Journal of Phonetics 45 (2014) 43–51

Contents lists available at ScienceDirect

Journal of Phonetics

journal homepage: www.elsevier.com/locate/phonetics

Acoustic correlates of allophonic versus phonemic dimensions in monolingual and bilingual infants' input

Amanda Seidl [a], Kristine H. Onishi [b], Golnoush Alamian [b,c], Alejandrina Cristia [d,*]

[a] Purdue University, West Lafayette, IN 47907, USA
[b] McGill University, Montreal, QC, Canada H3A 1B1
[c] University of British Columbia, Vancouver, BC, Canada V6T 1Z4
[d] Laboratoire de Sciences Cognitives et Psycholinguistique, CNRS, DEC-IEC-ENS-PSL*, EHESS Paris 75005, France

A R T I C L E I N F O

Article history:
Received 17 September 2013
Received in revised form
12 February 2014
Accepted 5 March 2014
Available online 12 April 2014

Keywords:
Allophones
Bilingualism
Infancy
Vowel nasalization
Vowel tenseness

A B S T R A C T

Allophones are diverse phonetic instantiations of a single underlying sound category. As such, they pose a peculiar problem for infant language learners: These variants occur in the ambient language, but they are not used to encode lexical contrasts. Infants' sensitivity to sounds varying along allophonic dimensions declines by 11 months of age, suggesting that there must be information to phonological status available to pre-lexical infants. The present work tests one specific type of information: acoustic implementation. It was hypothesized that the acoustic distance between two vowel categories is smaller when the dimension along which the two vowels differ is allophonic (e.g., vowel nasality in American English, vowel tenseness in Quebec French) compared to when it is phonemic (e.g., vowel tenseness in American English, vowel nasality in Quebec French). Monolingual mothers speaking either English or French and bilingual mothers speaking both languages were recorded while they described objects to their 11-month-olds. Results provided weak support for the main hypothesis.

1. Introduction

By the end of the first year of life, both monolingual and bilingual infants have begun to lay down the foundations for their native language(s). One of the key landmarks of this process involves perceptual attunement, the process by which infants allocate their attention more effectively by reducing sensitivity to non-native sound distinctions and increasing it for native ones (see Tsuji & Cristia, 2014, for a review). Recent work suggests that the decline even applies to pairs of sounds that are present in the input, but which differ on a dimension that is allophonic in the native language. Indeed, English-learning 11-month-olds fail to learn a pattern involving vowel nasality, a dimension that is allophonic in the ambient language, whereas both 4-month-old English learners and French-learning 11-month-olds succeed (Seidl, Cristia, Bernard, & Onishi, 2009). How might they achieve this early feat? Here, we evaluate a heretofore ignored potential contributor: the acoustic separation of sound categories. Specifically, we measure acoustic separation as a function of phonemic or allophonic status in infants' input.

Although there is increasing understanding of more intermediate cases (Hall, 2013), the classical description of allophony involves two clear scenarios. A pair of sounds are allophones of the same phoneme if they are in complementary distribution (for example, one variant occurs in one phonological environment, the other variant in all other environments) or in free variation (either is acceptable in a given context, and no meaning change ensues from swapping them; Kenstowicz, 1994). Children could use at least three strategies to learn about phonological status. The first two have been studied in previous research, which we summarize briefly before turning to the one focused in the present work.

The first and obvious strategy involves *lexical bootstrapping*. Hearing [tɛlɪtʌbi] and [tɛlɪd̰ʌbi] spoken in reference to the same Teletubby toy could indicate that the voiceless [t] and glottalized [d̰] alveolar stops are in free variation (Foulkes, Docherty, & Watt, 2005). Recent modeling work suggests that a similar strategy of mapping word forms to meanings or referents improves the determination of allophonic status even when applied pre-lexically, using 'pseudowords' (i.e., consistent sequences of sounds that may or may not correspond to actual words of the language; Fourtassi & Dupoux, in preparation; Martin, Peperkamp, & Dupoux, 2013). According to these proposals, infants would not need to have large semantic vocabularies to profit from lexical-type cues to allophony. As of yet, there is no experimental work directly supporting infants' *use* of such cues.

Second, *distributional information* is extremely informative to determine allophones in complementary distribution. The possibility that complementary distribution would affect perception is strengthened by evidence of reduced sensitivity in 12-month-olds after experimental exposure

* Corresponding author. Tel.: +33 1 44 32 26 23; fax: +33 1 44 32 26 30.
E-mail addresses: aseidl@purdue.edu (A. Seidl), kris.onishi@mcgill.ca (K.H. Onishi), g.alamian@alumni.ubc.ca (G. Alamian), alecristia@gmail.com (A. Cristia).

0095-4470/$ - see front matter
http://dx.doi.org/10.1016/j.wocn.2014.03.004

to sounds in complementary distribution (White, Peperkamp, Kirk, & Morgan, 2008). However, modeling work demonstrates that distributional information alone is not sufficient to establish which sounds are allophones of the same phoneme when complementary distribution affects whole sound classes. To take the example of vowel nasalization in English, a learner with access to distributional information alone cannot decide whether [i] and [ĩ] map onto the same phoneme and [i] and [ũ] to different ones, or whether [i] and [ũ] are allophones instead. Peperkamp, LeCalvez, Nadal, and Dupoux (2006) document precisely this behavior in a computational model having access only to complementary distribution. They also demonstrate that performance is improved when the learner constrained hypothetical allophones on the basis of phonetic similarity. In the next subsection, we explain how acoustic information on its own could potentially facilitate learning of phonological status in pre-lexical infants.

1.1. Phonetic instantiation as an additional cue to phonological status

A great deal of research in psychology suggests that category structure affects similarity judgments (e.g., Rosch, 1975). This general principle could also be evident in the phonetics–phonology interface (e.g., Kingston, 2007). A pair of sounds could come to be perceptually, articulatorily, and acoustically less different from each other if the dimension along which the sounds contrast is allophonic compared to when it is phonemic through the following perception–production loop.

First, a host of evidence shows that phonemic distinctions are more salient than comparable allophonic distinctions for adult perceivers. (For brevity, we refer to distinctions between allophones of the same phoneme as *allophonic* distinctions and distinctions between sounds that map onto different phonemic categories as *phonemic* distinctions.) To begin with, adult listeners rate sounds as more similar to each other if the sounds are used allophonically in the listeners' native language than if they are used phonemically (e.g., Boomershine, Hall, Hume, & Johnson, 2008; Johnson & Babel, 2010). Moreover, discrimination is faster, more accurate, and more categorical when a distinction is phonemic than when it is allophonic (Beddor & Strange, 1982; Boomershine et al., 2008; Harnsberger, 2001; Whalen, Best, & Irwin, 1997). Finally, sounds in an allophonic distinction do not block word recognition (e.g., the clear pronunciation of 'atom' [ætəm] and the flapped pronunciation [ærəm] prime each other and themselves to a similar extent, whereas sounds in a phonemic distinction do not, e.g., 'Adam' [ædəm] does not prime [ætəm]; McLennan, Luce, & Charles-Luce, 2003).

These perceptual differences could lead to less clear productions of allophonic distinctions since speakers tend to articulate less clearly distinctions that they find harder to discriminate (Perkell et al., 2004). Furthermore, given that listeners can still retrieve the lexical item even when a different allophone is uttered (McLennan et al., 2003), there is less pressure for talkers to hyperarticulate allophonic distinctions (Lindblom, 1990).

At this point, articulation feeds back on perception: If speakers hypo-articulate, then allophonic distinctions will tend to be acoustically less distinct than comparable phonemic distinctions. Indeed, such patterns have sometimes been reported (e.g., Gick, Pulleyblank, Campbell, & Mutaka, 2006, report that, in Kinande, the high vowels, phonemically contrastive in tenseness or advanced tongue root [ATR], are articulatorily more distinct than the low vowels that differ only allophonically in ATR; see also Johnson & Babel, 2010; Spears, 2006; Ussishkin & Wedel, 2009). If sounds come to be physically closer in the acoustic input of these speakers, then they will become perceptually more similar, reinforcing the perceptual step noted above.

Finally, the effects of this feedback loop could come to be 'phonologized' historically, such that sounds that are in an allophonic distinction would tend to become more similar to each other acoustically with passing generations. Regardless of whether phonological status affects acoustic implementation only synchronically, or whether additionally its effects come to be phonologized, the acoustic separation between two sounds would be smaller when they are in an allophonic, compared to when they are in a phonemic, distinction, in the spoken input available to the infant learner.

It is reasonable to suppose that such fine differences in acoustic instantiation could in turn affect infants' perception, since their perceptual categories are altered by the distributions of acoustic cues they encounter (Cristia, 2011; Cristia, McGuire, Seidl, & Francis, 2011; Maye, Weiss, & Aslin, 2008; Maye, Werker, & Gerken, 2002; Yoshida, Pons, Maye, & Werker, 2010). Thus, acoustic implementation could plausibly contribute to a decline in sensitivity for allophonic distinctions in infants, aside from any other cues that they may draw from their proto-lexicon and their knowledge of complementary distributions.

In the current study, we investigated whether there are differences in the instantiation of allophonic and phonemic distinctions in infants' input. We tested the hypothesis that two sounds would be more acoustically distinct when produced phonemically than allophonically by comparing vowel tenseness and vowel nasality in American English and Quebec French. Vowel tenseness is phonemic in American English (*seat* [sit] versus sit [sɪt]) and allophonic in Quebec French (where vowels are laxed in closed syllables; e.g., *ville* [vɪl] 'city' versus *village* [vi.laʒ] 'town'; Rose & Wauquier-Gravelines, 2007; Walker, 1984). Vowel nasality is allophonic in American English (where vowels are nasalized in syllables closed by nasal consonants; e.g., *bet* [bɛt] versus *Ben* [bɛ̃n]; Krakow, 1993) and phonemic in Quebec French (*monde* [mõd] 'world' versus *mode* [mod] 'fashion').

We measured acoustic–phonetic correlates of tenseness and nasality in the speech of three groups of mothers. One group lived in the American Midwest and spoke only American English. A second group lived in Montreal and spoke primarily Quebec French. The third group consisted of bilinguals living in Montreal, who spoke both Quebec French and Quebec English frequently, from an early age, and without a marked foreign accent. Each mother was recorded while talking with their 11-month-old about sets of objects, which had been selected so that the object labels contained specific target vowels. This infant-directed corpus contained data on two pairs of vowels differing in tenseness ([i-ɪ], for both English and French, and either [e-ɛ] for French or [eɪ-ɛ] for English) and two pairs of vowels differing in nasality ([ɛ-ɛ̃] and either [a-ã] for French or [æ-æ̃] for English). For each mother and for each vowel pair, we assessed the separation between the vowel centroids along several perceptually-relevant acoustic correlates.

If the perception–production loop described above is at least partially accurate, then we predict that the separation between [i-ɪ] will be larger in English than in French, as will the separation between [e(ɪ)-ɛ]. In contrast, the separation between [ɛ-ɛ̃] should be larger in French than in English, and the same should occur for [a/æ -ã/æ̃]. Our predictions only operate *within* contrast type, and not across them. Our key prediction is that the acoustic implementation of a given distinction might be affected by whether it is phonemic or allophonic (through the perception–production loop described above), and thus no predictions can or need to be made with respect to other contrasts in the same system. (Additionally, such a direct comparison would also be empirically inadequate, as we will see below.)

1.2. Phonetic instantiation in bilingual speech

The present study elaborates on data from both monolinguals and bilinguals. The latter population contributes to our research question in two ways.

First, the speech of bilinguals constitutes a strong test of our hypothesis. As noted above, a difference in acoustic–perceptual distance depending on phonological status should follow from the structure of the phonological and lexical systems, and it would be reinforced by perception and production practice. Therefore, such a difference should be evident *even in the speech of the same person* regularly using two different language systems. For example, a larger separation along tenseness for monolingual English speakers than monolingual French speakers could arise if the former speak more clearly than the latter. Such a difference in the speech of French–English bilinguals would constitute stronger evidence that phonological status can influence phonetic instantiation, bypassing alternative explanations based on individual differences.

Second, incorporating a group of bilingual mothers may provide unique evidence regarding the availability of such cues in bilingual language acquisition, given that bilingual learners may be facing different problems than monolingual ones (Werker, 2012). We can state two diametrically opposed predictions. On the one hand, some research suggests that phonological targets in bilinguals are less distinct than those in monolinguals – in other words, that under certain conditions bilinguals merge targets across their two languages. Indeed, late bilinguals with a high level of use of their first language tend to have targets that are intermediate between their two languages (e.g., Flege, Schirru, & MacKay, 2003) and compromises have also been reported for early balanced bilinguals, particularly in perception (for example, Pallier, Bosch, & Sebastián-Gallés, 1997). According to such reports, we might find that the differential realization of tenseness and nasality is reduced in bilinguals as compared to monolinguals. Alternatively, balanced bilingual caregivers may attempt to maintain or enhance the distinctiveness between the two languages when directly addressing their child. In this case, acoustic cues to phonological status might be maintained or even enhanced in bilingual speech.

2. Methods

2.1. Participants

Three groups of mothers whose children were about 11 months of age were recorded.[1] One monolingual group spoke Quebec French (henceforth QF). It consisted of 21 mothers whose children's age was on average 11 months and 3 days (range 10,20 to 11,18; 11 girls). Another 21 monolingual mothers spoke American English (henceforth AE). The average addressee age was 11,13 (range 11,00 to 12,00; 9 girls). Finally, 9 early balanced bilingual mothers were recorded speaking Quebec French (henceforth BF) and Quebec English (henceforth BE), to their child, whose age averaged 10,28 (range 10,15 to 11,10; 5 girls).

All children were born fullterm and had normal hearing according to parental report. Caregivers were all highly educated, with no differences across groups (QF: $M=15.8$ years of education; range 11–22; AE: $M=16$, range 12–22; bilinguals $M=16.4$, range 13–20). The three groups of talkers did differ in the percentage of time they spoke another language (QF spoke some other language $M=16\%$, range 0–40; AE spoke some other language $M=1\%$, range 0–20; bilinguals spoke English $M=45\%$, range 25–60; French $M=55\%$, range 40–75; some other language 0%). Families received a small gift and a diploma for their participation.

2.2. Elicitation materials

Both nasalization in AE (Krakow, 1993) and closed-syllable laxing in QF (Goldsmith, 2009) have been described as phonological processes whose application depends on syllable structure. Nonetheless, there is some variability in their application across different vowel qualities. In most dialects of French, closed-syllable laxing is applied systematically for [e ɛ] (e.g., Desrochers, 1994; see also Rose & Wauquier-Gravelines, 2007, for arguments that laxing is partially modulated by focus). The sounds in this allophonic distinction are phonemically distinct in another phonological position; that is, [e] and [ɛ] are distinctive in open, word-final syllables, such as *fée* [fe] 'fairy' versus *fait* [fɛ] 'fact'. In QF closed-syllable laxing also affects high vowels [i ɪ] (Goldsmith, 2009; Martin, 2002; Rose & Dos Santos, 2007). Since in one case the distinction is purely allophonic, and in the other it is not, Walker (1984) describes them as different processes (see e.g., p. 52). Therefore, we selected one pair of high vowels, namely [i ɪ], and another of mid vowels ([e(ɪ) ɛ]), with the knowledge that the former case may present a cleaner test of our predictions.[2] The vowels selected for nasality were the unrounded [ɛ ɛ̃] in QF and AE and [a ã] in QF and [æ æ̃] in AE.

Instances of the four pairs of vowels were elicited through pictures and objects representing target words, which contained the target vowels while trying to control for phonological context and lexical frequency across vowels within a pair and across the languages. We consulted several native speakers of each language group to confirm this classification. This was especially important for QF, where several processes aside from closed vowel laxing (such as diphthonguization) are active. The complete list of targets are given in Tables 1 and 2; notice that some items had to be dropped or reclassified post-hoc to avoid inflating differences across languages.[3]

2.3. Equipment and procedure

Recordings took place in a sound attenuated room. Monolingual English speakers were recorded in Indiana; monolingual French as well as bilingual speakers in Montreal. The mother was fitted with a Lavalier microphone (AKG WMS40 Pro Presenter Set Flexx UHF Diversity CK55), whose signal was recorded into a Marantz Professional Solid State Recorder (PMD660ENG). She was given 31 bags, each containing one object/picture set and with the target word written on each bag, and was then left alone with the child for about 30–45 min. Participants were told that we were interested in how caregivers talk about objects, and they were asked to describe the contents of each bag/set. In each set, there were two similar objects or pictures (e.g., two tambourines of different sizes) representing the target word and an oddball (a violin). They were provided an instance of the task, for

[1] Twenty-two additional caregivers could not be included for the following reasons: Caregiver spoke a different dialect (since laxing varies across different dialects of French, we only included mothers who had a Quebec accent according to self-report, which was confirmed by the French-speaking coders; 4 QF and 2 bilinguals were excluded based on this criterion); caregiver was male (3 AE); recording was not completed due to child fussiness (3 bilinguals; infants participated in two sessions); data loss or poor quality of the recording (2 AE and 1 bilingual); either the parent or the child had a disorder (1 AE, 1 QF, and 1 bilingual). Among bilinguals, 4 additional caregivers were excluded because they received high ratings of accentedness in either language (higher than 5.5 on a scale of 1 = no accent and 10 = heavy accent by 5 naïve listeners of each language).

[2] The last author has documented several stages of this project in Cristia (2013). This website includes: Annotated bibliographies on some relevant topics, coder training materials; Praat scripts for acoustic analyses; R scripts for statistical analyses; a full report of the tuning of all algorithms used in the current paper on the Hillenbrand corpus (Hillenbrand, Getty, Clark, & Wheeler, 1995), another naturalistic IDS corpus (Cristia, 2010), and a small subset of the current data (with 6 caregivers in each monolingual group).

[3] Preliminary analyses revealed that *picot* did not represent the intended vowel; together with its English counterpart, it was removed from further consideration. *Tissu* (originally lax), *dix-sept* (originally tense), and *tigre* (originally tense) were implemented as the opposite tenseness value. To avoid artificially reducing the distance spanned by tenseness contrasts in French, they were reassigned at the analysis stage.

Table 1
English target words used to elicit the target vowels.

			Phonemic [i-ɪ]				
beetle	piglet	peekaboo	picnic	teakettle	dictionary	teaspoon	disney
			[eɪ-ɛ]				
bacon	pegboard	basil	pesto	daycare	decker	N/A[a]	pepsi
			Allophonic [ɛ-ɛ̃]				
pepsi	Benji	best in show	pencil	pedal	pendant	Teddy	tender
			[æ-æ̃]				
baboon[b]	bamboo	tassle	dancer	bassinet	pansy	tapioca	tamborine

[a] Items that were inconsistently pronounced were removed from consideration.
[b] See the project website for analyses excluding baboon.

Table 2
French target words (and English gloss) used elicit the target vowels.

			Allophonic [i-ɪ]				
bicyclette *bicycle*	pistache *pistachio*	pyjama *pyjamas*	pique-nique *picnic*	N/A[a]	dictionnaire *dictionary* tigre[a] *tiger*	tissu[a] *cloth*	dix-sept[a] *seventeen*
			[e-ɛ]				
béquille *crutch*	bec de canard *duck beak*	bétail *cattle*	bestiole *bug*	décor *setting*	technicien *technician*	N/A[a]	pectoraux *pectorals*
			Phonemic [ɛ-ɛ̃]				
bec de canard *duck beak*	pain baguette *baguette*	pêcheur *fisherman*	pince à sucre *sugar tongs*	betterave *beet*	pentagone/ pinte de lait *pentagon/ pint of milk*	Teddy *(name)*	Tintin *(name)*
			[a-ã]				
bavette *bib*	bambou *bamboo*	tasse à café *coffee cup*	danseur *dancer*	bassinet *carriage*	pansement *bandaid*	tapioca *tapioca*	tambourin *tamborine*

[a] Items that were inconsistently pronounced were removed from consideration.

instance using two triangles and a square. Bilingual caregivers and their infants participated in the task twice within one session, once in each language, with language order counterbalanced across participants.

2.4. Coding and acoustic analyses

Coding and acoustic analyses were conducted in Praat (Boersma & Weenink, 2005; Linux version Praat 5.1.42). For each language, native or near-native speakers listened to the files and marked the approximate location of the syllable from the target word containing the target vowel. Three highly trained coders then marked the target vowel onset and offset for all of the syllables in both languages. The onset of the vowel was defined as the first upward crossing after the onset of periodicity following the burst or fricative release of the preceding stop. The offset of the vowels was determined as an abrupt attenuation of energy, evident in both the waveform and the spectrogram. Coders were particularly careful when determining the offset of a vowel followed by a nasal consonant, which is considered particularly difficult when the vowel is nasalized. In these cases, all sources of information (attenuation and increase in regularity in the waveform; sudden appearance of nasal zeros in the spectrogram; and auditory feedback) were used, and all cases of disagreement were resolved by consensus. Vowels were not tagged, coded, or analyzed if the formant structure was affected by background noise or talker overlap, or if the word had been whispered or glottalized. Tagged vowels shorter than 40 ms were excluded from further analyses, as the window of analysis used for nasality would then include neighboring sounds. This minimum length criterion was extended to the tenseness distinctions, to keep the vowel samples comparable.

The acoustic correlates considered for tenseness were the frequency of the first and second formants in Hertz ($F1$ and $F2$) and vowel duration. Both spectral and durational correlates affect listeners' identification of tense and lax vowels (e.g., Daniloff, Shriner, & Zemlin, 1968; Hillenbrand & Nearey, 1999), although their relative contribution varies with the vowel category (e.g., Hillenbrand et al., 1995), and the listeners' native language (even among listeners for whom vowel tenseness is phonemic, as in Northern German versus American English; Strange et al., 2007). Tense vowels have longer durations and more peripheral formant frequencies than lax vowels.

The correlates considered for nasality were $F1$ bandwidth ($F1$ bw) and the difference in amplitude between the first formant and the first (P0) and second (P1) nasal poles (henceforth A1P0 and A1P1). $F1$ bw and amplitude, and the relative increase of energy in the lower frequency regions, are among the most reliable acoustic–perceptual correlates of nasalization (e.g., Beddor, 1993; Hawkins & Stevens, 1985; Ohala & Ohala, 1993), as they are present (and perceived) regardless of vowel quality. Moreover, they characterize all types of nasalization, including coarticulatory (e.g., Chen, 1997) and phonemic nasalization (e.g., Maeda, 1993), as well as hypernasal and functionally nasalized speech (Dickson, 1962). Furthermore, direct comparisons suggest that such perceptually informed acoustic parameters outperform uninformed acoustic descriptors (such as Mel Frequency

Cepstral Coefficients; Pruthi, 2007). Nasal vowels have lower A1P0, lower A1P1 (both due to an increase in the intensity of the nasal poles, and a decrease in the intensity of the first formant) and a larger *F*1 bw than do oral vowels. It is clear that there are many alternative and additional correlates that could have been examined. We return to the difficulty of measuring vowel nasality from the acoustic signal in the discussion.

As in Escudero, Boersma, Rauber, and Bion (2009), the logarithm was applied to vowel duration to better reflect perception. Formant frequencies were measured using an implementation of the unsupervised ceiling optimization algorithm proposed by Escudero et al. (2009). All relevant formant measurements (frequencies for tenseness, amplitude and bandwidth for nasality) were carried out at two points in the vowel. The first point was at about 40% of the vowel duration, given that automatic measurements around this point correlate best with measurements carried out by extremely well-trained phoneticians (Evanini, 2009). The second point was at about 80%. This second point was important for tenseness because vowel-internal formant changes improve categorization particularly for diphthongs like [eɪ] (e.g., Hillenbrand & Nearey, 1999), and because nasalization also varies within the vowel (e.g., Chen, 1997). The poles P0 and P1 were identified using an implementation of the unsupervised algorithm proposed by Pruthi (2007).

2.5. Statistical analyses

Since the corpus is spontaneous speech, there was no control over how many tokens each caregiver produced of each target. Therefore, data for each vowel pair, (language, where relevant,) and speaker were only included in subsequent analyses if there were at least 4 tokens of each paired vowel type. The numbers of monolingual speakers that could be included were the following: [i ɪ] 15 AE and 14 QF; [e(ɪ) ɛ] 16 AE and 18 QF; [ɛ ɛ̃] 16 AE and 20 QF; [a/æ ã/æ̃] 17 AE and 19 QF. All 9 bilingual caregivers could be included in all contrasts. Analyses were carried out in R (Mac version 3.0.1, R Foundation for Statistical Computing, 2009).

In previous work, we have used a measure of Euclidean separation in the multidimensional space defined by all acoustical correlates (Cristia & Seidl, in press). This measure is limited in three ways.[4] First, it only represents how much two sounds differ within talkers, and therefore it does not take into account whether this difference is consistent across talkers. Second, a large separation could come about through a large separation along a single correlate (and no separation in all other correlates), or through small separations along several correlates that are compounded in the Euclidean distance. Yet we do not know whether infants can discriminate sounds equally well in both of these hypothetical situations (see a discussion in Cristia et al., 2011). Third, dimensions are compounded attributing all correlates the same weight. However, this might not represent the perception of English- and French-learning infants, since cue weighting varies across languages (Strange et al., 2007) and over the course of development (e.g., Mayo & Turk, 2004).

Therefore, the present study reports on effect sizes calculated within each of the two languages (English, French) and each of the two populations (monolinguals, bilinguals), separately for each acoustic correlate (5 correlates for tenseness, 6 for nasality) and vowel contrast ([i-ɪ] and [e(ɪ)-ɛ], or [ɛ-ɛ̃] and [a/æ-ã/æ̃], respectively). This process can be best understood through a specific example. A dataset was extracted containing all the tokens of both [i] and [ɪ] spoken by bilingual speaker 1 in her French session. After checking that there were at least 4 [i] and 4 [ɪ] tokens, we calculated the mean value for *F*1 at 40% of the vowel duration separately for [i] (*BF1.F1.40.i*) and [ɪ] (*BF1.F1.40.ɪ*). We then subtracted the tense value from the lax value (*BF1.F1.40.diff*=*BF1.F1.40.ɪ*−*BF1.F1.40.i*). The difference score was calculated for all the other speakers. Then the group effect size (*BF.F1.40.d*) was calculated for each language and each population as the difference score average divided by its standard deviation for the language and population. Following standard meta-analytic practice (Lipsey & Wilson, 2001), we applied Hedges' correction for bias in small samples.

Comparison of effect sizes across languages is only sensible if the two languages use a given correlate in a similar way. Therefore, effect sizes of different signs across languages were described as inconsistent. Furthermore, we excluded from consideration effect sizes that went against the established phonetic facts, noted in Section 2.4. Effect sizes depend on the size and direction of difference as well as on the stability across subjects. Our expectation is that effect sizes will be greater in English than French for the tenseness vowel pairs; and that they will be greater in French than English for the nasality vowel pairs. Should this general pattern be particularly evident in some correlates, this may indicate that these correlates are affected to a greater extent by the perception–production loop described in Section 1.1.

We complement our descriptive analyses with inferential statistics using Analysis of Variance for multiple dependent measures (MANOVA). While MANOVA suffers from some limitations (e.g., it assumes that all correlates are weighted equally), it allows us to provide a level of statistical significance for the answer to our research question. Averages for each relevant acoustic correlate were calculated for each speaker and vowel in a pair (e.g., *BF1.F1.40.i* and *BF1.F1.40.ɪ*). Separate MANOVAs were fit to the data from monolinguals and bilinguals. Additionally, since participants could maximally provide data for the 4 vowel pairs ([i-ɪ], [e(ɪ)-ɛ], [ɛ-ɛ̃], [a/æ-ã/æ̃], *p*-values were Bonferroni-corrected by multiplying by 4. Each of the 8 MANOVAs had the same internal structure: All relevant correlates were then declared as joint outcome measures, with language (English, French) and dimension (tense, lax for the tenseness vowel pairs; nasal, oral for the nasality vowel pairs) as predictors. Here, we report effects across correlates. Results separated by correlate can be found at the project website (Cristia, 2013). It should be noted that the Box's M test was significant for some factors in some models, suggesting that the assumption of homogeneity of variance–covariance matrices was not always met.

3. Results

We begin with the descriptive data. Fig. 1 shows the effect size for English and French separately for monolinguals and bilinguals, for each correlate. The first observation that can be made from this figure is that there is considerable variation in the absolute effect size across correlates for any given vowel pair, and thus not all correlates could provide equally strong cues. Second, some correlates are inconsistent across languages. For example, lax vowels have a higher *F*2 at 80% of the vowel duration than tense ones in French [i ɪ] whereas the opposite is true for English. Third, some correlates are inconsistent with previous phonetic reports. For instance, A1P0 at 80% was larger for nasal [ã/æ̃] than oral [a/æ] (cf. Chen, 1997).

For the remainder of this descriptive analysis, we set aside the inconsistent cases and concentrate on the remaining 16 tenseness and 18 nasality correlates. To facilitate inspection, we calculated a difference in effect sizes across languages for monolinguals and bilinguals separately. These are presented in Tables 3 and 4, where positive numbers indicate differences that fit our predictions.

[4] We are grateful to our reviewers for helpful discussion on these points.

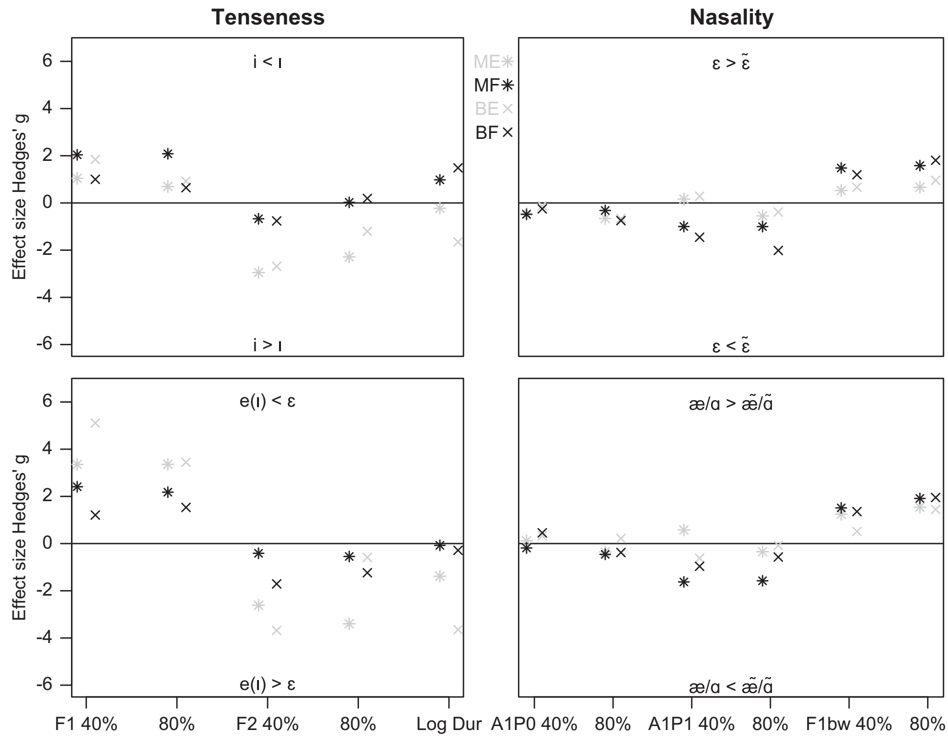

Fig. 1. Mean population-level effect sizes for each acoustic correlate relevant to the two vowel pairs differing in tenseness on the left panels (top [i-ɪ]; bottom [e(ɪ)-ε]); and those relevant to nasality on the right panels (top [ε-ε̃], bottom [a/æ-ã/æ̃]). English is shown in light gray, French in black. Monolinguals are represented by ⁎, bilinguals by x. In the left panels, a positive value indicates that the means for lax vowels were higher than those for tense ones (a negative value, the opposite). In the right panels, a positive value indicates that means for nasal vowels were higher than those for oral ones (a negative value, the opposite).

Table 3
Effect size differences (English minus French) for the two tenseness pairs for each acoustic correlate, for monolingual (M) and bilingual (B) speakers. Cells are blank if effect sizes were inconsistent across languages.

	*F*1 40%		*F*1 80%		*F*2 40%		*F*2 80%		Log Dur	
	M	B	M	B	M	B	M	B	M	B
[i-ɪ]	−1.00	0.84	−1.39	0.28	2.28	1.92				
[e(ɪ)-ε]	0.94	3.90	1.18	1.91	2.20	1.97	2.85	−0.65	1.31	3.36

Table 4
Effect size differences (French minus English) for the two nasality pairs for each acoustic correlate, for monolingual (M) and bilingual (B) speakers. Cells are blank if effect sizes were inconsistent across languages or with established phonetic results.

	A1P0 40%		A1P0 80%		A1P1 40%		A1P1 80%		*F*1 bw 40%		*F*1 bw 80%	
	M	B	M	B	M	B	M	B	M	B	M	B
[ε-ε̃]	0.02	0.13	−0.34	0.09			0.45	1.63	0.95	0.54	0.92	0.85
[a/æ -ã/ æ̃]		0.13				0.35	1.24	0.49	0.27	0.85	0.37	0.51

Inspection of Tables 3 and 4 suggests that the differences generally fit our predictions: 13 of the 16 tenseness correlates show larger effect sizes for English than French, whereas 17 out of the 18 nasality correlates show larger effect sizes for French than English. Nonetheless, some variation is evident across dimensions and correlates, as follows.

For the [i-ɪ] pair, bilinguals show larger effect sizes in English than French for all 3 of the consistent correlates, while only *F*2 at 40% fit our predictions in the monolingual comparison. For the [e(ɪ)-ε] pair, across both populations, effect sizes were larger for English than French for all correlates except *F*2 80% for bilinguals. The largest absolute effect size difference was for *F*2 80% for the monolinguals and *F*1 40% for bilinguals.

We illustrate these comparisons by plotting *F*1 as a function of *F*2 at 40% of the vowel duration, since the largest absolute effect size differences tended to be found in those correlates. It is clear from Fig. 2 that differences between [i] or [e(ɪ)] (closed circles) and [ɪ] or [ε] (open circles) are present in both English (gray) and French (black); and these differences are larger for English than French, for both monolinguals (left panel) and bilinguals (right panel).

For the [ε-ε̃] pair, across both populations, effect sizes were larger for French than English for all correlates except A1P0 80% for bilinguals. For the [a/æ-ã/æ̃] contrast, all correlates fit our predictions.

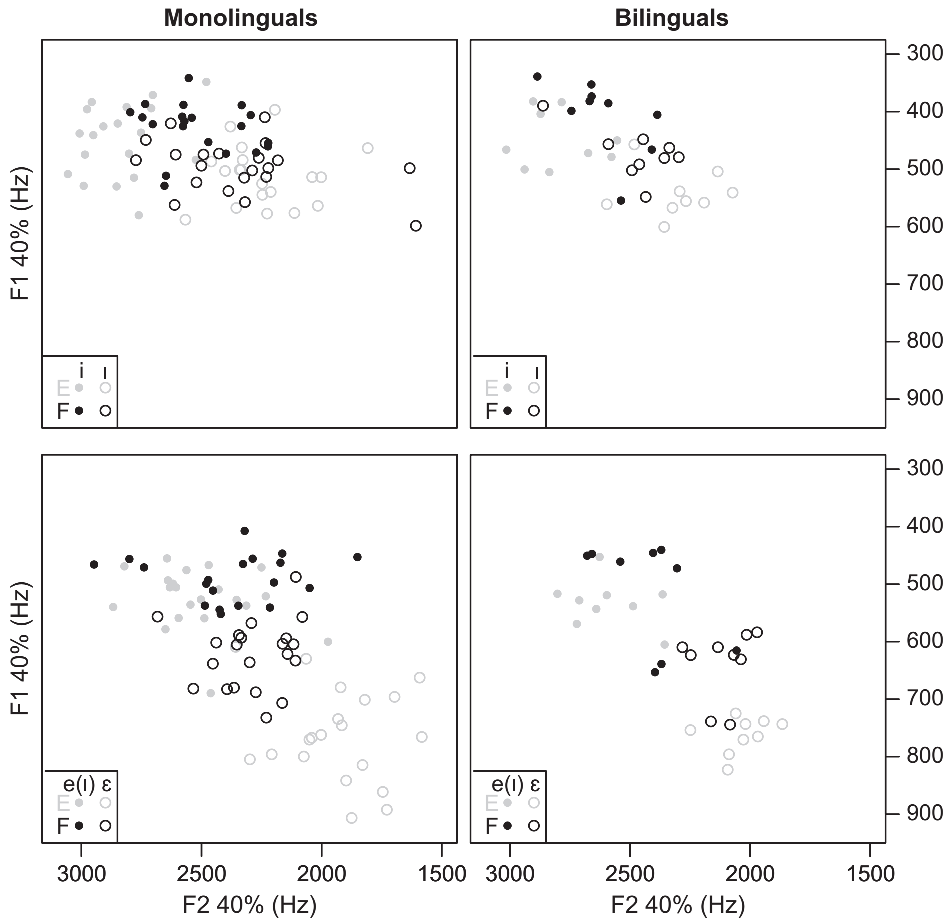

Fig. 2. Mean *F*1 and *F*2 values at 40% of the vowel for each speaker and vowel for the tenseness contrasts.

Table 5
F-values from MANOVAs predicting all correlates jointly in the monolingual (M) and bilingual (B) data, from the factors of Language (French, English), Dimension (tense, lax; or oral, nasal as relevant), and their interaction. Values in bold are significant after Bonferroni correction.

	[i-ɪ]		[e(ɪ)-ɛ]		[ɛ-ɛ̃]		[a/æ-ã/æ̃]	
	M	B	M	B	M	B	M	B
Language	3.25	3.08	**15.33**	**11.01**	**10.36**	1.27	**4.42**	2.18
Dimension	**21.78**	**14.00**	**51.61**	**42.74**	**5.18**	**4.51**	**12.69**	**6.88**
Interaction	**4.63**	**8.22**	**12.34**	**12.55**	1.57	2.43	**5.16**	0.69

In sum, descriptive analyses suggest that for tenseness, effect sizes tended to be larger for English than French, particularly for the [e(ɪ)-ɛ] contrast. For nasality, the opposite was true: effect sizes were larger for French than for English, in line with our predictions. There were no salient differences between monolinguals and bilinguals.

As explained in Section 2.5, our inferential statistics were MANOVAs, through which we attempted to predict all relevant correlates jointly (and not only consistent ones) from the factors of Language, Dimension, and their interaction. The results, portrayed in Table 5, confirmed that the realization of tenseness depended on language for both monolinguals and bilinguals for both [i-ɪ] and [e(ɪ)-ɛ]. For example, the interaction between Language (French, English) and Dimension (tense, lax) for [i-ɪ] was $F(5,50)=4.63$, $p=.006$. For nasality, only the monolingual data for [a/æ-ã/æ̃] evidenced a significant interaction between Language (French, English) and Dimension (oral, nasal). The other three MANOVAs did *not* show a significant interaction between Language and Dimension. As with the descriptive data, there were no salient differences between monolinguals and bilinguals.

4. Discussion

The present research speaks to a key theoretical question: Is there acoustic information available in infant-directed speech that could contribute to a loss of sensitivity to allophonic distinctions? Previous work has explored protolexical and distributional cues to phonological status, and we add to the picture by assessing availability of acoustic cues. Our analyses demonstrate a clear difference in acoustic implementation of tenseness in a language where the dimension is phonemic versus one where the dimension is used allophonically, evidenced by larger effect sizes in English than in French.

These cross-linguistic differences were evident in a between-subject comparison based on two groups of monolingual speakers. Had this been our only data point, such a difference could have been attributed to multiple causes, including talker-specific differences (i.e., Quebec speakers could have been less clear than their American counterparts). By including the within-participant comparisons instantiated in the bilingual speakers, we could rule this explanation out, because tenseness interacted with language even in within-participant analyses.

Additionally, our bilingual data speak to the input available to infants hearing two languages from the same speaker. It is remarkable that the contrast between languages was as large for the between-participant monolingual comparison as for the within-participant, bilingual comparison. Moreover, most correlates were used the same way across languages and across populations. Additional analyses of a range of acoustic features characterizing infant-directed speech (e.g., higher pitch, slower speech rate, and point vowel implementation) in the same three groups of talkers also shows that the two languages spoken by a bilingual speaker to her infant can be as different from each other as the English and French spoken to monolingual infants (Danielson, Seidl, Onishi, Alamian, & Cristia, 2014).

Yet the monolingual and bilingual tenseness data are not sufficient to demonstrate that there is a difference in acoustic implementation as a function of the phonological status of a contrast. Indeed, these results could still be explained through baseline differences across languages (i.e., French is spoken less clearly than English). To address this possibility, we had adopted a crossed design, in which another dimension (vowel nasality) was expected to pattern in the opposite manner, with larger effect sizes in French than English. While descriptive data based on effect sizes fit our predictions, inferential statistics largely failed to provide significant support for our hypothesis.

Why were the results less clear for nasality than tenseness? This is likely explained, in part, by the difficulty of measuring vowel nasality acoustically. Indeed, while human listeners are remarkably sensitive to nasality (e.g., Beddor, 1993), robust and stable acoustic correlates that can be estimated automatically have yet to be found. Over half a dozen have been proposed, including: overall amplitude, absolute $F1$ amplitude, amplitude of the energy in the $F2$/$F3$ frequency bands, the shape of the spectrum below 1 kHz, and others (see Pruthi, 2007, for discussion), but their reliability is variable for several reasons. First, nasalization involves coupling the vocal tract with the nasal tract, which is a complex and highly variable resonating cavity (or set of cavities). This coupling introduces additional poles and zeros whose frequency and number vary due to many factors, including the specific morphology of the talker's paranasal cavities and asymmetries between the nostrils (Pruthi & Espy-Wilson, 2007). Additionally, some speakers of French seem to modify their oral tract when producing nasal vowels, introducing additional differences in $F1$, $F2$, and $F3$ between the oral and nasal vowels, but not all talkers incorporate the same oral modifications (Delvaux, Demolin, Harmegnies, & Soquet, 2008; see also a discussion of previous reports in Beddor, 1993). These strategies result in complex acoustic correlates not yet captured by automated analyses. While nasalization may be more reliably measured through articulation (e.g., nasal/oral airflow; Delvaux, Metens, & Soquet, 2002) than acoustics, it is likely that the use of such equipment would compromise the naturalness of the infant-directed speech gathered.

We hope future work may explore phonological status effects on phonetic instantiation in several ways. First, while it is easy to predict a gradient between allophonic and phonemic distinctions in vowels, which vary along largely continuous dimensions, it is more difficult to make similarly specific predictions for sounds varying in non-continuous ways, such as manner of articulation in consonants.

Second, it would be of interest to understand how the phonological system as a whole shapes phonetic implementation. We studied only two vowel pairs in each dimension, and already differences across vowel quality were apparent. For example, we expected [i-ɪ] to provide clearer evidence for our hypothesis than [e(ɪ)-ɛ], since the latter pair is contrastive in Quebec French word-finally. However, the effect of language was equally large for both, and more stable across acoustic correlates in the latter than the former. Based on these findings, it might be tempting to conclude that phonological-status effects are tied to specific syllable positions, although other interpretations remain possible (e.g., perhaps speakers have greater freedom to vary their tenseness realization when producing mid vowels).

5. Conclusions

Allophones, unlike non-native sounds, occur in infants' input. Previous work shows that, in spite of this exposure, infants treat distinctions that are phonemic differently from those that are allophonic by 11 months of age. The current study demonstrated that, at least in certain cases, there are differences in the separation between two sounds depending on whether they participate in an allophonic or phonemic distinction in the speaker's language. This difference in acoustic implementation was found in a comparison of speech produced by two groups of monolingual caregivers, as well as within the speech of bilingual caregivers.

Acknowledgments

We thank the families who participated as well as Krista Ashmore, Allison Gladfelter, and Kat Montoya for the phonetic coding. We are grateful to Kyle Danielson for checking the coding for the bilingual speakers, and for the design and analysis of accentedness ratings. Portions of this work were discussed at the following conferences: 2010 JEP, November 2010 ASA, the 2011 workshop Testing Models of Phonetics and Phonology, 2013 ICPC, 2013 WILD. We thank these audiences and our colleagues at Purdue, the LSCP, and UPenn for their comments. We also appreciate the invaluable scientific input provided to us by Alex Francis, Jonah Katz, Maria Kondaurova, Titia Benders, Todd M. Bailey, and an anonymous reviewer on previous versions; and the advice on the statistical analyses furnished by the Purdue Statistical Consulting Service. AS, KO, and GA designed the elicitation material; AS, KO, and GA collected the recordings; AS and KO supervised the coders; AC generated the coder training materials, performed the acoustical and statistical analyses, and wrote the manuscript. This work was supported by funds from NSF 0843959 to AS. AC also acknowledges the financial support of the Fyssen Foundation, and the institutional support of ANR-10-LABX-0087 IEC and ANR-10-IDEX-0001-02 PSL*.

References

Beddor, P. S. (1993). The perception of nasal vowels. In: M. K. Huffman, & R. A. Krakow (Eds.), *Nasals, nasalization, and the velum, Vol. 5: Phonetics and phonology* (pp. 171–196). San Diego, CA: Academic Press.

Beddor, P. S., & Strange, W. (1982). Cross-language study of perception of the oral–nasal distinction. *Journal of The Acoustical Society of America*, *71*, 1551–1561.

Boersma, P., & Weenink, D. (2005). *Praat: Doing phonetics by computer (version 5.0.09)*. From: ⟨//www.praat.org/⟩ Retrieved 04.04.09.

Boomershine, A., Hall, K. C., Hume, E., & Johnson, K. (2008). The impact of allophony vs. contrast on speech perception. In: P. Avery, E. Dresher, & K. Rice (Eds.), *Contrast in phonology* (pp. 143–172). Berlin: de Gruyter.

Chen, M. (1997). Acoustic correlates of English and French nasalized vowels. *Journal of The Acoustical Society of America*, *102*, 2360–2370.
Cristia, A. (2010). Phonetic enhancement of sibilants in infant-directed speech. *Journal of the Acoustical Society of America*, *128*, 424–434.
Cristia, A. (2011). Fine-grained variation in caregivers' /s/ predicts their infants' /s/ category. *Journal of The Acoustical Society of America*, *129*, 3271–3280.
Cristia, A. (2013). *NSF 0843959: Acoustic cues to allophony*. URL: 〈http://sites.google.com/site/acrsta/Home/nsf_allophones_corpora〉 Last visited 26.08.13.
Cristia, A., McGuire, G. L., Seidl, A., & Francis, A. L. (2011). Effects of the distribution of cues on infants perception of speech sounds. *Journal of Phonetics*, *29*, 388–402.
Cristia A. and Seidl A. (in press). The hyperarticulation hypothesis of infant-directed speech. Journal of Child Language.
Danielson, K., Seidl, A., Onishi, K., Alamian, G., & Cristia, A. (2014). Infant-directed speech in bilinguals. *Journal of Acoustical Society of America Express Letters*, *135*, EL95–EL101.
Daniloff, R. G., Shriner, T. H., & Zemlin, W. R. (1968). Intelligibility of vowels altered in duration and frequency. *Journal of The Acoustical Society of America*, *44*, 700–707.
Delvaux, V., Demolin, D., Harmegnies, B., & Soquet, A. (2008). The aerodynamics of nasalization in French. *Journal of Phonetics*, *36*, 578–606.
Delvaux, V., Metens, T., & Soquet, A. (2002). Proprietés acoustiques et articulatoires des voyelles nasales du français. *Journées d'Etude sur la Parole*, *XXIV*, 357–360.
Desrochers, R. (1994). Le mecanisme des strategies de reparation en phonotactique generative et la diphtongaison en français montrealais. *Revue Québecoise de Linguistique*, *23*, 115–148.
Dickson, D. R. (1962). Acoustic study of nasality. *Journal of Speech and Hearing Research*, *5*, 103–111.
Escudero, P., Boersma, P., Rauber, A. S., & Bion, R. A. H. (2009). A cross-dialect acoustic description of vowels: Brazilian and European Portuguese. *Journal of the Acoustical Society of America*, *126*, 1379–1393.
Evanini, K. (2009). *The permeability of dialect boundaries: A case study of the region surrounding Erie, Pennsylvania [Ph.D. thesis]*. University of Pennsylvania.
Flege, J. E., Schirru, C., & MacKay, I. R. (2003). Interaction between the native and second language phonetic subsystems. *Speech communication*, *40*(4), 467–491.
Foulkes, P., Docherty, G., & Watt, D. (2005). Phonological variation in child-directed speech. *Language*, *81*, 177–206.
Fourtassi, A., & Dupoux, E. (2014). The role of bottom-up and top-down cues in the acquisition of phonemes [in preparation].
Gick, B., Pulleyblank, D., Campbell, F., & Mutaka, N. (2006). Low vowels and transparency in Kinande vowel harmony. *Phonology*, *23*, 1–20.
Goldsmith, J. (2009). The syllable. In: J. Goldsmith (Ed.), *The handbook of phonological theory*. Cambridge, Mass: Blackwell.
Hall, K. C. (2013). A typology of intermediate phonological relationships. *Linguistic Review*, *30*, 215–275.
Harnsberger, J. D. (2001). The perception of Malayalam nasal consonants by Marathi, Punjabi, Tamil, Oriya, Bengali, and American English listeners: A multidimensional scaling analysis. *Journal of Phonetics*, *29*, 303–327.
Hawkins, S., & Stevens, K. N. (1985). Acoustic and perceptual correlates of the nonnasal–nasal distinction for vowels. *Journal of the Acoustical Society of America*, *77*, 381–396.
Hillenbrand, J. M., Getty, L. A., Clark, M. J., & Wheeler, K. (1995). Acoustic characteristics of American English vowels. *Journal of the Acoustical Society of America*, *97*, 3099–3111.
Hillenbrand, J. M., & Nearey, T. M. (1999). Identication of resynthesized /hVd/ utterances: Effects of formant contour. *Journal of the Acoustical Society of America*, *105*, 3509–3524.
Johnson, K., & Babel, M. (2010). On the perceptual basis of distinctive features: Evidence from the perception of fricatives by Dutch and English speakers. *Journal of Phonetics*, *38*, 127–136.
Kenstowicz, M. (1994). *Phonology in generative grammar*. Cambridge: Blackwell.
Kingston, J. (2007). The phonetics–phonology interface. In: P. de Lacy (Ed.), *Cambridge handbook of phonology* (pp. 401–434). Cambridge: Cambridge University Press.
Krakow, R. A. (1993). Nonsegmental influences on velum movement patterns: syllables, sentences, stress, and speaking rate. In: M. K. Huffman, & R. A. Krakow (Eds.), *Nasals, nasalization, and the velum, Vol. 5: Phonetics and phonology,* (pp. 87–116). San Diego, CA: Academic Press.
Lindblom, B. (1990). Explaining phonetic variation: A sketch of the H&H theory. In: W. J. Hardcastle, & A. Marchal (Eds.), *Speech production and speech modelling* (pp. 403–439). Netherlands: Springer.
Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Thousand Oaks, CA: Sage.
Maeda, S. (1993). Acoustics of vowel nasalization and articulatory shift in French nasal vowels. In: M. K. Huffman, & R. A. Krakow (Eds.), *Nasals, nasalization, and the velum, Vol. 5: Phonetics and phonology* (pp. 147–167). San Diego, CA: Academic Press.
Martin, A., Peperkamp, S., & Dupoux, E. (2013). Learning phonemes with a proto-lexicon. *Cognitive Science*, *37*, 103–124.
Martin, P. (2002). Le système vocalique du français du Québec. De l'acoustique a la phonologie. *La Linguistique*, *38*, 71–88.
Maye, J., Weiss, D., & Aslin, R. (2008). Statistical phonetic learning in infants: Facilitation and feature generalization. *Developmental Science*, *11*, 122–134.
Maye, J., Werker, J. F., & Gerken, L. (2002). Infant sensitivity to distributional information can effect phonetic discrimination. *Cognition*, *82*, B101–B111.
Mayo, C., & Turk, A. (2004). Adult–child differences in acoustic cue weighting are influenced by segmental context. *Journal of the Acoustical Society of America*, *115*, 3184.
McLennan, C. T., Luce, P. A., & Charles-Luce, J. (2003). Representation of lexical form. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*, 539–553.
Ohala, J., & Ohala, M. (1993). The phonetics of nasal phonology: Theorems and data. In: M. K. Huffman, & R. A. Krakow (Eds.), *Nasals, nasalization, and the velum, Vol. 5: Phonetics and phonology* (pp. 225–249). San Diego, CA: Academic Press.
Pallier, C., Bosch, L., & Sebastián-Gallés, N. (1997). A limit on behavioral plasticity in speech perception. *Cognition*, *64*, B9–B17.
Peperkamp, S., LeCalvez, R., Nadal, J. P., & Dupoux, E. (2006). The acquisition of allophonic rules: Statistical learning with linguistic constraints. *Cognition*, *101*, B31–B41.
Perkell, J. S., Matthies, M. L., Tiede, M., Lane, H., Zandipour, M., Marrone, N., et al. (2004). The distinctness of speakers' /s/-/ʃ/ contrast is related to their auditory discrimination and use of an articulatory saturation effect. *Journal of Speech, Language, and Hearing Research*, *47*, 1259–1269.
Pruthi, T. (2007). *Analysis, vocal-tract modeling and automatic detection of vowel nasalization [Ph.D. thesis]*. University of Maryland.
Pruthi, T., & Espy-Wilson, C. Y. (2007). Simulation and analysis of nasalized vowels based on magnetic resonance imaging data. *Journal of The Acoustical Society of America*, *121*, 3858–3873.
Rosch, E. (1975). Cognitive representations of semantic categories. *Journal of Experimental Psychology: General*, *104*, 192–233.
Rose, Y., & Dos Santos, C. (2007). Stress domain effects in French phonology and phonological development. In: K. Arregi, Z. Fagyal, S. A. Montrul, & A. Tremblay (Eds.), *Romance linguistics 2008* (pp. 89–106). Philadelphia, PA: John Benjamins.
Rose, Y., & Wauquier-Gravelines, S. (2007). French speech acquisition. In: S. McLeod (Ed.), *The international guide to speech acquisition* (pp. 364–384). New York: Thomson Delmar Learning.
Seidl, A., Cristia, A., Bernard, A., & Onishi, K. H. (2009). Allophones and phonemes in infants' learning of sound patterns. *Language Learning and Development*, *5*, 191–202.
Spears, A. (2006). *Nasal coarticulation in the French vowel /i/: A phonetic and phonological study [Ph.D. thesis]*. University of North Carolina-Chapel Hill.
Strange, W., Weber, A., Levy, E. S., Shafiro, V., Hisagi, M., & Nishi, K. (2007). Acoustic variability within and across German, French, and American English vowels: Phonetic context effects. *Journal of the Acoustical Society of America*, *122*, 1111–1129.
Tsuji, S., & Cristia, A. (2014). Perceptual attunement in vowels: A meta-analysis. *Developmental Psychobiology*, *56*, 179–191.
Ussishkin, A., & Wedel, A. (2009). Lexical access, effective contrast and patterns in the lexicon. In: P. Boersma, & S. Hamman (Eds.), *Perception in phonology* (pp. 267–292). Amsterdam: Mouton de Gruyter.
Walker, D. (1984). *The pronunciation of Canadian French*. Ottawa: University of Ottawa Press.
Werker, J. (2012). Perceptual foundations of bilingual acquisition in infancy. *Annals of the New York Academy of Sciences*, *1251*, 50–61.
Whalen, D., Best, C., & Irwin, J. (1997). Lexical effects in the perception and production of American English /p/ allophones. *Journal of Phonetics*, *25*, 501–528.
White, K. S., Peperkamp, S., Kirk, K., & Morgan, J. (2008). Rapid acquisition of phonological alternations by infants. *Cognition*, *107*, 238–265.
Yoshida, K. A., Pons, F., Maye, J., & Werker, J. F. (2010). Distributional phonetic learning at 10 months of age. *Infancy*, *15*, 420–433.