-ms ISI.

3.1.3. Procedure

The test setup was the same as in Experiment 1. In this experiment, a sequential preferential looking procedure was employed to directly assess whether infants prefer to listen to a stop consonant over a fricative consonant. Research has shown that this relatively simple procedure is successful in ascertaining the presence or absence of preference for one type of stimulus over another in infants (e.g., Polka *et al.*, 2008; Shi & Werker, 2001; Vouloumanos & Gelfand, 2013; Vouloumanos & Werker, 2004). The test session was identical to Experiment 1 except that there was no habituation stage and the test stage consisted of 12 trials in which the syllable type presented was alternated across trials (6 /vas/ trials; 6/bas/ trials). The test session included one pre-test trial, twelve test trials, and one post-test trial. Half of the infants were assigned to a /bas/-first listening condition; their test trial order was /bas/ /vas/ /bas/ /vas/ etc. The other half were assigned to a /vas/-first listening condition; their test trial order was /vas/ /bas/ /vas/ /bas/ etc. Looking time was recorded on each trial to index listening time. As in Experiment 1, the pre-test and post-test trial were not analyzed; the pre-test trial was a warm-up trial and the post-test trial identified infants who were no longer attentive at the end of the task.

3.2. Results

For each infant, a mean listening time score was computed across the six trials of each syllable type (/bas/; /vas/). The scores were submitted to a mixed ANOVA with Language (English vs. French) as a between-subjects factor and Syllable Type (/bas/ vs. /vas/) as a within-subjects factor. The main effect of Language was not significant [$F(1, 18) = .68$, $p = .42$] and there was no interaction between Language and Syllable Type [$F(1, 18) = .14$, $p = .72$]. As expected, there was a robust main effect of Syllable Type [$F(1, 18) = 17.46$, , $p = .001$, partial $\eta^2 = .49$], showing that infants listened significantly longer to /bas/ trials than /vas/ trials [$M_{/bas/} = 11.41$ s, $SD = 4.39$ vs. $M_{/vas/} = 9.20$ s, $SD = 3.97$]. As shown in Figure 2, this pattern was evident in both English infants (left panel) and French infants (right panel).

For each infant mean listening time scores were re-calculated using only 5 trials, with the first trial (for each syllable type) removed. The ANOVA was repeated using the 5-trial scores. This was done to confirm that the preference is robust and not based on a transient initial reaction to each syllable type. As in the first analysis, there was no effect of Language [$F(1, 18) = .92$, $p = .35$] or Language × Syllable Type interaction [$F(1, 18) = .001$, $p = .98$]. However, as expected, the effect of Syllable type was highly significant [$F(1, 18) = 9.86$, $p = .006$, partial $\eta^2 = .3$], showing that infant preference was not due to the initial response to each syllable type.

A robust preference was also observed in individual performance. For both 5-trial and 6-trial scores, sixteen out of the 20 infants listened longer to /bas/ syllables. Binomial test indicates that this proportion is significantly higher than expected by chance ($p = .012$)

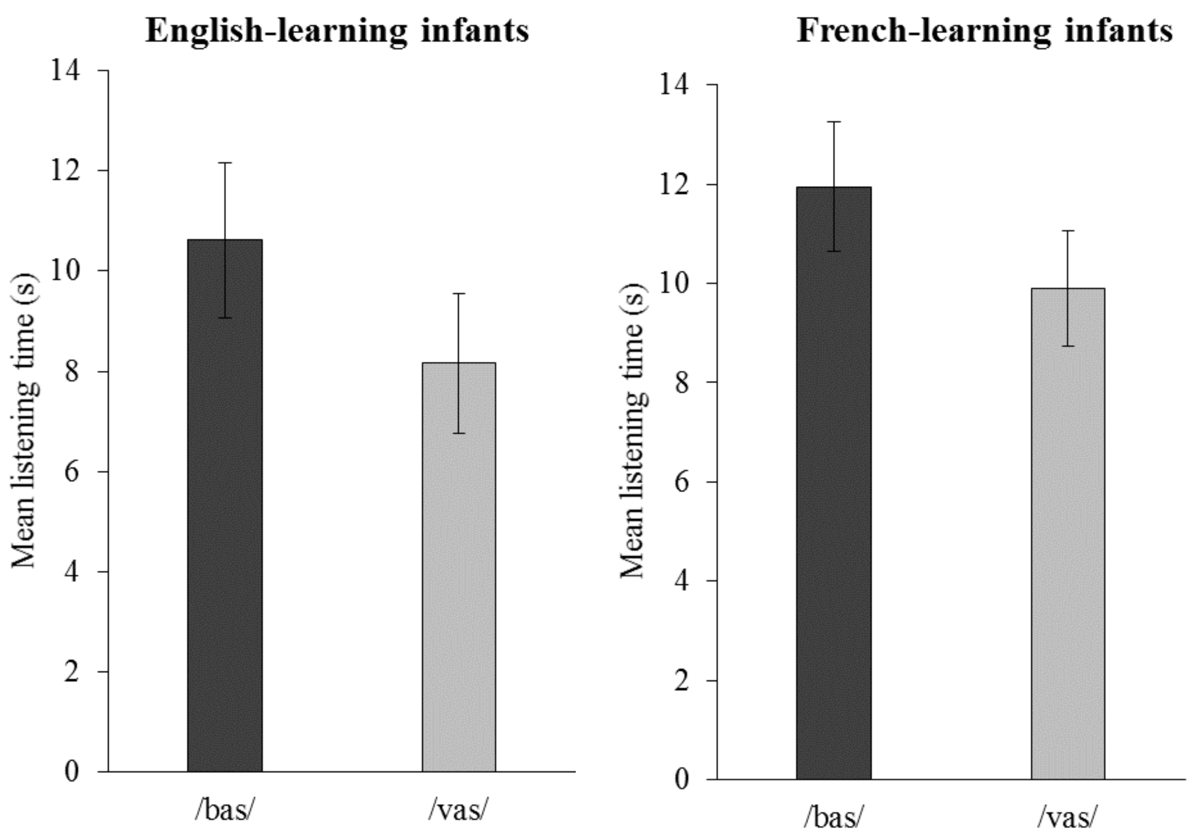

Figure 2. Mean listening times (6-trial scores) to /bas/ and /vas/ syllables within each language group. Error bars represent standard errors.

3.3. Discussion

As predicted, infants showed a robust listening preference for the stop-initial syllable /bas/. This pattern was evident in both English-learning and French-learning infants and was observed in 80% of the infants tested. Whereas directional asymmetries emerge from between-group comparisons in Experiment 1, these listening preference measures provide direct evidence of a perceptual bias within the same infants. Importantly, these finding show that infants can discriminate /bas/ and /vas/ syllables but were unable to reveal this in both habituation conditions in Experiment 1. A perceptual bias favoring /bas/ interfered with measuring a novelty response to the less-preferred /vas/ syllables.

4. General Discussion

The present findings provide clear evidence that young infants display a perceptual bias favoring /b/ over /v/. In Experiment 1, discrimination of the phonemic /bas/-/vas/ contrast was asymmetric in both English-learning and French-learning 5-to 6-month-olds. Infants habituated to fricative-initial /vas/ syllables responded to a novel /bas/ syllable, whereas infants habituated to stop-initial /bas/ syllables did not respond to a novel /vas/ syllable. In Experiment 2, listening preference was measured directly, confirming a perceptual bias favoring the stop-initial /bas/ syllables over the fricative-initial /vas/ syllables within the same infants. We observed this bias in young infants who cannot yet produce stop or fricative sounds and in both language groups, even though /b/ occurs more frequently than /v/ in English but the reverse distributional pattern occurs in French. Thus, an account based on lexical representations, speech production patterns, or native language attunement is not tenable. Overall, these findings uncover a language-independent pattern, consistent with the idea that young infants display a perceptual bias favoring stop consonants.

In previous studies, group comparisons revealed directional asymmetries pointing to language-independent biases in consonant perception; it is unknown whether these biases also drive a listening preference. This is the first direct evidence that infants selectively attend more to some consonants than others, bolstering the likelihood that perceptual biases influence infant listening behavior outside the laboratory. The present findings also align with several studies with adults (outlined in the introduction) which also reveal asymmetries, supporting the claim that stops are more perceptually salient than fricatives.

This study was motivated to explore the feasibility of extending the principles of the NRV framework to consonant perception. We will discuss how the present findings address this issue and then consider the broader implications for our understanding of infant speech perception development.

A core principle of the NRV framework is that young infants display language-universal perceptual biases grounded in acoustic-phonetic patterns that are shaped by speech production. Perceptual biases emerge because some speech sounds are more perceptually salient than others; salience differences may be indexed directly in listening preference measures, but they are also exposed as a directional asymmetry in discrimination. A link between perceptual salience and discrimination asymmetries has been observed in other domains (e.g. with letters, a O to Q change is easier to detect than a Q to O change, Gilmore, Hersh, Caramazza, & Griffin, 1979). This appears to reflect a general cognitive bias in which adding to or enhancing information is more attention-grabbing (increasing discriminability) than removing or reducing stimulus information (decreasing discriminability). It is beyond the scope of this paper to review the proposed mechanisms behind this bias (see Polka & Bohn, 2003; also Chang, Plauche & Ohala, 2001). In NRV, we propose that application of a general cognitive bias to phonetic patterns leads to universal asymmetries in speech perception.

Applying NRV principles to explain the data reported here, we propose that stops serve as natural referents in infant perception of consonant manner-of-articulation, and that this perceptual bias is a natural consequence of the acoustic-phonetic properties of stops and fricatives. To firmly establish this, we must isolate a specific acoustic-phonetic property that grounds this bias. Amplitude rise time (ART) is an obvious and compelling candidate. ART is a reliable cue distinguishing stops and fricatives. Moreover, perceptual advantages linked to ART as a prominent cue in consonant perception are well established in the literature. It is generally agreed that rapid, abrupt shifts in signal amplitude enhances auditory contrast and are easy to detect (e.g., Delgutte & Kiang 1984a, 1984b). Stevens and Keyser (1989) suggested that [- continuant] segments with abrupt onsets give rise to perceptual advantage over [+ continuant] segments with gradual onsets. The signal dynamics modulating auditory contrast for consonants are often gauged by measures of amplitude rise time, i.e., a shorter rise time indexes a more rapid, abrupt amplitude shift. Supporting this, Gage, Poeppel, Roberts and Hickok (1998) showed that stops with short rise times are perceptually more distinct than non-stop consonants with longer rise times. This was evident in a MEG study conducted with English adults. They found that M100 latencies were significantly shorter and higher in amplitude in response to stop-initial words (starting with /b, p, d, t, g, k/) compared to fricative-, liquid- and nasal-initial words (starting with /f, s, l, r, m/). Recently, Easwar, Glista, Purcell, and Scollie (2012) showed that rise time differences also affect perception of the same consonantal segment. Measuring cortical auditory evoked potential responses, they found that “sh” sounds with a shorter rise time (27.6 ms) elicited larger N1-P2 amplitudes with shorter latencies compared to stimuli “sh” sounds with longer rise times (113.2 ms). Similar findings have been reported using tonal stimuli that vary in rise time (Thomson, Goswami, & Baldeweg, 2009; see also Kodera, Hink, Yamada, & Suzuki, 1979; Onishi & Davis, 1968; Skinner & Jones, 1968). Recent research with language-impaired children also suggests that sensitivity to ART plays an important role in phonological development (Cumming, Wilson, & Goswami, 2015).

As expected, acoustic analyses of the stop /b/ and fricative /v/ in the present study reveal differences in amplitude rise time and slope, as well as noise duration and amplitude. As expected, the shift from consonant closure/constriction to open vocal tract in production of these consonants is acoustically marked by a rapid, sharp rise in amplitude for /b/ compared to a longer, shallow amplitude rise for /v/. The amplitude rise time cue was also more stable (less variable) for the /bas/ syllables than for the /vas/ syllables. This variability reflects natural inherent differences in stop and fricative production, given that stop closure involves an abrupt, blunt closure gesture compared to the continuous, nuanced constriction gesture involved in fricative production. The close correspondence between amplitude rise-time differences and perceptual biases in the present study is compelling. Never-the-less, further research manipulating ART is needed to confirm that this cue drives asymmetry.[2] Moreover, we tested infants who have had 5-6 months of listening experience with their native language. Research with younger infants will also reveal whether this bias is innate or emerges with exposure to speech in the early months of life.

Another core principle of NRV framework is that language experience (via perception and production) will reinforce or modulate initial universal biases, as needed, to optimize access

[2] Conceivably, the longer noise duration for /v/ compared to /b/ may be taken as evidence that /v/ may be more salient than /b/. However, this measure reflects only one static portion of the target consonants, whereas amplitude rise time and slope measures capture the onset dynamics properties of the consonant articulation that are associated with auditory contrast.

to native language categories. As such, NRV is compatible with other models such as NLM (Kuhl, 1991)[3] which focus on explaining native language attunement. For native contrasts, the NRV framework predicts that asymmetries observed in early infancy will fade as infants tune into their native language structure. For non-native contrasts, universal asymmetries may be maintained, disappear, or even reverse depending on how the contrasting sounds align with native categories.

The NRV principles outlined above can provide a coherent account for consonant asymmetries reported in the current literature. Considering infant asymmetric perception of /r/-/l/ reported by Kuhl *et al.* (2006), the converging second and third formant frequencies (F2 and F3) at the onset of /r/ may be more perceptually salient than the more dispersed formant frequencies of /l/. Thus, the same acoustic-phonetic property proposed to drive asymmetries in vowel perception (focalization) within the NRV framework may also account for the asymmetry observed for /r/ - /l/; notably, the /r/ and /l/ consonants are also classified as semi-vowels. To explain the labial-coronal (ompa-onta) place asymmetry, Tsuji *et al.* (2015) offered a salience-based interpretation grounded in acoustic-phonetic properties of the stop release bursts based on research by Chang, Plauche and Ohala (2001); this account converges conceptually with NRV principles. The data reported by Segal *et al.* (2016) suggests that discrimination of the /b/-/p/ voicing contrast is unbiased (symmetric) in young infants but become asymmetric once language attunement is clearly underway. These data align well with NLM and are also consistent with NRV principles.

Notably, we do not claim that universal perceptual asymmetries will be observed for all consonant contrasts. The present data suggest that universal biases emerge for phonetic features

[3] The NRV framework has not specified a precise developmental timeframe for these language effects.

distinguished by robust, dynamic amplitude cues that correspond to global vocal tract movements. This also aligns with the developmental progression in perceptual weighting of speech proposed and supported by Nittrouer (2002). However, we draw these broad conclusions with caution given the very limited data pertaining to asymmetries in infant consonant perception. Further research is needed to confirm that the bias favoring /b/ over /v/ is found in other syllable/word contexts and with other stop-fricative contrasts. It will also be informative to test predictions stemming from the NRV principles regarding developmental patterns and other consonant manner contrasts. With respect to the /b/-/v/ contrast, we predict that the perceptual asymmetry will decrease for both English and French perceivers; if native language distributional patterns also play a role, this decline will be earlier or larger for French perceivers since /v/ occurs more often than /b/. The fricative-affricate manner contrast “sha” - “cha” (/ʃ/ - /tʃ/) also provides an informative test case given that ART is also a reliable cue for this contrast (ART is shorter for “cha” than “sha”). In English, “sha” occurs more frequently than “cha”; and in French, this is a non-native contrast with “sha” occurring frequently and “cha” absent. Moreover, affricates are less common than fricatives across languages, thus for this contrast, perceptual salience based on ART and the status of these sounds across language are not aligned. If the proposed extension of NRV principles to consonant manner is correct, we should find a language-universal directional asymmetry in young French and English infants, i.e. easier discrimination in the “sha” to “cha” direction. Following native language attunement, this asymmetry should decrease for English infants and reverse for French infants.

As outlined in the introduction, the language-independent biases documented here and in Tsuji *et al.*, (2015) align with asymmetries observed in word processing research with toddlers and neural processing research with adults. Clearly, these early perceptual biases are a precursor rather than a product of lexical development. There appears to be a developmental continuity across the first and second year of life, with initial acoustic-phonetic biases feeding into early lexical processing and eventually becoming consolidated in the mature phonological architecture. As such, our interpretation of early perceptual asymmetries is compatible with accounts that claim under-specification to be innate. The current literatures reveals that different processes operate across development - acoustic-phonetic biases, native-language tuning, and phonological representations - and each contributes at different stages (and to varied degrees) to the acquisition of functional and efficient language processing skills. The NRV framework focuses on providing a grounded and mechanistic account of speech perception development that can enrich our understanding of spoken language acquisition and may inform our understanding of sound change. However, we need further research to establish the merits and limitations of this framework.

Taking a broader view, the present findings have important implications for our conceptual understanding of infant speech perception. We know that infants are born with perceptual capacities that allow them to discriminate all, or nearly all, linguistically relevant phonetic contrasts. The present findings confirm that there is another piece to this story – not all phonetic units have equal perceptual salience in early infancy. Perceptual asymmetries provide a way to tag salience differentials and offer a new window into the speech perception processes. Lahiri and Reetz (2010) have argued that, in mature perceivers, featurally under-specified representations afford processing efficiency and enable us to deal with the complex variability that we encounter in fluent speech. Similarly, in outlining the NRV framework, Polka and Bohn (2011) speculate that perceptual biases may ease the workload during acquisition. They proposed that the bias favoring natural referent vowels provides a default perceptual scaffold that infants can use to discover additional vowel categories in their native language. This is a versatile framework given that focal vowels define a vowel space that is used in every language (Maddieson, 1984). Having extra salient referent sounds may encourage infants to build or refine their speech processing skills in a step-wise fashion, focusing on some phonetic categories before others, rather than honing recognition of all categories simultaneously. This aligns with work highlighting the importance of processing constraints in other aspects of language acquisition (Newport, 1990).

In summary, the present study indicates that the stop /b/ is more perceptually salient than the fricative /v/ for young infants. These initial findings suggest that stops may serve as natural referents in infant consonant manner perception and point to amplitude rise time as the acoustic-phonetic cue grounding this bias. These data encourage further efforts to extend the NRV framework to consonants. The consonant asymmetry document here also confirms that we need to reform our view of infant speech perception to accommodate the fact that perceptual salience make some speech sounds more prominent than others in the early months of life. It is clear that infant perception affords access to an initial phonetic landscape with an irregular terrain - for consonants as well as vowels. In future research, our challenges are to uncover the universal biases that define this terrain, identify their origin(s), and understand their role in speech perception development.

Acknowledgments

We thank all participating infants and their parents, Raphaëlle Curis and Sunah Jeon for assistance with data collection, and AJ Orena and Jeff Lamontagne for help with the final manuscript. This research was supported by an NSERC grant (OPG105397) to Dr. Linda Polka. This study was completed as part of the first author's doctoral dissertation.

References

Adda-Decker, M. (2006). De la reconnaissance automatique de la parole à l'analyse linguistique de corpus oraux. *JEP*, 389-400.

Altvater-Mackensen, N., & Fikkert, P. (2010). The acquisition of the stop-fricative contrast in perception and production, *Lingua, 120*, 1898–1909.

Altvater-Mackensen, N., van der Feest, S. V., & Fikkert, P. (2014). Asymmetries in Early Word Recognition: The Case of Stops and Fricatives. *Language Learning and Development*, *10*(2), 149-178.

Best, C. T., McRoberts, G. W., & Sithole, N. M. (1988). Examination of perceptual reorganization for nonnative speech contrasts: Zulu click discrimination by English speaking adults and infants. *Journal of Experimental Psychology: Human perception and performance*, *14*(3), 345.

Boersma, P., & Weenink, D. (2010). Praat: doing phonetics by computers [Computer program]. Version 5.1.44.

Chang, S., Plauché, M. C., & Ohala, J. J. (2001). Markedness and consonant confusion asymmetries. *The role of speech perception in phonology*, 79-101.

Cohen, L. B., Atkinson, D. J., & Chaput, H. H. (2000). *Habit 2000: A new program for testing infant perception and cognition* (Version 2.2.5c). [Computer software]. Austin, the University of Texas.

Cornell, S. A., Lahiri, A., & Eulitz, C. (2013). Inequality across consonantal contrasts in speech perception: Evidence from mismatch negativity. *Journal of Experimental Psychology: Human Perception and Performance, 39*(3), 757.

Côté, M. H. (*to appear*). Du projet PFC à la géolinguistique du français laurentien. In Jacques Durand, Gjert Kristoffersen & Bernard Laks (Ed.) *Phonologie du français: de la norme aux variétés périphériques*. Presses de l'Université de Paris Ouest.

Cumming, R., Wilson, A., & Goswami, U. (2015). Basic auditory processing and sensitivity to prosodic structure in children with specific language impairments: a new look at a perceptual hypothesis. *Frontiers in psychology, 6*.

Delgutte, B., & Kiang, N. Y. (1984a). Speech coding in the auditory nerve: III. Voiceless fricative consonants. *The Journal of the Acoustical Society of America*, *75*(3), 887-896.

Delgutte, B., & Kiang, N. Y. (1984b). Speech coding in the auditory nerve: IV. Sounds with consonant-like dynamic characteristics. *The Journal of the Acoustical Society of America, 75*(3), 897-907.

Denes, P. B. (1963). On the statistics of spoken English. *The Journal of the Acoustical Society of America*, *35*(6), 892-904.

Dijkstra, N., & Fikkert, P. (2011). Universal Constraints on the Discrimination of Place of Articulation? Asymmetries in the Discrimination of 'paan'and 'taan'by 6-month-old Dutch Infants. In N. Danis, K. Mesh, & H. Sung (Eds.). *Proceedings of the Boston University conference on language development* (Vol. 35, pp. 170–182).

Durand, J., Laks, B., & Lyche, C. (2002). La phonologie du français contemporain : usages, variétés et structure. In C. Pusch & W. Raible. (Ed.), *Romanistische Korpuslinguistik-Korpora und Gesprochene Sprache/Romance Corpus Linguistics- Corpora and Spoken Language* (pp. 93-106). Tübingen: Gunter Narr Verlag.

Durand, J., Laks, B., & Lyche, C. (2009). Le projet PFC : une source de données primaires structurées. In Jacques Durand, Bernard Laks, & Chantale Lyche (Ed.) *Phonologie,* *variation et accents du français* (pp. 19-61). Paris: Hermès.

Easwar, V., Glista, D., Purcell, D. W., & Scollie, S. D. (2012). The effect of stimulus choice on cortical auditory evoked potentials (CAEP): consideration of speech segment positioning within naturally produced speech. *International journal of audiology*, *51*(12), 926-931.

Gage, N., Poeppel, D., Roberts, T. P., & Hickok, G. (1998). Auditory evoked M100 reflects onset acoustics of speech sounds. *Brain research*, *814*(1), 236-239.

Gilmore, G. C., Hersh, H., Caramazza, A., & Griffin, J. (1979). Multidimensional letter similarity derived from recognition errors. *Perception & Psychophysics*, *25*(5), 425-431.

Kodera, K., Hink, R. F., Yamada, O., & Suzuki, J. I. (1979). Effects of rise time on simultaneously recorded auditory-evoked potentials from the early, middle and late ranges. *International Journal of Audiology*, *18*(5), 395-402.

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. (1992). Infants’ perception and representation of speech: Development of a new theory. In J. J. Ohala, T. M. Nearey, B. L. Derwing, M. M. Hodge, & G. E. Wiebe (Eds.), *Proceedings of the International Conference on Spoken Language Processing* (pp. 449-456).

Kuhl, P. K. (1994). Learning and representation in speech and language. *Current Opinion in Neurobiology*, *4*, 812-822.

Kuhl, P. K., Stevens, E., Hayashi, A., Deguchi, T., Kiritani, S., & Iverson, P. (2006). Infants show facilitation for native language phonetic perception between 6 and 12 months.

Developmental Science, *9*, 13–21.

Ladefoged, P. (1993). *A Course in Phonetics (Third Edition)*. New York: Harcourt Brace Jovanovich.

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

Lee, S. A. S., Davis, B., & MacNeilage, P. (2010). Universal production patterns and ambient language influences in babbling: A cross-linguistic study of Korean-and English-learning infants*. *Journal of Child Language*, *37*(02), 293-318.

Le Calvez, R. (2004). *Modélisation de l'acquisition des catégories phonémiques*. Master's thesis, EHESS, Paris.

Maddieson, I. (1984). *Patterns of sounds*. Cambridge: Cambridge University Press.

Malécot, A. (1974). Frequency of occurrence of French phonemes and consonant clusters. *Phonetica*, *29*(3), 158-170.

Miller, J. L., & Eimas, P. D. (1996). Internal structure of voicing categories in early infancy. *Perception & psychophysics*, *58*(8), 1157-1167.

Mines, M. A., Hanson, B. F., & Shoup, J. E. (1978). Frequency of occurrence of phonemes in conversational English. *Language and speech*, *21*(3), 221-241.

Nam, Y. (2015). *The role of acoustic-phonetic bias in consonant manner perception*. Unpublished doctoral dissertation, McGill University.

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

Newport, E. L. (1990). Maturational constraints on language learning. *Cognitive science*, *14*(1), 11-28.

Nittrouer, S. (2002). Learning to perceive speech: How fricative perception changes, and how it stays the same. *The Journal of the Acoustical Society of America*, *112*(2), 711-719.

Onishi, S., & Davis, H. (1968). Effects of duration and rise time of tone bursts on evoked V potentials. *The Journal of the Acoustical Society of America*, *44*(2), 582-591.

Polka, L., & Bohn, O. S. (2003). Asymmetries in vowel perception. *Speech Communication*, *41*(1), 221-231.

Polka, L., & Bohn, O.,-S. (2011). Natural Referent Vowel (NRV) framework: An emerging view of early phonetic development. *Journal of Phonetics, 39*, 467-478.

Polka, L., Jusczyk, P. W., & Rvachew, S. (1995). Methods for studying speech perception in infants and children. In W. Strange (Ed.), *Speech perception and linguistic experience: Theoretical and methodological issues in cross-language speech research* (pp. 49–89). Timonium, MD: York Press.

Polka, L., Rvachew, S., & Molnar, M. (2008). The effect of a nonspeech distractor on infant speech perception. *Cognition*, *13*, 421–439.

Pons, F., Albareda-Castellot, B., & Sebastián-Gallés, N. (2012). The interplay between input and initial biases: Asymmetries in vowel perception during the first year of life. *Child development*, *83*(3), 965-976.

202, 321–435.

Polka, L., & Werker, J. F. (1994). Developmental changes in perception of nonnative vowel contrasts. *Journal of Experimental Psychology: Human Perception and Performance, 202*, 321–435.

Reetz, H., & Jongman, A. (2011). *Phonetics: Transcription, production, acoustics, and perception* (Vol. 34). John Wiley & Sons.

Repp, B. H., Liberman, A. M., Eccardt, T., & Pesetsky, D. (1978). Perceptual integration of acoustic cues for stop, fricative, and affricate manner. *Journal of Experimental Psychology: Human Perception and Performance*, *4*, 621–637.

Saffran, J. R., Werker, J. F., & Werner, L. A. (2006). The infant's auditory world: Hearing, speech, and the beginnings of language. Handbook of child psychology.

Schwartz, J. L., Abry, C., Boë, L. J., Ménard, L., & Vallée, N. (2005). Asymmetries in vowel perception, in the context of the Dispersion-Focalisation Theory. *Speech Communication*, *45*, 425-434.

Segal, O., Hejli-Assi, S., & Kishon-Rabin, L. (2016). The effect of listening experience on the discrimination of/ba/and/pa/in Hebrew-learning and Arabic-learning infants. *Infant Behavior and Development*, *42*, 86-99.

Shi, R., & Werker, J. F. (2001). Six-month-old infants' preference for lexical words. *Psychological Science*, *12*(1), 70-75.

Skinner, P. H., & Jones, H. C. (1968). Effects of signal duration and rise time on the auditory evoked potential. *Journal of Speech, Language, and Hearing Research*, *11*(2), 301-306.

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

Stevens, K. N., & Keyser, S. J. (1989). Primary Features and their Enhancement in Consonants. *Language*, *65*, 81-106.

Thomson, J. M., Goswami, U., & Baldeweg, T. (2009). The ERP signature of sound rise time changes. *Brain research*, *1254*, 74-83.

Tsuji, S., Mazuka, R., Cristia, A., & Fikkert, P. (2015). Even at 4 months, a labial is a good enough coronal, but not vice versa. *Cognition*, *134*, 252-256.

Tsushima, T., Shiraki, S., Yoshida, K., & Sasaki, M. (2003). On stimulus order effects in discrimination of nonnative consonant contrasts. *Acoustical Science and Technology***,** *24*(6), 410-412.

Tsushima, T., Shiraki, S., Yoshida, K., & Sasaki, M. (2005). Stimulus order effects in discrimination of a nonnative consonant contrast, English /b-v/, by Japanese listeners in the AX discrimination procedure. *Paper presented at the First Acoustical Society of America Workshop on L2 Speech Learning*, Vancouver, Canada.

Tyler, M. D., Best, C. T., Faber, A., & Levitt, A. G. (2014). Perceptual assimilation and discrimination of non-native vowel contrasts. *Phonetica*, *71*(1), 4-21.

Vouloumanos, A., & Gelfand, H. M. (2013). Infant perception of atypical speech signals. *Developmental psychology*, *49*(5), 815.

Vouloumanos, A., & Werker, J. F. (2004). Tuned to the signal: the privileged status of speech for young infants. *Developmental science*, *7*(3), 270-276.

Werker, J. F., Cohen, L. B., Lloyd, V. L., Casasola, M., & Stager, C. L. (1998). Acquisition of word–object associations by 14-month-old infants. *Developmental psychology*, *34*(6), 1289.

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, 49–63.

Yamaguchi, N. (2007). Markedness, Frequency: Can We Predict the Order of Acquisition of Consonants?. In *LingO 2007, the second Oxford Postgraduate Conference in Linguistics* (pp. 236-243).

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.

Zhang, Y., Imada, T., Kawakatsu, M., & Kuhl, P.K., (2006). Neural basis of perceptual asymmetry for the stimulus order effect: A cross-language MEG study. *Poster presented at the Biomag 2006*, Vancouver, Canada.