J. Child Lang. **38** (2011), 606–627.
doi:10.1017/S0305000910000243

Does size matter? Subsegmental cues to vowel mispronunciation detection*

NIVEDITA MANI
University of Göttingen

AND

KIM PLUNKETT
University of Oxford

(*Received* 17 *November* 2008 – *Revised* 28 *September* 2009 – *Accepted* 18 *April* 2010 – *First published online* 1 *November* 2010)

ABSTRACT

Children look longer at a familiar object when presented with either correct pronunciations or small mispronunciations of consonants in the object's label, but not following larger mispronunciations. The current article examines whether children display a similar graded sensitivity to different degrees of mispronunciations of the vowels in familiar words, by testing children's sensitivity to 1-feature, 2-feature and 3-feature mispronunciations of the vowels of familiar labels: Children aged 1;6 did not show a graded sensitivity to vowel mispronunciations, even when the trial length was increased to allow them more time to form a response. Two-year-olds displayed a robust sensitivity to increases in vowel mispronunciation size, differentiating between small and large mispronunciations. While this suggests that early lexical representations contain information about the features contributing to vocalic identity, we present evidence that this graded sensitivity is better explained by the acoustic characteristics of the different mispronunciation types presented to children.

INTRODUCTION

During the second year of life, infants demonstrate comprehension of a substantial repertoire of words. The average infant aged 1;0 knows as many as 80 words, a number which increases rapidly to around 500 words by the

[*] This research was supported by an ESRC grant RES-000-23-1322 awarded to Kim Plunkett. Address for correspondence: Nivedita Mani, Free-Floater (Junior) Research Group, 'Language Acquisition', University of Göttingen, Gosslerstrasse, 14, 37073 Göttingen. e-mail: nmani@gwdg.de

end of her second year. Given the apparent difficulties involved in learning new words, the rapid increase during this period is remarkable, and has led to a number of studies questioning the robustness of children's representations of the words learned in this early phase of vocabulary development. A prominent study by Stager & Werker (1997) reported that children at 1;2 cannot simultaneously learn two words that differ by a single consonant (e.g. *bih–dih*). The authors concluded that the complications inherent in word learning may cause children to fail to pay attention to the phonetic detail of early words: although words may be phonologically well represented in the lexicon, children may not be able to access these representations early on, due to the cognitive demands imposed during a word learning task. Similarly, Swingley & Aslin (2007) have argued that children aged 1;6 have difficulty learning novel words that sound similar to familiar words.

Other research, however, indicates that children can access the phonological detail of FAMILIAR words (Swingley & Aslin, 2000, 2002; Bailey & Plunkett, 2002; Mani & Plunkett, 2007): as early as age 0;10 (Mani & Plunkett, 2008a) children can differentiate between correct pronunciations and minimal mispronunciations of the vowels and consonants of familiar monosyllabic words. In these studies, children are presented with two images of familiar objects, followed by a label for one of the objects. The label is either correctly or incorrectly pronounced, where mispronunciations change a single phonological feature of either the word-initial consonant (e.g. *book–dook*) or the word-medial vowel (e.g. *book–bok*). Children look longer and more quickly at the object when it is correctly labelled than when it is mispronounced. These results suggest that children possess phonologically well-specified representations of familiar words and that they can readily access these representations. This level of phonetic detail does not appear to be restricted to children's representations of familiar words, but extends to their representations of the vowels (with 3-feature mispronunciations; Mani & Plunkett, 2008b) and consonants in newly learned words (Ballem & Plunkett, 2005). Mani & Plunkett (2007; 2008b) suggest that children's sensitivity to vowel and consonant mispronunciations of familiar and newly learned words provides evidence for a symmetry in the specification of vowels and consonants in lexical entries early in life.

Some mispronunciations, however, appear to be more salient compared to others. Mani, Coleman & Plunkett (2008) report that children's sensitivity to different kinds of mispronunciations of vowels depends on the type of vowel changes involved: children at 1;6 are more sensitive to mispronunciations of vowel height and vowel backness, compared to mispronunciations of vowel roundedness, suggesting that height and backness are well specified in Southern British English. This is unsurprising, since specification of vowel roundedness is relatively redundant due to the strong correlation between vowel backness and roundedness in English.

One might also expect differences in children's sensitivity to small and large mispronunciations – children may be more sensitive to mispronunciations which change many features of either the vowel or the consonant compared to mispronunciations which change only one feature. A recent study provides evidence in support of this degree of specification of subphonemic consonantal features in children's lexical representations (White & Morgan, 2008; henceforth referred to as W&M (2008)). Children aged 1;7 were presented with an image of a familiar object and a novel object side by side on a screen, followed by a label for one of the images. The label for the novel image was a word that children were not expected to know (e.g. *barrel*). The familiar label, on the other hand, was either correctly pronounced or mispronounced. Mispronunciations changed one feature (place of articulation), two features (place and voicing) or three features (place, voicing and manner) of the word-initial consonant. W&M (2008) reported a significant linear trend in children's sensitivity to mispronunciations, with children being more sensitive to 3-feature mispronunciations compared to 2-feature mispronunciations, which in turn were more salient than 1-feature mispronunciations.[1]

This finding was surprising, since previous studies report no systematic differences in children's sensitivity to 1- and 2-feature mispronunciations of consonants in familiar words (Bailey & Plunkett, 2002). However, Bailey & Plunkett presented children at 1;6 and 2;0 with images of two familiar objects, followed by correct pronunciations, 1-feature or 2-feature mispronunciations of the label for one of the familiar images. W&M (2008) argue that this may be confusing, since the labels for both images are known to the children. The mispronunciation, therefore, does not match either image. In contrast, presenting children with an image of a familiar and novel object is compatible with the notion that the mispronunciation is a label for the novel object (The Principle of Mutual Exclusivity; Halberda, 2003), and may provide a more reliable estimate of children's sensitivity to different kinds of mispronunciations.

Given that children are sensitive to variations in the size of mispronunciations of consonants in familiar words (W&M, 2008), are children similarly sensitive to variations in the size of vowel mispronunciations? As mentioned above, Mani & Plunkett (2007) provide evidence that there is a symmetry in children's sensitivity to vowel and consonant mispronunciations of familiar words. Some studies even suggest that vowel changes may be

[1] Similar results are found in studies comparing adults' sensitivity to minimal (approximately one feature) and larger (5-feature) changes to the phonemes in a word (Connine, Titone, Deelman & Blasko, 1997). However, this sensitivity is restricted by the phoneme under investigation. Adults differentiated between minimal and maximal changes to the phoneme /t/ in a word. However, a similar difference was not found between minimal and maximal changes to the phoneme /k/.

more salient than consonant changes. Curtin, Fennell & Escudero (2009) report that English children aged 1;1 can simultaneously learn some words that differ by a single vowel (i.e. the vowel change from /i/ to /ɪ/). In contrast, there is currently no evidence suggesting that English children can simultaneously learn two words differing by a single consonant at this age or, indeed, until 1;5 (Werker, Fennell, Corcoran & Stager, 2002). Similarly, infants aged 0;6 can discriminate between native and non-native language vowels, while a similar sensitivity to native language consonants is displayed only between 0;9 and 1;0 (Kuhl, Williams, Lacerda, Stevens & Lindblom, 1992; Werker & Tees, 1984). The native vocalic repertoire appears to be in place earlier than the consonantal range. Furthermore, Gerken, Murphy & Aslin (1995) found that preschoolers were more sensitive to vowel changes in bisyllabic words than consonant changes, though no differences were found for monosyllabic words. Given the prominence of vowels in phonological acquisition, one might expect children to focus on the featural or acoustic detail differentiating vowels early in life. In keeping with this view, therefore, one might suppose that children would be at least equally if not more sensitive to variations in the size of vocalic than consonantal mispronunciations.

In contrast to this view, however, Nazzi and colleagues (2005; Nazzi, Floccia, Moquet & Butler, 2009; Havy & Nazzi, 2009) report that French children at 1;8 more easily categorize objects whose labels differ by a single consonant (*duk–guk*) compared to a single vowel (*duk–dok*). Nazzi's results suggest that consonants may be more important for lexical acquisition compared to vowels, and perhaps more robustly represented in children's lexical representations compared to vowels. Caramazza, Chialant, Capasso & Micell (2000) found a double dissociation between the processing of vowels and consonants in aphasic patients, and argued that this demonstrated categorically distinct representations of vowels and consonants that could not be reduced to a featural level.

Nazzi's and Caramazza *et al.*'s results support typological analysis by Nespor, Pena & Mehler (2003) suggesting 'that the task of distinguishing lexical items rests more on consonants than on vowels' (p. 209). Nespor *et al.* argue that consonants are specialized for conveying information about the lexicon whereas vowels provide information about prosody and grammar. From this perspective, one might expect children to show less sensitivity to variations in the pronunciations of vowels compared to consonants.

There are also differences in the acoustic and articulatory characteristics of vowels and consonants that might lead to differences in children's sensitivity to variations in the size of vocalic and consonant mispronunciations. Consonants, on the one hand, are usually described in terms of their place and manner of articulation and voicing. Consonant features are categorically distinct from each other, i.e. a change in place of articulation need not involve any change in the voicing or manner of articulation of the consonant. Vowels, on the other hand, are typically described in terms of the position of the tongue and shape of the mouth during articulation, providing three main dimensions of variation: vowel height, backness and roundedness. These dimensions are not completely distinct from each other, i.e. a change in vowel height may also cause a small change in vowel roundedness (*book–bok*), or a change in vowel backness may also cause a small change in vowel height (e.g. in Southern British English: *bed–bud*). Consequently, there may not be as clear a separation between different sizes of vowel mispronunciations as may be the case with consonant changes. Given this lack of distinctiveness of vocalic feature changes, would children show a graded sensitivity to an increase in the number of vocalic features contributing to the mispronunciation, as has previously been shown with consonants (W&M, 2008)?

Experiment 1 examines children's sensitivity to differences in the sizes of vowel mispronunciations of familiar words at ages 1;6 and 2;0. We employ the W&M (2008) modification of the standard infant testing paradigm in which a familiar object is paired with an unfamiliar object. We present children with 1-, 2- and 3-feature mispronunciations of the vowels in familiar words, where a 1-feature mispronunciation changes the height, backness or roundedness of the vowel. A 2-feature mispronunciation changes the height and backness, height and roundedness, or backness and roundedness of the vowel. Finally, a 3-feature mispronunciation changes all three vowel features. Comparison of children's performance in the three mispronunciation conditions permits an initial test of the psychological reality of vocalic phonological features. Note that changes to phonological features naturally lead to changes in the acoustic characteristics of the mispronunciation. If children show a graded sensitivity to vowel mispronunciations, there are a number of further issues to consider. First, we must ask whether this graded sensitivity is driven by the acoustic or phonological characteristics of the mispronunciation. Second, by comparing children's performance at 1;6 and 2;0, we test whether there are any developmental differences in children's sensitivity to the size of vocalic changes presented.

EXPERIMENT 1: CHILDREN AT 1;6 AND 2;0

METHOD

Participants

The participants in this experiment were twenty-seven children at 1;6 (*M* = 1;5·27, Range = 1;5·6 to 1;6·1) and twenty-seven children at 2;0 (*M* = 1;11·25, Range = 1;11·6 to 2;0·12). Ten additional children were tested but were excluded due to fussiness, parental interference or experimenter

TABLE 1. *Stimuli presented to children*

Target label	Mispronunciation	Feature change	Acoustic change	% comprehended 1;6	% comprehended 2	Type	Novel label
Bed	Bud /bʌd/	B	198	83	100	1-feature	Kig
Bib	Beb /bɛb/	H	253	76	96	1-feature	Daz
Bread	Brud /brʌd/	B	246	71	92	1-feature	Daz
Duck	Dock /dɔk/	R	206	93	96	1-feature	Bint
Fish	Fesh /fɛʃ/	H	230	76	97	1-feature	Bron
Book	Buck /bʌk/	RH	154	95	97	2-feature	Rad
Brush	Broosh /brʊʃ/	RH	253	73	90	2-feature	Rad
Cup	Cip/kɪp/	HB	254	79	98	2-feature	Bint
Foot	Fit /fɪt/	BR	207	70	97	2-feature	Bron
Keys	Kous /kʊs/	BR	267	74	96	2-feature	Bint
Cat	Cout /kʊt/	BRH	285	93	96	3-feature	Bif
Doll	Deal /di:l/	BRH	283	65	85	3-feature	Gek
Ball	Beal /bi:l//	BRH	266	96	100	3-feature	Rad
Hat	Hout /hʊt/	BRH	251	88	100	3-feature	Bif
Spoon	Span /spæn/	BRH	310	78	100	3-feature	Daz

error (six at 1;6 and four at 2;0). All children had no known hearing or visual problems and were recruited via the local maternity ward. Children came from homes where British English was the primary language in use.

Stimuli

The speech stimuli were produced by a female speaker of British English in an enthusiastic, child-directed manner. The audio-recordings were made with a solid state compact flash card recorder in a sound-treated recording booth. The audio stimuli were digitized at a sampling rate of 44·1 kHz and a resolution of 16 bits and spliced using Goldwave v. 5.10. The stimuli presented to children were fifteen monosyllabic (CVC) nouns taken from the British Communicative Developmental Inventory (Hamilton, Plunkett & Schafer, 2000), with which 50% of children at age 1;2 are reported to be familiar. According to the CDI data collected, the words were known to an average of 80% of children at 1;6 and 96% of children at 2;0. In addition, we created seven phonotactically legal novel words for the novel word condition. Based on W&M (2008), it was expected that children would look at the novel object when presented with the novel labels. Six mispronunciations resulted in non-words and nine mispronunciations resulted in real words with which children were unlikely to be familiar (see Table 1). Due to restrictions on the number of possible single feature changes resulting in legal English vowels, not all words could change in all of the features to yield all kinds of mispronunciations. Consequently, across children, five words yielded 1-feature mispronunciations, five words were

https://doi.org/10.1017/S0305000910000243 Published online by Cambridge University Press

changed to result in 2-feature mispronunciations, and five words resulted in 3-feature mispronunciations. We ensured that there was no systematic difference in the word durations of the correct and mispronounced labels ($t(14)=0{\cdot}52$; $p=0{\cdot}61$).

Visual stimuli were computer images created from photographs, with one image for each familiar word, and an image of a novel object paired with each familiar object across conditions. All subjects saw the same image pairs. Familiar images were judged by three adults (the authors and an independent observer) as typical exemplars of the labelled category. The novel images selected for the study were real objects, which children were not expected to have a name for (according to the BCDI; Hamilton *et al.*, 2000), e.g. an accordion, binoculars, old-fashioned perfume bottles.

Procedure

All children sat on their caregiver's lap during the experiment, facing a projection screen. Two cameras mounted directly above the visual stimuli recorded children's eye movements. Auditory stimuli were presented through a centrally located loudspeaker. Synchronized signals from the two cameras were then routed via a digital splitter to create a recording of two separate time-locked images of the infant.

Each child was presented with fifteen trials. In each trial, children saw an image of a familiar object and a novel object, side by side, for five seconds. Children were then presented with either correct pronunciations or mispronunciations of the familiar label, or novel words, inserted after the carrier phrase 'Look!' Onset of the target word began halfway into the trial at 2500 ms. The onset of the target word divided the trial into a prenaming and postnaming phase. Children saw each object only once during the experiment. Familiar and novel object pairings were maintained across pronunciation conditions. Children were presented with six correct pronunciations, two 1-feature mispronunciations, two 2-feature mispronunciations, two 3-feature mispronunciations and three novel label trials. Children never heard the same object labelled twice or heard the same word twice. Since each infant was presented with two of the three possible 1- and 2-feature mispronunciations, we ensured that the image pairs were counterbalanced so that each image pair appeared with a correct and an incorrect pronunciation equally often, and each mispronounced word appeared equally often across children. Familiar objects appeared equally often to the left and to the right. Likewise, correct and incorrectly pronounced words identified left and right targets equally often. Across children, image pairs appeared equally often with correct pronunciations, mispronunciations and novel words. Order of presentation of trials was randomized across children.

A digital-video scoring system was used to assess visual events on a frame-by-frame basis (every 40 ms). This technique enabled tracking of every single eye fixation. For analysis, we use the Proportional Target Looking measure (PTL), which is the amount of time children spent looking at the target (T) over the amount of time children spent looking at the target and distracter (T+D) in order to determine the proportion of time children spent looking at the target, i.e. T/[T+D]. The dependent variable used in both familiar and novel label trials is the difference in children's preference for the target image between the prenaming and postnaming phase (Postnaming $_{(PTL\ (T/[T+D]))}$ – Prenaming $_{(PTL\ (T/[T+D]))}$). We refer to this difference as the NAMING EFFECT. A positive value for this difference can be interpreted as a measure of the child's appreciation of the association between the heard label and the familiar image. A negative value would indicate the child's association of the heard label with the novel image. Only those trials in which children fixated both the target and the distracter during the prenaming phase were included in the analysis. We also calculated the Longest Look measure (LLK), which is the difference between children's single longest fixation at the target or familiar image (t) and distracter or novel image (d), i.e. t–d. Since both measures revealed a similar pattern of results, the results will be presented using the PTL measure, as in W&M (2008).

RESULTS

Children aged 2;0

Figure 1 suggests that two-year-olds showed a differentiation of 1-feature and 2- and 3-feature mispronunciations. To further examine these effects, we carried out a repeated measures ANOVA to see whether there was a significant difference between the three main pronunciation conditions (correct, mispronounced, novel words) and found a significant main effect of pronunciation type ($F(2,25)=3{\cdot}46$, $p=0{\cdot}04$). Planned comparisons found that there was a significant difference between children's performance following correct pronunciations and mispronunciations ($t(26)=2{\cdot}43$, $p=0{\cdot}02$) and a near-significant difference between correct pronunciations and novel word trials ($t(26)=1{\cdot}9$, $p=0{\cdot}06$), but not between mispronunciations and novel word trials ($t(26)=-0{\cdot}35$, $p=0{\cdot}72$). The effect of naming following correct pronunciations was significantly different from chance ($t(26)=3{\cdot}16$, $p=0{\cdot}004$), but not following mispronunciations ($t(26)=-0{\cdot}35$, $p=0{\cdot}7$) or novel words ($t(26)=0{\cdot}18$, $p=0{\cdot}8$). Note that Experiment 1 presented infants with more correct trials compared to incorrect trials in order to avoid infants getting frustrated with the experiment. This raised the concern that the effect of naming for correct pronunciations was driven by the greater number of trials in this condition.

https://doi.org/10.1017/S0305000910000243 Published online by Cambridge University Press

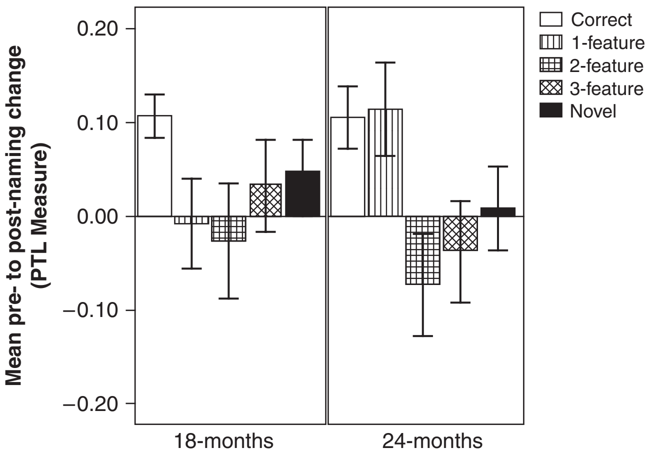

Fig. 1. Experiments 1 and 2: mean effect of naming for different pronunciation conditions at 1;6 and 2;0.

However, the effect of naming persisted even when only the first three correct pronunciation trials presented to children were considered ($t(26)=3{\cdot}36$, $p=0{\cdot}002$).

A repeated measures ANOVA with mispronunciation size (i.e. 1-, 2- and 3-feature mispronunciations) as a within-subjects factor revealed a significant main effect of mispronunciation size ($F(2, 25)=4{\cdot}57$, $p=0{\cdot}02$). Post-hoc comparisons revealed that the effect of naming following 1-feature mispronunciations was significantly different from chance ($t(26)=2{\cdot}31$, $p=0{\cdot}02$), but not following 2-feature ($t(26)=1{\cdot}35$, $p=0{\cdot}2$) or 3-feature mispronunciations ($t(26)=-0{\cdot}7$, $p=0{\cdot}4$). In addition, there was a significant difference in children's preference for the familiar image between 1- and 2-feature mispronunciations ($t(26)=3{\cdot}08$, $p=0{\cdot}005$) and between 1- and 3-feature mispronunciations ($t(26)=2{\cdot}36$, $p=0{\cdot}03$), but not between 2- and 3-feature mispronunciations ($t(26)=-0{\cdot}59$, $p=0{\cdot}5$).

The differences between 1-feature and 2-/3-feature mispronunciations suggest a marked distinction between smaller and larger mispronunciations, rather than a graded sensitivity to mispronunciations at age 2;0. This suggestion is borne out by the absence of a significant difference between 2- and 3-feature mispronunciations. One interpretation of this apparently non-linear difference between 1-feature and 2-/3-feature mispronunciations is that a quantification of mispronunciation size in terms of features may not provide a complete explanation of infants' behaviour. Therefore, we investigated whether the acoustic characteristics of the different mispronunciations were more crucial in driving infants' responses.

https://doi.org/10.1017/S0305000910000243 Published online by Cambridge University Press

We compared the acoustic characteristics of the three mispronunciation types, using the power spectrum of the midpoint of the vowel. The power spectrum of the vowel can be defined as the amount of vibration (in dB) at each individual frequency (in Hz), i.e. a plot of how power varies with frequency. We calculated the spectral energy at the midpoint of the steady state of the vowels of all the words presented to infants (correct and incorrect pronunciation). We then computed the difference between the spectra of the correct and incorrect pronunciations of the same word, using the formula

$$\text{Acoustic characteristics of a mispronunciation} = \sqrt{\sum_{i=1}^{n} (C_i - M_i)^2}$$

where n is the number of samples (bits) at which the spectral energy is recorded (256), C is the spectral energy recorded at the midpoint of the vowel of the correct pronunciation of a word for each sample, and M is the spectral energy recorded at the midpoint of the vowel in the vowel mispronunciation of the same word. This difference indexes the acoustic characteristics of each mispronunciation token. We then examined whether there was a correlation between the acoustic characteristics of each mispronunciation and infants' sensitivity to the mispronunciations. We use this Euclidean distance as our acoustic measure, since a single formant measure cannot provide information about the acoustic variance caused by different kinds of vocalic features, i.e. height, backness, roundedness. A power spectral measure, on the other hand, can characterize changes to vowel height, backness and roundedness as a single quantity.

Using unaggregated data, we found a significant correlation between the spectral quality of the mispronunciations (i.e. the spectral difference between correct and incorrect pronunciations) and the naming effect ($r = -0{\cdot}18$, $p = 0{\cdot}02$). This result implies that an increase in the acoustic deviation of the mispronunciation leads to an increase in the salience of the mispronunciation (decrease in effect of naming). However, we note that the acoustic characteristics of the mispronunciations correlate with the number of features involved in the mispronunciation ($r = 0{\cdot}55$, $p = 0{\cdot}03$). In contrast, the non-linear nature of the effect of increasing the number of features on infants' sensitivity to mispronunciations (i.e. the difference between 1-feature and 2-/3-feature mispronunciations) suggests that featural distance does not explain infants' behaviour. We analyzed this possibility in two ways.

First, using data aggregated by items, we carried out an analysis of covariance using the naming effect of mispronunciations as the dependent variable and the number of features (FEATURES) as the independent variable, covarying out the acoustic difference between the correct and incorrect pronunciations (ACOUSTIC DIFFERENCE). This led to a significant effect of acoustic difference ($F(1, 2) = 5{\cdot}35$, $p = 0{\cdot}04$), but not of features ($F(1, 2) = 0{\cdot}43$,

https://doi.org/10.1017/S0305000910000243 Published online by Cambridge University Press

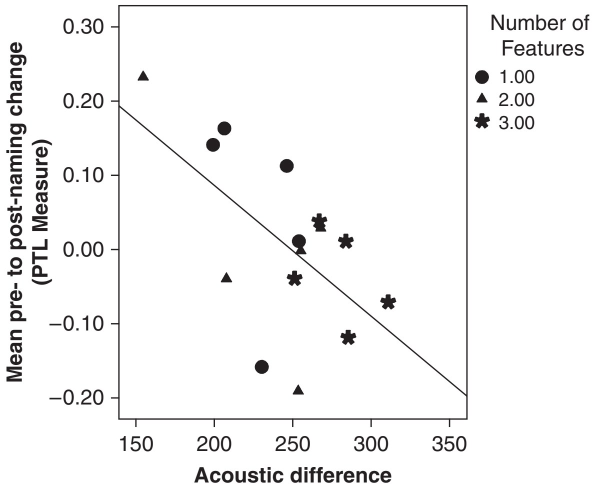

Fig. 2. Experiment 1: correlation between acoustic difference and effect of naming for different mispronunciation types at 2;0.

$p=0{\cdot}6$). Second, given the small number of degrees of freedom of the previous analysis, we also ran a stepwise multiple regression to investigate the individual contribution of ACOUSTIC DIFFERENCE and FEATURES. In this model, predictors are added or removed from the regression equation based on the predictive value of the two variables. The decision of the stepwise programme was to remove FEATURES from the regression equation due to the lack of predictive power of this variable, while retaining ACOUSTIC DIFFERENCE ($F(1, 13)=6{\cdot}949$, $p=0{\cdot}02$, $R^2=35\%$). According to the regression equation, a unit change in ACOUSTIC DIFFERENCE produces a change of 0·59 in the z score of the effect of naming. Only the ACOUSTIC DIFFERENCE between the correct and incorrect pronunciations was a worthwhile predictor of the effect of naming. See Figure 2 for a scatter plot of the acoustic characteristics of the mispronunciation against the size of the mispronunciation effect with the data aggregated by items.[2]

[2] Note that the current experiment describes the manipulations of the vowels in terms of vowel height, backness and roundedness alone, i.e. not using tenseness. Unfortunately, due to the constraints of the English vowel space and the limited lexical repertoire of young children, we could not systematically manipulate tenseness. We therefore ran a separate analysis including tenseness as one of the features in the analysis and found a similar pattern of results. As in the main analysis presented, a stepwise regression analysis removed FEATURES from the equation due to the lack of predictive power of this variable, while retaining ACOUSTIC DIFFERENCE. Once again, the acoustic difference between the correct and incorrect pronunciations was a better predictor of the effect of naming.

Children aged 1;6

As Figure 1 shows, there was no evidence of a graded sensitivity to vowel mispronunciations at this age. Children only demonstrated systematic looking at the familiar object when the label was correctly pronounced. There were no systematic preferences expressed in any of the other conditions. In order to examine these results further, we ran a repeated measures ANOVA with pronunciation type as a within-subjects factor (3: correct, mispronounced, novel word condition). The results indicated a significant main effect of pronunciation type ($F(2, 25)=4{\cdot}704, p=0{\cdot}018$). Post-hoc tests confirmed that there was a significant difference between children's performance following correct pronunciations and mispronunciations ($t(26)=3{\cdot}12$, $p=0{\cdot}004$), but not between correct pronunciations and novel words ($t(26)=1{\cdot}31$, $p=0{\cdot}2$) or between mispronunciations and novel words ($t(26)=-1{\cdot}27$, $p=0{\cdot}2$). In addition, there was a significant effect of naming following correct pronunciations ($t(26)=4{\cdot}647, p<0{\cdot}001$), but not following mispronunciations ($t(26)=-0{\cdot}17$, $p=0{\cdot}86$) or novel words ($t(26)=1{\cdot}44$, $p=0{\cdot}15$). As with the older children, we found that the effect of naming for correct pronunciations persisted even when the analysis only considered the first three correct pronunciations presented to children ($t(26)=2{\cdot}21$, $p=0{\cdot}026$).

We also compared children's performance in the three mispronunciation conditions with mispronunciation size as a within-subjects factor (3: 1-feature, 2-feature, 3-feature). The ANOVA revealed no significant main effect of mispronunciation size ($F(2, 25)=0{\cdot}26$, $p=0{\cdot}76$). There were no significant effects of naming following 1-feature ($t(26)=-0{\cdot}16$, $p=0{\cdot}87$), 2-feature ($t(26)=-0{\cdot}43$, $p=0{\cdot}67$) or 3-feature mispronunciations ($t(26)=0{\cdot}665$, $p=0{\cdot}5$). There were no differences in children's responding following 1- and 2-feature mispronunciations ($t(26)=0{\cdot}25$, $p=0{\cdot}7$), 2- and 3-feature mispronunciations ($t(26)=-0{\cdot}65$, $p=0{\cdot}5$) or 1- and 3-feature mispronunciations ($t(26)=-0{\cdot}68$, $p=0{\cdot}5$).

As expected, given the absence of difference between 1-, 2- and 3-feature mispronunciations, there was no correlation between the acoustic characteristics of the mispronunciation and the size of the mispronunciation effect ($p>0{\cdot}2$).

Vocabulary analysis

Note that in the current experiment, due to the constraints of the English vowel space, not all words were presented to children in all the conditions. Some words were presented as correct pronunciations and 1-feature mispronunciations alone, some as correct pronunciations and 2-features mispronunciations and so on. Given that children aged 1;6 knew an average of 80% of the words presented to them, we repeated the analyses using only those words that children were reported to be familiar with, using

https://doi.org/10.1017/S0305000910000243 Published online by Cambridge University Press

individual CDI data.[3] Three children were not included in this analysis due to the unavailability of CDI data for these children or their not providing enough trials per condition in this reduced dataset. We ran a repeated measures ANOVA with pronunciation type as a within-subjects factor (3: correct, mispronounced, novel word condition). The results again indicated a significant main effect of pronunciation type ($F(2, 22)=5{\cdot}71$, $p=0{\cdot}01$). Post-hoc tests confirmed that there was a significant difference between children's performance following correct pronunciations and mispronunciations ($t(23)=3{\cdot}45$, $p=0{\cdot}002$), a near-significant difference between correct pronunciations and novel words ($t(23)=1{\cdot}84$, $p=0{\cdot}08$) but no difference between mispronunciations and novel words ($t(23)=0{\cdot}71$, $p=0{\cdot}48$). In addition, there was a significant effect of naming following correct pronunciations ($t(23)=4{\cdot}68$; $p<0{\cdot}001$), but not following mispronunciations ($t(23)=-0{\cdot}48$, $p=0{\cdot}63$) or novel words ($t(23)=0{\cdot}53$, $p=0{\cdot}59$).

Comparing children's performance in the three mispronunciation conditions with mispronunciation size as a within-subjects factor (3: 1-feature, 2-feature, 3-feature) revealed no significant main effect of mispronunciation size ($F(2, 22)=0{\cdot}08$, $p=0{\cdot}92$). There were no significant effects of naming following 1-feature ($t(23)=0{\cdot}17$, $p=0{\cdot}9$), 2-feature ($t(23)=-0{\cdot}52$, $p=0{\cdot}6$) or 3-feature mispronunciations ($t(23)=-0{\cdot}09$, $p=0{\cdot}9$). There were no differences in children's responding following 1- and 2-feature mispronunciations ($t(23)=0{\cdot}41$, $p=0{\cdot}6$), 2- and 3-feature mispronunciations ($t(23)=-0{\cdot}29$, $p=0{\cdot}7$) or 1- and 3-feature mispronunciations ($t(23)=0{\cdot}79$, $p=0{\cdot}9$).

DISCUSSION

Experiment 1 investigated whether children aged 1;6 and 2;0 show a graded sensitivity to mispronunciations of the vowels in familiar words, i.e. whether children would be more sensitive to 3-feature mispronunciations, compared to 2-feature mispronunciations which in turn, may be more salient than 1-feature mispronunciations. The mispronunciations were counterbalanced so that 1-feature mispronunciations changed the height, roundedness or backness of the vowel. Two-feature mispronunciations changed the height and backness, the height and roundedness, or the backness and roundedness of the vowel. Three-feature mispronunciations necessarily changed the height, backness and roundedness of the vowel.

[3] Vocabulary reduced analysis is only presented for children aged 1;6 since the older children knew a much higher percentage of words presented to them (96%), leading to few changes to the dataset in the vocabulary reduced analysis.

Children aged 2;0

Children showed an effect of naming in correct pronunciation trials and 1-feature mispronunciation trials. However, they did not show an effect of naming for 2- or 3-feature mispronunciation trials. In addition, there was a significant difference in children's looking behaviour following 1-feature and 2- or 3-feature mispronunciations. However, the results suggested a marked difference between 1-feature and larger mispronunciations, with no difference between 2- and 3-feature mispronunciations. Therefore, we analyzed whether children's responding could be better explained by the acoustic characteristics of the mispronunciations. Indeed, children's sensitivity to mispronunciations correlated significantly with an increase in the acoustic difference between correct and incorrect pronunciations. However, the increase in the number of features contributing to the mispronunciation also correlated significantly with an increase in the acoustic differences between correct and incorrect pronunciations. Three-feature mispronunciations were acoustically more different from correct pronunciations compared to 2-feature mispronunciations. In turn, 1-feature mispronunciations were acoustically more similar to correct pronunciations than either 2- or 3-feature mispronunciations. Consequently, it is difficult to disentangle the contribution of acoustic and featural distance on children's responding. In an attempt to do so, we ran a multiple regression analysis, which confirmed that the acoustic difference between correct and incorrect pronunciations was a more accurate predictor of the variance in the effect of naming compared to the number of features contributing to the mispronunciation. Furthermore, an analysis of the influence of featural distance, covarying out acoustic difference, also found the acoustic difference to be a better predictor of children's responses. These findings suggest that acoustic characteristics are a better estimate of the impact of a vowel mispronunciation compared to the number of features – children's responding can be more appropriately described as a graded sensitivity to increasing acoustic differences between the correct and incorrect pronunciations.[4]

The results of Experiment 1 extend the previous work by W&M (2008), who report a graded sensitivity to consonant mispronunciations by age 1;7. With vowel mispronunciations, on the other hand, we find that the acoustic characteristics of the mispronunciations, rather than the featural characteristics, provide a good predictor of children's responding to vowel

[4] It is important to note that these results do not provide evidence against feature-based responding in two-year-olds. Such a claim is difficult to substantiate in the presence of the strong correlation between acoustic and featural information ($p = 0.03$) and the correlation between featural difference and mispronunciation sensitivity ($p = 0.05$, without covarying acoustic distance). Our results, therefore, suggest only that acoustic difference was a BETTER predictor of children's responding than featural distance.

mispronunciations.[5] By age 2;0, children are sensitive to differences between acoustically more or less salient mispronunciations.

Children aged 1;6

As with the two-year-olds, children aged 1;6 also discriminated between correct pronunciations and mispronunciations. However, they did not discriminate between small and large mispronunciations of the vowel, nor show a significant effect of naming for 1-, 2- or 3-feature mispronunciations, nor were there any differences in the effect of naming for the three kinds of mispronunciations. These results suggest that although children are sensitive to vowel mispronunciations at 1;6 (see also Mani & Plunkett, 2007; Mani *et al.*, 2008), they do not display a graded sensitivity to vowel mispronunciations at this age. One interpretation of these findings is that children at 1;6 are not sensitive to graded increases in the size of mispronunciations, suggesting that they do not possess a featural representation for vowels in familiar words. Neither do children at this age group differentiate between acoustically less or more salient mispronunciations. This finding contrasts with W&M's (2008) finding that children at 1;7 show graded sensitivity to the size of consonant mispronunciations and with the two-year-olds in Experiment 1, who demonstrated graded sensitivity to acoustic changes in the vowel mispronunciations. These results indicate that children aged 1;6 demonstrate a feature-based sensitivity to consonant mispronunciations but not vowel mispronunciations and that children aged 2;0 show an acoustically based sensitivity to vowel mispronunciations.

The results also provide a comparison of some interest in the performance of the children aged 1;6 and 2;0 following 1-feature mispronunciations. The younger children appear, in fact, to be more sensitive to 1-feature mispronunciations than the two-year-olds. The older children do not differentiate between correct pronunciations and 1-feature mispronunciations. W&M (2008) report a similar finding with children showing an effect of naming only for 1-feature mispronunciations, but not for 2- or 3-feature mispronunciations, i.e. looking to the familiar image following 1-feature mispronunciations, but not 2- or 3-feature mispronunciations. We suggest that for the two-year-olds who are able to pick up on the differing degrees of mispronunciations presented to them, the smaller difference between the

[5] Note that the notion of an important role for acoustic information is not inconsistent with White & Morgan's (2008) results. However, the nature of consonant mispronunciations makes it difficult to examine the influence of acoustic change on consonant mispronunciation sensitivity, since no single acoustic feature can adequately represent different kinds of consonant changes.

1-feature mispronunciation and the correct pronunciation may be obscured, relative to the more salient 2- and 3-feature mispronunciations and the novel label trials. For the younger children, the larger mispronunciations are not more salient than the smaller mispronunciations, and consequently may not obscure the difference between 1-feature mispronunciations and correct pronunciations. This is not to suggest that two-year-olds cannot detect 1-feature mispronunciations. Rather, it highlights the dynamic impact of the current task where two-year-olds may ignore a less salient mispronunciation when concurrently presented with more salient mispronunciations in other trials. In other studies where only small differences are used (Mani & Plunkett, 2007), two-year-olds can and do readily detect single-feature vowel mispronunciations.

Finally, we note that interpretation of the data from the children at 1;6 in Experiment 1 is complicated by the fact that these children showed little evidence for a difference in performance between novel label trials and correct pronunciations. One possibility is that the children aged 1;6 in the current study did not display graded sensitivity or mutual exclusivity due to the shorter duration of trials in the current study (5 s) compared to W&M (9 s).[6] This might suggest that the absence of a graded sensitivity to mispronunciations at 1;6 is due to children not being given an adequate opportunity to display fully their range of sensitivity to the stimuli presented to them. Indeed, the results of Mather & Plunkett (2009) and Halberda (2003) both suggest that children at this age need to be presented with the images for over 7 s to display mutual exclusivity. Experiment 2, therefore, examines the graded sensitivity of children aged 1;6 to vowel mispronunciations using longer trials than in Experiment 1, i.e. 8-s trials, in order to give children enough time to fully examine the relationships between the visual and auditory stimuli presented to them.

EXPERIMENT 2: CHILDREN AT 1;6

METHOD

Participants

The participants in this experiment were twenty-six children at 1;6 (M=1;5·27, Range=1;5·6 to 1;6·1). Five additional children were tested but were excluded due to fussiness, parental interference or experimenter error. All children had no known hearing or visual problems and were recruited via the local maternity ward. Children came from homes where British English was the primary language in use.

[6] Although White & Morgan (2008) presented infants with 14 s trials, only the first 9 s of each trial were incorporated into their reported analyses.

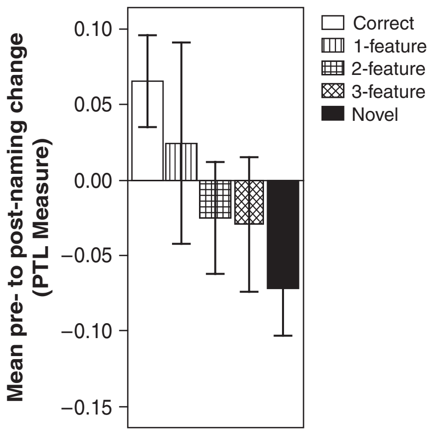

Fig. 3. Experiment 2: mean effect of naming for different pronunciation conditions at 1;6.

Procedure

The stimuli were identical to Experiment 1. The only difference between Experiments 1 and 2 is that in each trial in Experiment 2, children saw an image of a familiar object and a novel object, side by side, for EIGHT seconds. Children were then presented with either correct pronunciations or mispronunciations of the familiar label, or novel words, inserted after the carrier phrase 'Look!'. Onset of the target word began halfway into the trial at 4000 ms. The onset of the target word divided the trial into a prenaming and postnaming phase. All other factors were counterbalanced as in Experiment 1.

RESULTS

Children aged 1;6

As Figure 3 shows, while there was no evidence of a graded sensitivity to vowel mispronunciations at 1;6, children tended to look less at the familiar object when the label was mispronounced than when the label was correctly pronounced. Furthermore, they tended to look longer at the novel object when they heard a novel label. In order to examine these results further, we ran a repeated measures ANOVA with pronunciation type as a within-subjects factor (3: correct, mispronounced, novel word condition). The results indicated a significant main effect of pronunciation type ($F(2, 24)=3{\cdot}76$, $p=0{\cdot}038$). Planned comparisons confirmed that there was a significant difference between children's performance following correct pronunciations and mispronunciations ($t(25)=2{\cdot}08$, $p=0{\cdot}04$) and between correct pronunciations and novel words ($t(25)=2{\cdot}73$, $p=0{\cdot}01$), but not between mispronunciations and novel words ($t(25)=-1{\cdot}48$, $p=0{\cdot}15$). In addition, there was a significant effect of naming following correct pronunciations ($t(25)=2{\cdot}15$, $p=0{\cdot}04$) and novel words ($t(25)=-2{\cdot}3$, $p=0{\cdot}029$), but not following mispronunciations ($t(25)=-0{\cdot}32$, $p=0{\cdot}74$).

We also compared children's performance in the three mispronunciation conditions with mispronunciation size as a within-subjects factor (3: 1-feature, 2-feature, 3-feature). The ANOVA revealed no significant main effect of mispronunciation size ($F(2, 24)=0{\cdot}27$, $p=0{\cdot}76$). There were no significant effects of naming following 1-feature ($t(25)=0{\cdot}36$, $p=0{\cdot}72$), 2-feature ($t(25)=-0{\cdot}69$, $p=0{\cdot}49$) or 3-feature mispronunciations ($t(25)=-0{\cdot}66$, $p=0{\cdot}51$). There were no differences in children's responding following 1- and 2-feature mispronunciations ($t(25)=0{\cdot}63$, $p=0{\cdot}5$), 2- and 3-feature mispronunciations ($t(25)=0{\cdot}06$, $p=0{\cdot}94$) or 1- and 3-feature mispronunciations ($t(25)=0{\cdot}73$, $p=0{\cdot}46$).

As expected from the absence of differences between 1-, 2- and 3-feature mispronunciations, there was no correlation between the acoustic characteristics of the mispronunciation and the size of the mispronunciation effect ($p>0{\cdot}2$).

Vocabulary analysis

We repeated the analysis using only those words that individual CDI data indicated children were familiar with. This resulted in the exclusion of four children who did not provide data for one of the conditions tested using this reduced dataset. A repeated measures ANOVA with pronunciation type as a within-subjects factor (3: correct, mispronounced, novel word condition) yielded a near-significant main effect of pronunciation type ($F(2, 20)=3{\cdot}29$, $p=0{\cdot}058$). Planned comparisons confirmed that there was a near-significant difference between children's performance following correct pronunciations and mispronunciations ($t(21)=1{\cdot}91$, $p=0{\cdot}06$) and a significant difference between correct pronunciations and novel words ($t(21)=2{\cdot}38$, $p=0{\cdot}026$), but not between mispronunciations and novel words ($t(21)=-1{\cdot}14$, $p=0{\cdot}26$). In addition, there was a significant effect of naming following correct pronunciations ($t(21)=2{\cdot}60$, $p=0{\cdot}01$), but not following mispronunciations ($t(21)=0{\cdot}34$, $p=0{\cdot}73$) nor novel words ($t(21)=-1{\cdot}43$, $p=0{\cdot}16$).

Comparing children's performance in the three mispronunciation conditions with mispronunciation size as a within-subjects factor (3: 1-feature, 2-feature, 3-feature) revealed no significant main effect of mispronunciation size ($F(2, 20)=0{\cdot}69$, $p=0{\cdot}51$). There were no significant effects of naming following 1-feature ($t(21)=0{\cdot}81$, $p=0{\cdot}4$), 2-feature ($t(21)=0{\cdot}04$, $p=0{\cdot}9$) or 3-feature mispronunciations ($t(21)=-0{\cdot}5$, $p=0{\cdot}6$). There were no differences in children's responding following 1- and 2-feature mispronunciations ($t(21)=0{\cdot}63$, $p=0{\cdot}5$), 2- and 3-feature

https://doi.org/10.1017/S0305000910000243 Published online by Cambridge University Press

mispronunciations ($t(21)=0{\cdot}41$, $p=0{\cdot}6$) or 1- and 3-feature mispronunciations ($t(21)=1{\cdot}18$, $p=0{\cdot}2$).

DISCUSSION

The results of Experiment 2 suggested that even when presented with longer trials, children aged 1;6 do not differentiate between small and large mispronunciations of the vowels in familiar words. Despite displaying sensitivity to 1-, 2- and 3-feature mispronunciations of the vowels in familiar words, children did not display a graded sensitivity to the increasing number of features contributing to the mispronunciation. Neither did children display sensitivity to increases in the acoustic characteristics of the mispronunciations. The results of the current study suggest that, at 1;6, children appear to consider small and large vowel mispronunciations equivalently. This does not extend to complete label mismatches, however. Unlike Experiment 1, children in Experiment 2 differentiated between the novel label trials and correct pronunciation trials. Given the difference in children's responding to novel label trials in Experiments 1 and 2, as has previously been reported by Mather & Plunkett (2009), children at this age require longer exposure to the familiar image–novel image pairing (i.e. longer than the 5 s in Experiment 1) to display mutual exclusivity. Furthermore, the similarity between children's responding to the novel label trials in W&M (2008) and Experiment 2 provides a more consistent backdrop against which to interpret children's sensitivity to vowel and consonant mispronunciations at 1;6.

This difference in children's responding to vowel and consonant changes may be explained in terms of the differences in the nature of vocalic and consonantal features – vocalic features tend to be more fluid and distributed, with a change in one feature almost invariably leading to a change in another feature. For instance, the change from *bed* to *bud* is ostensibly a change in backness, but also includes a small change in the height of the vowel. In contrast, consonantal features tend to be more discrete or perceptually independent (Miller & Nicely, 1955: 348), such that a change in place of articulation need not necessarily involve a change in voicing or manner. Consequently, it may be easier for very young children to note a change in the size of consonant mispronunciations than changes to the size of vowel mispronunciations. This contrast may also be indicative of the more variable acoustic characteristics of vowels produced in natural speech, compared to consonants (Liberman, Delattre, Cooper & Gerstman, 1954; Pisoni, 1973) and the greater influence of acoustic characteristics on vowel perception: if the acoustic characteristics of vowels are more important than the feature-based representation, then the variability of the acoustic characteristics (relative to the feature-based representation) may have prevented children at 1;6 displaying a graded sensitivity to vowel mispronunciations, compared to consonants in W&M (2008).

However, it is also worth noting that the failure of children aged 1;6 in the current study may be task-specific. As highlighted by different models of child language development (notably, PRIMIR; Werker & Curtin, 2005), different tasks place very different cognitive demands on children, and poor performance in the current study may not be indicative of a poorly specified feature-based representation of vowels at 1;6, but of difficulties related to completing the task. Furthermore, the current study contrasts with previous work by Mani *et al.* (2008) showing that at 1;6 children are sensitive to the acoustic characteristics of mispronunciations. The introduction of the familiar image–novel image pairing in the current study may impact the ability of these children to differentiate between different kinds of mispronunciations.

It is difficult to draw strong conclusions regarding the older infants' graded sensitivity to vocalic FEATURES in the current study due to the strong correlation between acoustic and featural differences. This makes it difficult to directly compare infants' graded sensitivity to consonantal and vocalic features. It is possible that an experiment with longer trials may be able to disentangle the relative contribution of acoustic and featural characteristics to two-year-olds' responding to vocalic features, given the greater clarity in the data on children aged 1;6 using longer trials (Experiment 2). However, any comparison of the relative salience of acoustic and featural salience in vowels and consonants would still be plagued by the difficulty of obtaining an acoustic metric of consonantal differences. The W&M data, for instance, cannot differentiate between the acoustic and featural contributions to infants' graded sensitivity to consonantal changes as examined with vocalic features in the current article.

CONCLUSION

The current set of experiments attempted to investigate the underlying nature of children's lexical representations by examining whether children at 1;6 and 2;0 display a graded sensitivity to an increase in the number of features contributing to mispronunciations of vowels. We found that children show a marked distinction in their sensitivity to small and large mispronunciations at 2;0 but not at 1;6. This provides strong evidence for subsegmental representation of vowels by, at least, as early as two years of age. We have argued further that this subsegmental representation owes more to the acoustic than the featural characteristics of the mispronunciation. Note that we do not claim that this undermines the view that phonological features play an important role in characterizing children's lexical representations, but rather we highlight the quality or acoustic characteristics of the features distinguishing vowels more than the number of features. Furthermore, at least in the context of the current study, this attention to acoustic subsegmental detail does not appear to set in until at least age 2;0, inasmuch as younger children at 1;6 do not show a graded sensitivity to vowel mispronunciations, i.e. discriminate between smaller and larger mispronunciations of vowels. This provides an interesting contrast to W&M (2008), who find that children at this age can discriminate between smaller and larger mispronunciations of consonants, and raises questions about differences in the underlying representations of vowels and consonants. For example, does this contrast suggest that, early in development at least, consonants are represented more categorically, or with further detail than vowels?

REFERENCES

Bailey, T. & Plunkett, K. (2002). Phonological specificity in early words. *Cognitive Development* **17**, 1265–82.

Ballem, K. & Plunkett, K. (2005). Phonological specificity in 14-month-olds. *Journal of Child Language* **32**, 159–73.

Caramazza, A., Chialant, D., Capasso, R. & Micell, G. (2000). Separable processing of consonants and vowels. *Nature* **403**, 428–30.

Connine, C. M., Titone, D., Deelman, T. & Blasko, D. (1997). Similarity mapping in spoken word recognition. *Journal of Memory and Language* **37**, 463–80.

Curtin, S., Fennell, C. & Escudero, P. (2009). Weighting of acoustic cues explains patterns of word-object associative learning. *Developmental Science* **12**, 725–31.

Gerken, L., Murphy, W. D. & Aslin, R. N. (1995). 3-year olds' and 4-year-olds' perceptual confusions for spoken words. *Perception and Psychophysics* **57**, 475–86.

Halberda, J. (2003). The development of a word-learning strategy. *Cognition* **87**, B23–B34.

Hamilton, A., Plunkett, K. & Schafer, G. (2000). Infant vocabulary development assessed with a British communicative development inventory. *Journal of Child Language* **27**, 689–705.

Havy, M. & Nazzi, T. (2009). Better processing of consonantal over vocalic information in word learning at 16 months of age. *Infancy* **14**(4), 439–56.

Kuhl, P., Williams, K., Lacerda, F., Stevens, K. & Lindblom, B. (1992). Linguistic experience alters phonetic perception in infants by 6 months of age. *Science* **255**(5044), 606–608.

Liberman, A., Delattre, P., Cooper, F. & Gerstman, L. (1954). The role of consonant–vowel transitions in the perception of the stop and nasal consonants. *Psychological Monographs* **68**, 1–13.

Mani, N., Coleman, J. & Plunkett, K. (2008). Phonological specificity of vocalic features at 18-months. *Language and Speech* **51**, 3–21.

Mani, N. & Plunkett, K. (2007). Phonological specificity of vowels and consonants in early lexical representations. *Journal of Memory and Language* **57**(2), 252–72.

Mani, N. & Plunkett, K. (2008a). Vowels and consonants, difference or no difference. Paper presented to the International Conference on Infant Studies, Vancouver, Canada.

Mani, N. & Plunkett, K. (2008b). 14-month-olds pay attention to vowels in novel words. *Developmental Science* **11**(1), 53–59.

Mather, E. & Plunkett, K. (2009). Learning words over time: The role of stimulus repetition in mutual exclusivity. *Infancy* **14**(1), 60–76.

Miller, G. A. & Nicely, P. (1955). An analysis of perceptual confusions among some English consonants. *Journal of the Acoustical Society of America* **27**(2), 338–52.

Nazzi, T. (2005). Use of phonetic specificity during the acquisition of new words, differences between consonants and vowels. *Cognition* **98**(1), 13–30.

Nazzi, T., Floccia, C., Moquet, B. & Butler, J. (2009). Bias for consonantal over vocalic information in French- and English-learning 30-month-olds: Crosslinguistic evidence in early word learning. *Journal of Experimental Child Psychology* **102**, 522–37.

Nespor, M., Pena, M. & Mehler, J. (2003). On the different role of vowels and consonants in speech processing and language acquisition. *Lingue e Linguaggio* **2**, 203–229.

Pisoni, D. B. (1973). Auditory and phonetic memory codes in the discrimination of consonants and vowels. *Perception and Psychophysics* **13**(2), 253–60.

Stager, C. L. & Werker, J. F. (1997). Children listen for more phonetic detail in speech perception than word learning tasks. *Nature* **388**, 381–82.

Swingley, D. & Aslin, R. N. (2000). Spoken word recognition and lexical representation in very young children. *Cognition* **76**, 147–66.

Swingley, D. & Aslin, R. N. (2002). Lexical neighbourhoods and the word form representations of 14-month-olds. *Psychological Science* **13**(5), 480–84.

Swingley, D. & Aslin, R. N. (2007). Lexical competition in young children's word learning. *Cognitive Psychology* **54**, 99–132.

Werker, J. F. & Curtin, S. (2005). PRIMIR, A Developmental Framework of Infant Speech Processing. *Language Learning and Development* **1**(2), 197–234.

Werker, J. F., Fennell, C. T., Corcoran, K. M. & Stager, C. L. (2002). Infants' ability to learn phonetically similar words: Effects of age and vocabulary size. *Infancy* **3**(1), 1–30.

Werker, J. & Tees, R. (1984). Cross-language speech perception, Evidence for perceptual reorganisation during the first year of life. *Infant Behaviour and Development* **7**, 49–63.

White, K. S. & Morgan, J. (2008). Sub-segmental detail in early lexical representations. *Journal of Memory and Language* **59**, 114–32.