2014-09

Utterance-Final Lengthening Is Predictive of Infants' Discrimination of English Accents

White, Laurence

http://hdl.handle.net/10026.1/9946

10.1111/lang.12060
Language Learning
Wiley

LANGUAGE LEARNING
A Journal of Research in Language Studies

Language Learning ISSN 0023-8333

Utterance-Final Lengthening Is Predictive of Infants' Discrimination of English Accents

Laurence White,[a] Caroline Floccia,[a] Jeremy Goslin,[a] and Joseph Butler[b]

[a]School of Psychology, Plymouth University, UK and [b]Laboratório de Fonética & Lisbon Baby Lab, Universidade de Lisboa, Portugal

Infants in their first year manifest selective patterns of discrimination between languages and between accents of the same language. Prosodic differences are held to be important in whether languages can be discriminated, together with the infant's familiarity with one or both of the accents heard. However, the nature of the prosodic cues that actually facilitate infant discrimination has not been directly examined. We analyzed the accent discrimination of 5- and 7-month-old British English infants, looking for durational features that could predict a range of discrimination results using the Headturn Preference Procedure. We previously found that both 5- and 7-month-olds based in Plymouth could discriminate Plymouth-accented and Welsh-accented English, while 5-month-olds could not discriminate Welsh-accented and Scottish-accented English. Most surprisingly, 7-month-olds failed to discriminate Plymouth-accented and French-accented English. From half of all utterances used in these four experiments, we calculated a range of durational metrics, both globally across utterances and locally at utterance-final edges. Utterance-edge metrics were the relative duration of the final consonant+vowel interval (nfinalCV) and the final vowel alone (nfinalV). Separately for 5- and 7-month-olds, we determined the difference in scores for all durational metrics between: (a) familiar training and novel test utterances and (b) familiar test and novel test utterances. Regression analyses showed unique predictors of discrimination. For 5-month-olds, the predictor was the difference in nfinalCV between familiar training and novel test utterances. For 7-month-olds, the predictor was the difference in finalV between familiar test and novel test utterances. This demonstrates an early sensitivity to precisely those localized timing features that are useful across languages for segmentation. We therefore predict that similar sensitivity should be shown in infants of other language backgrounds.

Keywords infant speech perception; accent discrimination; prosody; final lengthening

Correspondence concerning this article should be addressed to Caroline Floccia, School of Psychology, Plymouth University, Plymouth, UK PL4 8AA. E-mail: caroline.floccia@plymouth.ac.uk

DOI: 10.1111/lang.12060

Introduction

Prosodic features are informative about speech structure. Furthermore, the valency of much prosodic variation is language independent by default: sounds that are higher in pitch, louder, or longer are perceived as more salient. Thus, infants at an early stage of linguistic development may exploit prosody to analyze the structure of speech, in particular, for segmentation of speech into words and phrases (see Mattys, White, & Melhorn, 2005, for a review).

The primary focus regarding early infant segmentation has been on the use of stress (e.g., Höhle & Weissenborn, 2003; Johnson, 2005; Jusczyk, Houston, & Newsome, 1999). In English—the language of the infant learners in much foregoing research—the predominant location for lexical stress is the first syllable of the word, making metrical segmentation a useful if suboptimal strategy in the absence of word knowledge (Cutler & Carter, 1987).

Across the range of human languages to which infants are exposed, systems of linguistic prominence are, however, highly variable in terms of the existence of lexical stress, its distribution, and its phonetic correlates. Thus, metrical segmentation would be of marginal benefit, at best, for infant learners of many languages. By contrast, the marking of the edges of prosodic domains—words and phrases—by selective lengthening of speech sounds appears universal (e.g., Beckman, 1992). In particular, many languages have been shown to manifest substantial lengthening of segments, especially vowels, immediately preceding phrase and utterance boundaries (see Fletcher, 2010, for a review).

The use of lengthening as a cue to a preceding boundary accords with the iambic-trochaic law (Hayes, 1995), whereby, other things being equal, sounds made salient by virtue of higher pitch or greater loudness are perceived as sequence-initial, while sounds that are salient through greater duration are perceived as sequence-final. Although White, Mattys, Stefansdottir, and Jones (2014) showed that lengthened consonants tend to be perceived as word-initial rather than word-final, the use of vowel lengthening as a cue to an upcoming boundary appears to be a segmentation strategy with universal applicability (e.g., Saffran, Newport, & Aslin, 1996; Price, Ostendorf, Shattuck-Hufnagel, & Fong, 1991).

Gout, Christophe, and Morgan (2004) showed that American English–learning infants at 10 months old recognized a sequence of sounds corresponding to a disyllabic word better when the two syllables were not separated by a phrase boundary (e.g., *paper* vs *pay per*). Cues to this boundary included lengthening of the preboundary vowel and of the postboundary consonant. Materials eliminating the other naturally occurring cues to boundaries would be necessary to determine the salience of final lengthening alone for infants at this age.

Testing younger infants' sensitivity to final lengthening through their segmentation behavior is more problematic, however, as extraction of words from the speech stream has only reliably been shown in experimental studies from 6 months of age and may be focused on a very small number of familiar items, such as the child's name, baby terms for "mother," and so on (Bortfeld, Morgan, Golinkoff, & Rathbun, 2005; Tincoff & Juszcyk, 1999). Furthermore, early evidence for segmentation is inconsistent between cultures (see Nazzi, Mersad, Sundara, Iakimova, & Polka, 2013, comparing infants acquiring Canadian and Parisian French) and may, for example, arise several months later in children learning British English rather than American English (Floccia et al., 2014).

Rather than examining segmentation directly, early sensitivity to durational variation may also be evidenced through patterns of discrimination between and within languages, consistently demonstrated in the first year of life. Indeed, even neonates have been shown to discriminate between certain languages (Nazzi, Bertoncini, & Mehler, 1998). Furthermore, as Nazzi et al. used low-pass filtered speech and found that patterns of discrimination accorded with hypothesized rhythmic differences between languages, they proposed that prosodic cues, in particular rhythm and timing, might underlie the observed behavior.

Studies of adult language discrimination have focused even more narrowly on timing, minimizing segmental cues to language identity by using a restricted, formulaic set of phonemes to resynthesize the original signal. In the most widely used paradigm, all vowels are replaced by /a/ and all consonants by /s/, but the timing of the original speech is maintained: using such *sasasa* stimuli, Ramus, Dupoux, and Mehler (2003) showed similar patterns of adult between-language discrimination to those found for young infants.

Following this methodology, White, Mattys, and Wiget (2012) showed that British English adults could not only discriminate Castilian Spanish and Standard Southern British (SSB) English, two rhythmically different languages, but could also discriminate between certain varieties of British English. Additionally, White et al. used regression analyses to find which temporal properties of the experimental utterances were most predictive of discrimination behavior. In the Spanish versus SSB English comparison, the Spanish utterances had substantially higher mean speech rate, due, at least in part, to the lower phonological complexity of Spanish syllables. Speech rate proved to be the best predictor of discrimination: thus, Spanish utterances with high speech rate or English utterances with low speech rate were more likely to be categorized correctly than vice versa.

Where speech rate differences were eliminated through linear time-warping of utterances, other durational cues were salient. In particular, in two pairwise comparisons there were significant differences in the duration of the final consonant-plus-vowel interval (CV) relative to the mean pre-final CV duration. Thus, SSB English had much longer utterance-final CVs, relative to pre-final CVs, than did Spanish, and Welsh English had much more lengthening of utterance-final CVs than did Orkney English. These differences in the magnitude of final lengthening were found to be predictive: SSB English utterances were better discriminated from Spanish, and Welsh English utterances were better discriminated from Orkney, if they had longer utterance-final CV intervals (White et al., 2012).

There has been no direct investigation to date of the timing cues that infants use in language discrimination. To this end, we revisited our data from previous studies of discrimination between accents of English. Butler, Floccia, Goslin, and Panneton (2011) reported three experiments using an adaptation of the Headturn Preference Procedure (HPP; see Nazzi, Jusczyk, & Johnson, 2000) with 5-month-olds (Experiment 1 and Experiment 2) and 7-month-olds (Experiment 3). In each experiment, 20 infants born and raised in southwest England were presented with passages produced in one accent of English (e.g., Welsh English). During the test phase they heard an alternation of passages read in the same, familiarized, accent and passages read in a novel accent (e.g., Plymouth English). Passages were randomly presented on the left- or right-hand-side of the child, and looking times toward the source of sound were recorded and taken as evidence of interest in the corresponding passage. Any significant difference between looking times to the familiarized versus the novel accent was taken as evidence of discrimination between the two accents. Discrimination was found between Plymouth English and Welsh English for 5-month-olds (Experiment 1) and 7-month-olds (Experiment 3). However, 5-month-olds did not discriminate between Welsh English and Scottish English (Experiment 2).

In Butler et al. (2011), we interpreted these findings as showing early learning of native dialect properties, and we suggested that infants rely on a combination of segmental and suprasegmental information to perform the task, an inclusive interpretation informed by somewhat contradictory previous studies. First, Nazzi et al. (2000), who showed discrimination of British English and American English in 5-month-old American babies, argued for a strong role of prosodic cues: the British English utterances tended to have greater variability in syllable duration, together with more frequent high sentence-initial pitches and more terminal rises. Furthermore, American 6-month-old babies failed at the apparently easier discrimination of British English and Dutch when only presented with word lists (Jusczyk, Friederici, Wessels, Svenkerud, & Jusczyk, 1993), suggesting that sentence prosody is required over and above phonemic and phonotactic contrasts. However, Diehl, Varga, Panneton, Burnham, and Kitamura (2006) found that 6-month-old American infants did not show a preference for Australian English over American English when utterances were low-pass filtered at 400 Hz, thereby preserving utterance-level prosody but compromising lower-level (e.g., phonemic) information. This indicates that prosodic information alone may not be sufficient to maintain differential attention to a non native accent, although Kitamura, Panneton, and Best (2013) suggest that the most salient differences between Australian and American English accents are segmental rather than prosodic, also pointing to accent familiarity to explain differential discrimination results between American and Australian English 6-month-olds. Therefore, based on the studies just reviewed, a reasonable conclusion is that infants may utilize both prosodic and segmental cues for discrimination, with the balance of cues depending on the nature of the accent contrasts.

As discussed above, White et al. (2012), in an adult study, used regression analyses to examine the nature of the prosodic (specifically durational) cues that best predict discrimination. Thus, we decided to revisit Butler et al.'s (2011) data using White et al.'s battery of durational metrics, allowing us to examine more directly what kind of durational cues the infants could have extracted to distinguish between accents.

In addition to the Butler at al. (2011) series of experiments, we added an unpublished study (discussed in Butler, 2010), which examined 7-month-olds' discrimination of French-accented English versus the children's native Plymouth accent. The weight of evidence at the time led to an expectation of robust discrimination between French-accented and Plymouth-accented English. The two speech styles differ segmentally and suprasegmentally, with little overlap between the French and English vowel systems (Kitamura et al., 2013), together with salient differences at the consonant level (e.g., voice-onset-time and coronal fricative sounds; see Arslan & Hansen, 1996). In addition, French speakers find it notoriously difficult to produce English-like stress patterns (Dupoux, Pallier, Sebastian-Galles, & Mehler, 1997). Furthermore, informed by developmental and adult studies, we have argued elsewhere that discrimination between foreign and regional accents could be categorical rather than linear and based on perceptual distance (Goslin, Duffy, & Floccia, 2012;

Floccia, Goslin, Girard, & Konopczynski, 2006; Floccia, Butler, Goslin, & Ellis, 2009). Thus, the distinction between a native accent and a foreign accent would potentially be very salient in young infants. However, against these expectations, 7-month-olds failed to discriminate their Plymouth home accent from French-accented English (see Butler, 2010).

Thus, we examine whether there are durational cues that could account for not only infants' behavior in this surprising fourth experiment, but also in the three Butler et al. (2011) studies. The methods for Experiments 1–3 are reported in Butler et al., so we present in detail here the method only for Experiment 4, which tested 7-month-olds' discrimination between their Plymouth home dialect and French-accented English (very similar procedures were used in all four experiments). We then outline our method of calculation of durational metrics for the stimuli from all four experiments. Detailed discrimination results are presented for Experiment 4; analysis of the durational predictors of accent discrimination is presented for all four experiments.

Method

Participants

Twenty 7-month-old monolingual infants participated in this study (mean age = 7.3 months; range 6.8–8.1 months; 6 females, 14 males; all healthy by parental report; all were raised in the West Country region of England from birth). At least one parent in three cases originated from the north of England, whereas both parents of the 17 remaining children originated from either the West Country or the south of England. In all cases parents reported that the children had no significant exposure to French-accented speakers, that is, no more than occasionally (e.g., when traveling on vacation). Seven additional infants were excluded from the study: three due to crying or failure to pay attention to the lights or sounds used in the experiment and four because at least one of the parents originated from outside England.

Materials

The stimuli were similar to those used by Nazzi et al. (2000) and Butler et al. (2011). Eight passages, consisting of five unrelated sentences each (see Appendix) were recorded by four female speakers with a West Country accent (ages 20–30 years, all resident in Plymouth throughout their lives) and four native French female speakers (age 36–42, resident in Plymouth for 3–12 years). Each speaker recorded two of the eight passages, such that all passages were recorded by one speaker of each accent. To make the passages interesting to infants, the speakers were instructed to read them in typical infant-oriented style. Passages were recorded using a digital Dictaphone and microphone, using 16-bit, 44.1 kHz sampling rate. The average duration for the passages was 21.1 seconds (West Country passages: 20.6 seconds; French passages: 21.6 seconds).

An accent identification task was conducted with eight southern England native speakers naïve to the aim of the experiment (see Butler et al., 2011, for details). It was found that French-accented passages were identified as such with a mean accuracy of 98.4% (range 87.5–100%) and with a mean confidence of 3.88 (on a scale ranging from 1 to 4, where 1 = *no confidence* and 4 = *very high confidence*). Similarly, participants identified the West Country passages in 98.4% (87.5–100%) of cases, with a mean confidence of 3.13. The West Country recordings of the passages were those also used in Butler et al. (2011, Experiments 1 and 3).

Procedure

After obtaining full ethical consent from the parents, accent discrimination responses were collected using an adaptation of the HPP. Infants were seated on their caregiver's lap in the center of the test booth. At the beginning of each trial, a flashing green light was presented at the front of the booth to focus infants' attention to the middle of the test area. This green light was then turned off and replaced with a flashing red light, which could either be to the left or right side of the booth. The location of the red light was chosen on a pseudorandom basis, such that the light could not appear on the same side for more than two consecutive trials. Once the infant turned to look at the flashing light, one of the recorded passages was played from a speaker next to the light (the red light continued flashing during the presentation of the passage). If the passage ended, or the infant looked away from the light for more than 2 seconds, then all lights and sounds were terminated and after a few seconds a new trial began. Control of lights, speech recordings and the monitoring of the infant's looking times were all synchronized and remotely controlled by the experimenter using a computer program. The infant's caregiver wore headphones playing music during the experiment so that she/he was unaware of the accent of the speech stimuli presented to the infant. The experimenter sat outside the booth and monitored the child's looking behavior on a monitor with the sound turned off.

During the familiarization phase, each infant was presented with a particular accent of English, using four passages from two speakers of that accent. Each infant was required to accumulate a total of 20 seconds of looking time to each of the passages during familiarization. Half of the infants were familiarized with West Country-accented passages, and the other half were familiarized with French-accented passages. Once this time-locked familiarization was complete, the test phase began, with a randomly ordered presentation of four West Country accent and four French accent passages (two speakers of each accent; neither the passages nor the speakers were those heard during familiarization). Passages were terminated whenever the infant looked away for more than 2 seconds or when the five-utterance passage ended.

Extraction of Durational Metrics

The durational metrics used as candidate predictors of language discrimination were extracted from half of the passages recorded by each speaker and each accent of English (Plymouth, Welsh, Scottish, and French). The measured passages were those corresponding to the sentences in the A blocks (see the Appendix), thus derived from 20 different sentences in four blocks (1A, 2A, 3A, 4A). Each block of sentences (e.g., *The young boy got up quite early in order to watch the sun rise*) had been produced by one speaker of each of the four accents, with four speakers of each accent in total. Thus, in total, 80 different utterances were annotated.

Following the procedure of White and Mattys (2007) and White et al. (2012), we identified the boundaries between consonant and vocalic intervals based on visual inspection of the waveform and spectrogram in Praat (Boersma & Weenink, 2010), occasionally relying on the auditory signal. Immediately adjacent vowels and immediately adjacent consonants were treated as belonging to the same interval. Pauses were excluded and, where intervals of the same type (vocalic or consonantal) were separated by a pause, these intervals were summed. Following annotation, the duration of each vocalic and consonant interval was extracted automatically and for each utterance a range of durational metrics was calculated (see Table 1, reproduced from White et al., 2012).

Regression Analysis of Predictors of Discrimination

Due to the counterbalancing of speakers and passages, the subset of stimuli heard within each experiment varied between infants. For example, in Experiment 1, some infants heard French-accent speakers 1 and 2 during familiarization and French-accent speakers 3 and 4, alternated with Plymouth-accent speakers 3 and 4, during the test. We estimated the distance between the familiarized accent and the novel one, as experienced by the infant, in two ways. DFT refers to the distance between the passages heard during the familiarization (familiarized accent) and those heard during the test (novel accent). DT refers to

Table 1 Durational metrics derived from the experimental utterances: Mean (and standard error) of each metric for each accent type (reproduced from White et al., 2012)

	Plym	Welsh	Scot	French	
%V	**45.9**	**44.5**	**44.5**	**45.0**	Percentage of utterance
	1.4	0.6	1.4	2.0	duration comprised of vocalic intervals
ΔC	**73.5**	**72.7**	**78.3**	**74.5**	Standard deviation of
	8.5	9.0	4.5	5.9	consonantal intervals
ΔV	**63.9**	**60.6**	**65.1**	**63.1**	Standard deviation of vocalic
	4.8	3.3	4.2	4.5	intervals
nPVI-V	**72.5**	**74.6**	**71.7**	**68.4**	Normalized PVI for vocalic
	3.1	2.2	2.7	1.6	intervals. Mean of the differences between successive intervals divided by their sum (×100)
rPVI-C	**81.5**	**79.1**	**91.3**	**84.4**	PVI for consonantal intervals.
	7.3	9.5	7.2	8.9	Mean of the differences between successive intervals
MeanV	**115.7**	**117.3**	**116.9**	**110.9**	Mean duration of vocalic
	6.6	8.4	5.1	4.6	intervals (ms)
MeanC	**130.4**	**128.2**	**139.1**	**128.9**	Mean duration of consonantal
	11.7	10.0	7.8	8.6	intervals (ms)
VarcoV	**56.5**	**52.3**	**56.0**	**56.6**	Standard deviation of vocalic
	1.6	1.3	3.1	3.5	intervals divided by the mean (×100)
VarcoC	**56.1**	**56.2**	**56.0**	**57.9**	Standard deviation of
	2.2	3.2	2.0	3.0	consonantal intervals divided by the mean (×100)
rPVI-CV	**117.1**	**106.6**	**110.4**	**132.6**	PVI for consonant + vowel
	10.2	7.6	11.6	4.9	intervals. Mean of the differences between successive intervals
nFinalC	**1.14**	**0.99**	**0.94**	**0.96**	Duration of final consonantal
	0.09	0.12	0.11	0.08	interval divided by the mean consonantal interval duration for the utterance
nFinalV	**1.69**	**1.52**	**1.56**	**1.64**	Duration of final vocalic interval
	0.05	0.21	0.17	0.11	divided by the mean vocalic interval duration for the utterance
nFinalCV	**1.41**	**1.27**	**1.22**	**1.30**	Duration of final consonant +
	0.05	0.16	0.13	0.09	vowel interval divided by the mean consonant + vowel interval duration for the utterance
Speech rate	**4.21**	**4.26**	**4.03**	**4.34**	Number of syllables per second
	0.31	0.24	0.15	0.20	

Note. For a mathematical definition of the Pairwise Variability Index (PVI), see Grabe and Low (2002).

the distance between the passages heard during the test phase only (familiarized accent versus novel accent). For each infant, for each of the durational metrics (Table 1), we calculated the normalized difference between the familiarized accent and the novel accent, both DFT and DT, that is, we calculated the mean scores for all relevant utterances from each accent and took the absolute value of the difference between the two accents.

The dependent variable in the regression analyses was each infant's discrimination score, that is, for 7-month-olds, looking time (novel - familiar); for 5-month-olds, looking time (familiar - novel); see below for discussion of the reversed direction of discrimination responses in 5- and 7-month-olds. These looking time differences were normalized by dividing them by the average looking times for the novel and familiarized accents.

Before injecting individual discrimination scores and durational metrics in a stepwise regression analysis, we inspected data for collinearity and rejected through an iterative process every variable with the highest variance inflation factor (VIF) until no variable showed a VIF greater than 5 (Rakotomalala, 2011). Because of their different discrimination responses (see below), we ran regression analyses separately for 5-month-olds (Experiments 1 and 2) and 7-month-olds (Experiments 3 and 4).

Results

Discrimination Performance in Experiment 4: Plymouth- vs French-Accented English

Overall, the average looking time was 5.65 seconds ($SD = 1.42$) to the novel dialect and 6.08s to the familiarized dialect ($SD = 1.98$). Of the 20 infants tested, 7 had longer looking times for the new dialect over the familiarized dialect.

A repeated-measures analysis of variance was carried out on looking times, with a within-participants variable of accent status (familiar vs. novel) and a between-participants variable of order of familiarization (Plymouth first vs. French first). Evidence for discrimination would be obtained with a significant effect of accent status, with looking times expectedly longer for the novel accent versus the familiarized accent (see Butler et al., 2011). However, there was no effect of accent status, $F(1, 18) < 1$, no effect of order of familiarization, $F(1, 18) < 1$, and no interaction between the two variables, $F(1,18) < 1$.

Figure 1 shows the mean looking times for novel versus familiarized accents for this experiment (Experiment 4), together with those for the related experiments reported in Butler et al. (2011). The direction of the expected discrimination effect is reversed for Experiments 1 and 2, which tested

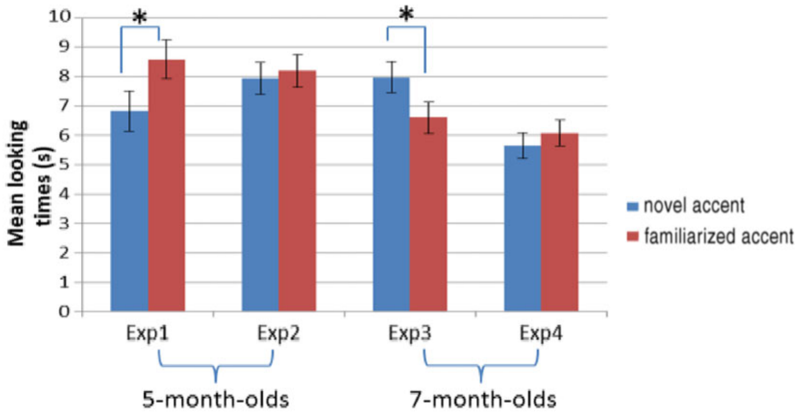

Figure 1 Mean looking times (seconds) for the novel accent (blue) and the familiarized accent (red) during test phase. Significant differences are indicated by asterisks. Experiment 1 contrasts Plymouth and Welsh accents in 5-month-olds; Experiment 2 contrasts Welsh and Scottish accents in 5-month-olds; Experiment 3 contrasts Plymouth and Welsh accents in 7-month-olds; Experiment 4 contrasts Plymouth and French accents in 7-month-olds. Experiments 1–3 are from Butler et al. (2011).

5-month-olds, versus Experiments 3 and 4, which tested 7-month-olds. That is, in 5-month-olds, discrimination is found with longer looking times for the familiarized accent (familiarity effect, Exp1) whereas 7-month-olds prefer to listen to novel accent (novelty effect, Exp3). This reversal accords with Houston-Price and Nakai's (2004) review of the factors influencing novelty versus familiarity effects in HPP-related procedures. Habituation becomes faster as children age, which would lead to a better representation of the familiarized accent in 7-month-olds, which in turn leads to a greater reaction to novelty.

Predictors of 5-Month-Olds' Discrimination Scores for Experiments 1 and 2

For the set of DFT metrics, six variables were iteratively removed due to VIF values above 5 (meanC, nfinalC, rate, nfinalV, nPVI-V, and rPVI-CV). With the remaining eight variables, a stepwise regression analysis gave nfinalCV as the unique significant predictor of infants' discrimination scores ($\beta = 1.32$, $p = .014$, $R^2 = .15$). Thus, as illustrated in Figure 2, the greater the difference in nfinalCV between the familiarized accent and the novel one, the greater the infants' discrimination response, where nfinalCV is the duration of final consonant+vowel interval divided by the mean consonant+vowel interval duration for the utterance.

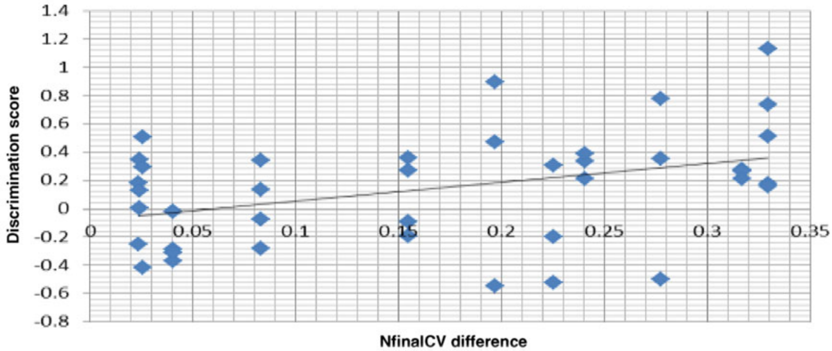

Figure 2 Correlation between the nfinalCV distance (familiar training *vs* novel test) and 5-month-olds' discrimination scores (Experiments 1 and 2 combined).

For the set of DT metrics, six variables were iteratively removed due to VIF values above 5 (rPVI-C,%V, ΔC, speech rate, nfinalCV, and nfinalC). With the remaining eight variables, a stepwise regression analysis failed to provide any significant predictor of discrimination scores.

Predictors of 7-Month-Olds' Discrimination Scores for Experiments 3 and 4
For the set of DFT metrics, five variables were iteratively removed due to VIF values above 5 (rPVI-CV, nfinalCV, speech rate, nfinalC, and ΔC). With the remaining nine variables, a stepwise regression analysis failed to provide any significant predictors of discrimination scores.

For the set of DT metrics, six variables were first removed due to VIF values above 5 (nfinalCV, ΔC,%V, speech rate, ΔV, and meanC). With the remaining eight variables, a stepwise regression analysis gave nfinalV as the unique significant predictor of infants' discrimination scores ($\beta = 1.27$, $p =$ .038, $R^2 = .11$). Thus, as illustrated in Figure 3, the greater the difference in nfinalV between the familiarized test accent and the novel one, the greater the infants' discrimination response, where nfinalV is the duration of the final vowel interval divided by the mean vowel interval duration for the utterance.

Discussion

Revisiting a series of language discrimination experiments, we showed that 5-month-old infants raised in Plymouth could distinguish Plymouth- from Welsh-accented English, but not Plymouth- from French-accented English. At 7 months old, infants could still distinguish their home accent from Welsh

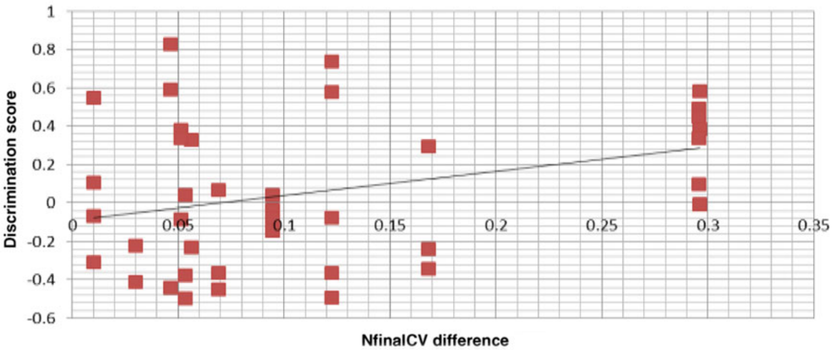

Figure 3 Correlation between the nfinalV distance (familiar test *vs* novel test) and 7-month-olds' discrimination scores (Experiments 3 and 4 combined).

English, but could not discriminate Welsh and Scottish accents of English (Butler, 2010; Butler et al., 2011).

Seeking to identify prosodic predictors of these patterns of discrimination, we calculated a number of metrics of global and local durational variation. In line with adult experiments where speech rate was controlled (White et al., 2012), we found that relative lengthening of the utterance-final consonant+vowel sequence was predictive of discrimination for 5-month-olds, while relative lengthening of the final vowel alone predicted 7-month-olds' ability to distinguish two accents of English. These results suggest that even infants of this age are sensitive to localized durational cues to prosodic structure, which accords with their potential importance for segmenting the speech stream into words and phrases, thereby bootstrapping lexical acquisition.

The power of the final lengthening cues is impressive given the relatively crude nature of the metric. We simply took the ratio of the final interval (vowel or vowel+consonant) relative to the mean of the same interval across the utterance (excluding the final interval for the mean). No account was taken of the phonological or metrical structure of the final words in the utterance. Thus, for a stressed open vowel like the final one in *ago*, the nfinalV metric would well reflect the large degree of lengthening expected at the end of an utterance. This would be somewhat diminished in a closed syllable such as *rise*, where the coda consonant might be expected to absorb some of the lengthening, while an unstressed final syllable, such as in *city*, should be longer than in phrase-medial position, but probably not longer than the mean of all non-final vowel durations. A more sophisticated measure might take metrical structure into account, given that final lengthening tends to begin with the final stressed vowel and may increase with proximity to the boundary (see White & Turk, 2010, for a discussion of the locus of final lengthening effects).

Thus, the fact that these noisy metrics are useful predictors reinforces the salience of utterance-final prosodic phenomena for infants. Of course, this study only considered durational cues to language discrimination, and other prosodic parameters might well be found to be perceptually important for infants in such tasks, together with measures of segmental distance. However, the results reported here do accord with expectations given the potential universality of preboundary lengthening as a segmentation cue. Additionally, the phonetic salience conferred by final lengthening is likely to be a factor in the finding that 8-month-old American English–learning infants recognize words that are previously presented at utterance edges (initial or final) better than those presented utterance-medially (Seidl & Johnson, 2006).

Finally, considering the nature of evidence in language discrimination studies, it has often been argued that patterns of discrimination accord with so-called rhythm classes. Thus, infants are held to be able to distinguish stress-timed English from syllable-timed Spanish, but not two syllable-timed or two stress-timed languages (see White et al., 2012, for a review), unless they are native learners of one of the two rhythmically similar languages (e.g., Bosch & Sebastián-Gallés, 1997; Nazzi et al., 2000). Of course, even if there are meaningful prosodic classes, these differences must be realized in the speech signal, potentially through local suprasegmental effects such as final lengthening. Moreover, White, Delle Luche, and Floccia (2014) have shown discrimination between syllable-timed French and Spanish by 5-month-olds, arguing against existing notions of intrinsic rhythm class and prompting the prediction that discrimination should also be observed between other accent or language pairs where there are reliable and perceptible differences in local timing. Furthermore, given the apparently universal nature of final lengthening in spoken language, such early sensitivity to these localized durational cues would be expected in infants regardless of their native language background.

References

Arslan, L. M., & Hansen, J. H. (1996). Language accent classification in American English. *Speech Communication*, *18*, 353–367.

Beckman, M. E. (1992). Evidence for speech rhythms across languages. In Y. Tohkura, E. Vatikiotis-Bateson, & Y. Sagisaka (Eds.), *Speech perception, production and linguistic structure* (pp. 457–463). Oxford, UK: IOS Press.

Boersma, P., & Weenink, D. (2010). Praat: Doing phonetics by computer (Version 4.4.24) [Computer software]. Available at http://www.praat.org/

Bortfeld, H., Morgan, J. L., Golinkoff, R. M., & Rathbun, K. (2005). Mommy and me: Familiar names help launch babies into speech-stream segmentation. *Psychological Science*, *16*, 298–304.

Bosch, L., & Sebastián-Gallés, N. (1997). Native-language recognition abilities in 4-month-old infants from monolingual and bilingual environments. *Cognition*, *65*, 33–69.

Butler, J., Floccia, C., Goslin, J., & Panneton, R. (2011). Perception of regional accents in 5 to 7 months old British infants. *Infancy*, *16*, 392–417.

Butler, J.P. (2010). Perception of accents and dialects in adults and infants. *Unpublished Plymouth University PhD dissertation*.

Cutler, A., & Carter, D. M. (1987). The predominance of strong initial syllables in the English vocabulary. *Computer Speech & Language*, *2*, 133–142.

Diehl, M., Varga, K., Panneton, R., Burnham, D., & Kitamura, C. (2006, June). *Six-month-old infants' perception of native speech accent*. Paper presented at the annual meeting of the XVth Biennial International Conference on Infant Studies, Kyoto, Japan.

Dupoux, E., Pallier, C., Sebastian-Galles, N., & Mehler, J. (1997). A destressing "deafness" in French? *Journal of Memory and Language*, *36*, 406–421.

Fletcher, J. (2010). The prosody of speech: Timing and rhythm. In W. Hardcastle, J. Laver, & F. Gibbon (Eds.), *The handbook of phonetic sciences* (pp. 523–602). Chichester, UK: Wiley-Blackwell.

Floccia, C., Butler, J., Goslin, J., & Ellis, L. (2009). Regional and foreign accent processing in English: Can listeners adapt? *Journal of Psycholinguistic Research*, *38*, 379–412.

Floccia, C., Goslin, J., Girard, F., & Konopczynski, G. (2006). Does a regional accent perturb speech processing? *Journal of Experimental Psychology: Human Perception and Performance*, *32*, 1276–1293.

Floccia, C., Keren-Portnoy, T., Duffy, H., Delle Luche, C., Durrant, S., White, L., et al. (2014). British-English learning infants fail in word segmentation tasks. Manuscript in preparation.

Goslin, J., Duffy, H., & Floccia, C. (2012). An ERP investigation of regional and foreign accent processing. *Brain and Language*, *122*, 92–102.

Gout, A., Christophe, A., & Morgan, J. L. (2004). Phonological phrase boundaries constrain lexical access II. Infant data. *Journal of Memory and Language*, *51*, 548–567.

Grabe, E., & Low, E. L. (2002). Durational variability in speech and the rhythm class hypothesis. In N. Warner & C. Gussenhoven (Eds.), *Papers in laboratory phonology 7* (pp. 515–546). Berlin, Germany: Mouton de Gruyter.

Hayes, B. (1995). *Metrical stress theory: Principles and case studies*. Chicago: University of Chicago Press.

Höhle, B., & Weissenborn, J. (2003). German-learning infants' ability to detect unstressed closed-class elements in continuous speech. *Developmental Science*, *6*, 122–127.

Houston-Price, C., & Nakai, S. (2004). Distinguishing novelty and familiarity effects in infant preference procedures. *Infant and Child Development*, *13*, 341–348.

Johnson, E. K. (2005). English-learning infants' representations of word forms with iambic stress. *Infancy*, *7*, 99–109.

Jusczyk, P. W., Friederici, A. D., Wessels, J., Svenkerud, V. Y., & Jusczyk, A. M. (1993). Infants' sensitivity to the sound patterns of native language words. *Journal of Memory and Language*, *32*, 402–420.

Jusczyk, P. W., Houston, D. M., & Newsome, M. (1999). The beginnings of word segmentation in English-learning infants. *Cognitive Psychology*, *39*, 159–207.

Kitamura, C., Panneton, R., & Best, C.T. (2013). The development of language constancy: attention to native versus nonnative accents. *Child Development*, *84*, 1686–1700.

Mattys, S. L., White, L., & Melhorn, J. F. (2005). Integration of multiple speech segmentation cues: A hierarchical framework. *Journal of Experimental Psychology: General*, *134*, 477–500.

Nazzi, T., Bertoncini, J., & Mehler, J. (1998). Language discrimination by newborns: Towards an understanding of the role of rhythm. *Journal of Experimental Psychology: Human Perception and Performance*, *24*, 756–766.

Nazzi, T., Jusczyk, P. W., & Johnson, E. K. (2000). Language discrimination by English learning 5 month olds: Effects of rhythm and familiarity. *Journal of Memory and Language*, *43*, 1–19.

Nazzi, T., Mersad, K., Sundara, M., Iakimova, G., & Polka, L. (2013). Early word segmentation in infants acquiring Parisian French: task-dependent and dialect-specific aspects. *Journal of Child Language*, *41*, 600–633.

Price, P. J., Ostendorf, M., Shattuck-Hufnagel, S., & Fong, C. (1991). The use of prosody in syntactic disambiguation. *Journal of the Acoustical Society of America*, *90*, 2956–2970.

Rakotomalala, R. (2011). *Pratique de la regression lineaire multiple. Diagnostic et selection de variables* [PDF document]. Retrieved from Online Lecture Notes at http://eric.univ-lyon2.fr/˜ricco/cours/cours/La_regression_dans_la_pratique.pdf

Ramus, F., Dupoux, E., & Mehler, J. (2003). The psychological reality of rhythm classes: Perceptual studies. In *Proceedings of the 15th International Congress of Phonetic Sciences, Barcelona* (pp. 337–342). Barcelona: ICPhS Organizing Committee, 2003.

Saffran, J. R., Newport, E. L., & Aslin, R. N. (1996). Word segmentation: The role of distributional cues. *Journal of Memory and Language*, *35*, 606–621.

Seidl, A., & Johnson, E. K. (2006). Infant word segmentation revisited: Edge alignment facilitates target extraction. *Developmental Science*, *9*, 565–573.

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

White, L., Delle Luche, C., & Floccia, C. (2014). Language discrimination within a "rhythm class" by five-month-old infants. Manuscript in preparation.

White, L., & Mattys, S. L. (2007). Calibrating rhythm: First language and second language studies. *Journal of Phonetics*, *35*, 501–522.

White, L., Mattys, S. L., Steffansdottir, L., & Jones, V. (2014). Lengthened consonants are interpreted as word-initial. In N. Campbell, D. Gibbon, & D. Hirst (Eds.), *Proceedings of the VIIth International Conference on Speech Prosody, Dublin: TCD* (pp. 477–481).

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

White, L., & Turk, A.E. (2010). English words on the Procrustean bed: Polysyllabic shortening reconsidered. *Journal of Phonetics*, *38*, 459–471.

Appendix

Passages Recorded for the Discrimination Task

1A	The young boy got up quite early in order to watch the sun rise.
	This supermarket had to close due to economic problems.
	The committee will meet this afternoon for a special debate.
	Having a big car is not something I would recommend in this city.
	Mothers usually leave the maternity unit two days after giving birth.
1B	The next local elections will take place during the winter.
	Some more money will be needed to make this project succeed.
	Artists have always been attracted by the life in the capital.
	Your welcome speech will be delivered without the press office's agreement.
	The latest events have caused an outcry in the international community.
2A	The local train left the station more than five minutes ago.
	The first flowers have bloomed due to the exceptional warmth of March.
	Trade unions have lost a lot of their influence during the last ten years.
	The green parties unexpectedly gained strong support from middle class people.
	This is the first time an international exhibition takes place in this town.
2B	In this case the easier solution seems to appeal to the court.
	The last concert given at the opera was a tremendous success.
	They didn't hear the good news until last week on their visit to their friends.
	This year's Chinese delegation was not nearly as impressive as last year's.
	In spite of technical progress predicting the weather is still very difficult.
3A	The art gallery in this street was opened only last week.
	In this famous coffee shop you will eat the best doughnuts in town.
	Most European banks close extremely early on Friday afternoons.
	The government is planning a reform of the educational program.
	The recent rainfall has caused very severe damage in the higher valleys.
3B	A hurricane was announced this afternoon on the TV.
	This rugby season promises to be a very exciting one.
	Science has acquired an important place in western society.
	The rebuilding of the city started the very first day after the earthquake.
	It is getting very easy nowadays to find a place in a nursery school.
4A	My grandparents' neighbour is the most charming person I know.
	Nobody noticed when the children slipped away just after dinner.
	The library is open every day from 8am to 6pm.
	The city council has decided to renovate the medieval centre.
	Seven paintings of great value have recently been stolen from the museum.
4B	The parents quietly crossed the dark room and approached the boy's bed.
	Finding a job is difficult in the present economic climate.
	There is an important market twice a week on the main square of the village.
	The woman over there is an eminent specialist in plastic surgery.
	Most of the supporters of the football club had to travel for an entire day.