Developmental Changes in Cue Weighting: The Case of Voiceless Coronal Fricatives in Brazilian Portuguese

Fabiana A. Penido and Rui Rothe-Neves

Universidade Federal de Minas Gerais

The research presented here was made possible by a CNPq/PQ grant to the first author (309773/2012-1) and by financial support of the Fundo de Apoio Acadêmico, Faculdade de Letras, and Pró-Reitoria de Pesquisa, Universidade Federal de Minas Gerais.

Corresponding author: Rui Rothe-Neves, Laboratório de Fonética, Faculdade de Letras, Universidade Federal de Minas Gerais, Avenida Antônio Carlos, 6627 - 31270-217, Belo Horizonte/MG - Brazil, email: rothe-neves@ufmg.br

Abstract

Purpose: This study investigated the developmental changes in cue weighting for the distinction of voiceless coronal fricatives in Brazilian Portuguese. It examines the hypothesis that children weight vowel formant transitions relatively more than adults, whereas adults depend more on the spectral characteristics of the frication noise (Nittrouer, Manning & Meyer, 1993). Methods: The sample was composed of adults and 4 and 7-year-old children. Two continua were synthesised for the minimal pair [a'ʃa] 'to find' and [a'sa] 'to bake' and for the contrast [ʃu]-[su]. Two auditory cues were manipulated: *centre frequency of the frication noise* and *vowel formant transition*. Results: The results do not completely agree with the development pattern proposed by Nittrouer and colleagues. The cue weighting by both the 4-year-olds and 7-year-olds changed as a function of vowel and/or lexical status, and this was not observed among the adults. Furthermore, the adults weighted the F2 transition more heavily than the noise frequency. Conclusions: The results of this study suggest that younger children who are native speakers of Brazilian Portuguese do not always weight the F2 transition more heavily in the classification of voiceless coronal fricatives.

Keywords: cue weighting, speech perception, language development, Brazilian Portuguese, phonetics.

Introduction

In speech perception, acoustic characteristics related to phonetic features of speech sounds are used to distinguish sound categories of the listener's language. Following Wright (2001), we use the term "auditory cue" to mean information in the acoustic signal that allows the listener to apprehend the existence of a phonological contrast. Auditory cues are not equally important to every sound but are weighted according to the sound being perceived as well as its articulatory and communicative contexts, a phenomenon referred to as "cue weighting" (Holt & Lotto, 2006). In a series of experiments with English speakers, Nittrouer and colleagues have shown that children and adults differently weight the auditory cues to the same sound contrast [s]:[ʃ] followed by a vowel: children weight the vowel formant transitions from the consonant relatively more than do adults, whereas adults depended more on the spectral characteristics of the frication noise (Nittrouer, 1992, 1996; Nittrouer & Miller, 1997; Nittrouer & Studdert-Kennedy, 1987). Such a strategy may facilitate the analysis of the incoming signal and may indicate the articulatory gestures necessary to produce such a contrast in English. This pattern seems to change with increasing language experience because older children behave more similarly to adults.

In most studies by Nittrouer and colleagues, cue weighting was investigated by means of the contrast [s]:[ʃ] (Nittrouer, 1992, 1996; Nittrouer & Miller, 1997; Nittrouer, Miller, Crowther & Manhart, 2000). In this case, two auditory cues were manipulated: the *centre frequency of the frication noise* and *vowel formant transition*. In a "trading relations" paradigm, two sound continua are typically made by varying noise frequency, from a more [ʃ]-like (2.2 kHz) to a more [s]-like (3.8 kHz) frequency. The difference between the two continua lies in the formant transition to the vowel: a transition appropriate for the sound pair [sa] in one continuum and for [ʃa] in the other. This type of design allows for the investigation of the perceptual effects of these two cues. A listener not affected by the difference between the continua – that is, the formant transition cue – would show the same response to both continua.

As an explanation for that change pattern, Nittrouer, Manning and Meyer (1993; see also Nittrouer et al., 2000) proposed the Developmental Weighting Shift (DWS) model. According to the DWS model, changes of weighting in auditory cues along with increasing language experience may be related to an increase in sensitivity to phonetic structure. The rationale is as follows (see Lindblom, MacNeilage & Studdert-Kennedy, 1983). Because children's lexicons are initially limited to a few words, the distinctive "features" to be dealt with are more global. Speech processing at an early age thus proceeds by chunking larger units such as syllables and monosyllabic words. Thus, children tend to pay more attention than adults to cues that signal the limits of syllables, i.e., to vowel formant transitions. As the lexicon increases in size, increasingly more subtle differences must be dealt with to keep oppositions functional. At some point approximately 7 years of age, children process speech in a more analytical, adult manner, chunking shorter units and thus relying on auditory cues more robustly related to each single sound.

An alternative proposal (Sussman, 2001) explains weighting differences between children and adults as a function of general auditory processing. Because children have a less mature auditory system, they should be less able to process acoustic information that is insufficient, incomplete or physically not clearly distinguishable. Thus, children should be more confident in using cues that are higher, louder or spectrally more informative as formant transitions. Methodological differences, however, hinder the direct comparison of results. Mayo and Turk (2004, 2005) confronted both explanations using the same trading relations paradigm. In their first study, they investigated the weighting strategies of auditory cues by children aged 3-7 years and by adults, contrasting both fricatives and plosives. For each contrast, a transitional and a non-transitional cue were tested. Their results indicated that, contrary to what Nittrouer and colleagues proposed, children were not always biased towards transitional cues compared with adults. “The results of the current study suggest that even when methods are held constant, the observed types of adult–child cue weighting differences are likely to change depending on the segmental context being tested” (Mayo & Turk, 2004, p. 3192). Mayo and Turk (2005) focused on Sussman’s sensory explanation, and their study involved very careful acoustic manipulations of the spectral distinctiveness or informativeness of auditory cues in various contrasts. They found that some patterns of response adhered to the sensory explanation, in that children were indeed biased towards more salient cues. However, the children did not behave this way for all cues. The noticeable exception was observed in /t/:/d/ contrast, where the children were found to weight the spectrally distinctive transitions less heavily than the adults or to exhibit no differences. Mayo and Turk (2005) concluded that any explanation for cue weighting differences between children and adults cannot be solely based on sensory or strategy differences in these groups, as cue weighting seems to be strongly dependent on segmental context.

If context plays a role in developmental cue weighting differences, a natural question arises as to what extent this influence varies in different languages. It has long been recognised that languages may differ in how they implement a given feature value phonetically (Lindau & Ladefoged, 1986). Thus, an acoustic parameter that might be important for English speakers to differentiate a contrast may be less salient when the same contrast is produced by speakers of a different language. To the best of our knowledge, developmental cue weighting differences for native contrasts in any language other than

English have been reported only by Gerrits (2001). She investigated cue weighting for contrasting fricatives, plosives and vowel duration in Dutch adults and children (4, 6 and 9 years). In contrast to the authors mentioned thus far, Gerrits used minimal pairs of existing words, such as *sjok* 'plod' versus *sok* 'sock', *pop* 'doll' versus *kop* 'cup', and *zak* 'bag' versus *zaak* 'shop'. To decide to which fricative category a sound belongs, children at 6 and 9 years behaved like adults in using only the noise cue, a result we would expect if DWS holds. This was not the case when a stop was at stake: all groups depended on both cues, but only 4-year old children weighted the formant transition more heavily than the noise burst. Moreover, the adults showed a significant effect of the transitional cue. Lastly, all groups used both cues for vowel classification, whereby children at 4 and 6 years weighted the formant transition *less* heavily than older children or adults – a result we would *not* expect if DWS holds. Gerrits points to a clear difference in the use of formant transitions for the perception of fricatives compared with plosives and vowels, a fact more compatible with the claims of Mayo and Turk (2004, 2005). In the perception of fricatives, it is important to note that the English-speaking children investigated by Nittrouer still relied heavily on formant transitions at seven years of age, whereas Dutch-speaking children already showed a pattern similar to adults at 6 years. Because English and Dutch are both Germanic languages, this finding makes even more important the question of the role of language experience in the investigation of developmental changes in speech perception.

In the present study, we investigated developmental changes in cue weighting for the distinction of voiceless coronal fricatives in Brazilian Portuguese. In this language, [s] is acquired at approximately 2:6 years, earlier than [ʃ] at approximately 2:10 years (Oliveira, 2002; Sávio, 2001). Thus, it is expected that in normal acquisition, 4-year old children would be able to pronounce both fricatives correctly and to use them distinctively. Following the

paradigm extensively used by Nittrouer and colleagues and by Gerrits (2001), we used two continua of sounds manipulated in the centre frequency of frication noise along the continua, and the transition of the second formant (F2) across the continua, both adapted to values described for Brazilian Portuguese (Escudero, Boersma, Rauber & Bion, 2009). Contrary to all previously published studies, however, we used a *logit mixed-effects model* for the data analyses. Differences among groups of subjects can validly be examined by means of analysis of variance for any dependent measure as long as it does not violate its underlying principles (continuous and normal distribution and homogeneity of variance). However, as convincingly argued by Jaeger (2008), violations of the model's restrictions and the inflation of type I error have already been noted in the literature in respect to ANOVAs as applied to binomial data, even when data are transformed to fit within a logit model (see also, Baayen, Davidson & Bates, 2008; Quené & Van den Bergh, 2004, 2008). In previous studies on cue weighting, two different measures have been extracted from the same response curves and then used as dependent measures: the phoneme boundary, taken to be the 50% response point, and the slope. The phoneme boundary has typically been used to investigate whether the separation of phonic categories occurs earlier or later in each continuum, whereas the slope indicates whether responses are more or less categorical. Using a logit mixed-effects model, it is possible to conjointly model separation and slopes for the response curves as well as to quantify the importance of variance at the subjects' level. Morrison (2007; see also Escudero, Benders & Lipski, 2009) suggested that if the investigated factors are represented at the same scale, they can be taken as a direct measure of the importance of each factor; that is, it is a way of quantifying the weight attributed by each subject to each cue.

Methods

Participants

Three groups of ten participants each volunteered to participate in the study. The children groups (4 and 7 years of age) included 50% male participants. Adults aged 20-40 (mean: 29.5 year) included nine female participants. The examiners recruited the participants from schools in Belo Horizonte, MG, Brazil, and all participants were native Brazilian Portuguese speakers. The selection criteria were normal speech development, the absence of present or past hearing problems, and never having had tympanostomy tubes inserted. All of the participants were submitted to an auditory screening test in both ears. The test consisted of presenting pure tones at frequencies of 0.5, 1, 2, 4 and 8 kHz at 25 dB to 7-year-olds and adults (Davis & Silverman, 1970) and at 15 dB to 4-year-olds (Northern & Downs, 1989) using a portable Auditec/VSD2090 audiometer and Philips headphones (SHM1900). Participants were only included if they could hear the tones at the specified intensity. The 4-year-old children were additionally tested for phonological development (Yavas, Hernandorena & Lamprecht, 2002) to include only those for which the investigated fricatives [s]-[ʃ] were already part of their phonological inventory.

Stimuli

Two continua were synthesised for the minimal pair *achar* [a'ʃa] 'to find' and *assar* [a'sa] 'to bake', and two continua were synthesised for the nonword contrast [ʃu]-[su], both at a sampling rate of 16 kHz. Ten 160-ms-long fricative noise sounds were produced using the Klatt synthesiser implemented in Praat (Boersma & Weenink, 2011) with centre frequencies between 3360 and 6240 Hz appropriate for [ʃ] and [s], respectively (Santos, 1987), in steps of 320 Hz, and bandwidth of 230 Hz. Each sound was then faded-in after 5 ms and faded-out 5 ms from the end to avoid a sensation of burst. Vocalic portions of 270 ms duration were

synthesised in HLSyn (Sensimetrics Inc.) with the F2 transition manipulated for [a] and [u] (see Figure 1). The fundamental frequency declined between 100 and 80 Hz for [a] and between 120 and 100 Hz for [u]. The fricative noises and vocalic portions were then concatenated.

FIGURE 1 ABOUT HERE

Procedure

The same procedure as described by Gerrits (2001) was used. The participants responded in a classification task using the Brazilian Portuguese version of PercEval (André, Ghio, Cavé & Teston, 2003). In this approach, one stimulus at a time selected in random order from both continua is classified as either [a'ʃa] or [a'sa] in one session and as either [ʃu] or [su] in the other. Testing the children involved two pairs of figures that are associated with the sound categories, one pair for each session. In one pair, each figure depicted the actions described by the verbs that served as the best exemplars. The other figure pair depicted two robots named [ʃu] and [su]. Before testing, a training session was performed with the best exemplars. Stimuli were presented in two sessions, one for each contrast, with a self-terminated response time of up to 5 s. In each session, the stimuli were presented in three blocks of 50 trials each (10 frequency levels x 2 transitions x 2 repetitions for each test stimuli + 10 training stimuli) with a short pause between blocks. Children at 4 years had an additional training session during which they became familiar with figures and best exemplars. Each session varied from approx. 35 min (4 yr) to 25 min (7 yr) and 15 min (adults).

Data analysis

Of a possible total of 3600 data values for each vowel, 3587 were included for vowel [a] and 3585 for [u] (a total of 0.4% missing values due to lack of response in the appropriate time). The data were analysed using the *lmer* package in R (R Development Core Team, 2012; Bates, Maechler & Bolker, 2011). Because a direct comparison between vowels was not relevant, we analysed the results separately for [a] and [u]. The noise frequency was coded in ten levels from 1 to 10, and the F2 transition, in level 1 and 2 (appropriate for [ʃ] or [s], respectively). The noise frequency and the F2 transition levels were nested within Age (4, 7, adults) as independent variables. In the first series of models reported below (Model1, Table 2), age group differences were tested with reference to adults in the “Age Group” factor. In this way, for 4 and 7-year-olds each new factor was tested within the same age group.

The estimates for the F2 transition indicate a change in [sa] response log odds as a function of transition. It thus provides information on the separation of response curves, the same as the “phoneme boundary” in previous ANOVA-based studies. The estimates for noise frequency indicate how much the [sa] response log odds change from one step in frequency level to the next step, and thus, indicates the slope of the curve. A second series of models was then calculated without interaction with the age factor (Model 2, Table 3), in order to estimate how frequency noise and F2 transition would explain variance in each group separately. The importance or weight attributed to each cue (frication noise versus formant transition) was directly calculated by dividing the second model's estimate (β) for each cue by the sum of estimates, as suggested by Morrison (2007). In logistic regression models, β's represent how much the log odds of a participant's response change when one of the

predictors changes in one step. Morrison (2005, 2007) and Escudero, Benders and Lipski (2009) provide a rationale for the method, and, according to these authors, the comparison is straightforward: the weight of a cue is indicated by how much its β amounts relatively to all β's. In the present case, the main interest relay on the weight attributed to the frication noise frequency by the different age groups is calculated by:

$$\text{cue weighting} = \beta_{\text{frequency}} / (\beta_{\text{frequency}} + \beta_{\text{transition}})$$

The cue weighting would be 0.5 if frequency and transition were both attributed the same weight. While the z test on the estimates and the significance attributed to it refers to the difference caused by each factor and its importance in explaining variance in each model, the weighted β can be taken as a participant's reliance on each of the investigated cues (Morrison, 2007).

Results

The response scores (see Figure 2) show a clear age effect. The adults' classification curves are steeper than those of the children for [a] and [u], which means a more categorical response. There is also a clear difference between vowels. The per cent [su] responses for each group were not as different from each other as for [sa], indicating that even for adults there was not a perfect categorical effect, as shown by a sigmoid-shaped curve. This was more true for [su] than for [ʃu] in adults; in the other groups, the curves were more or less indistinguishable from transition to transition. Nevertheless, in all groups the [s]-transition caused more [su]-responses that crossed the 50% response boundary between categories earlier than did the [ʃ]-transition. This clear and consistent transition effect was more pronounced for adults than for children and for 7-year-olds than 4-year-olds.

FIGURE 2 ABOUT HERE

One of the advantages of mixed-effects models is that we can directly test the hypothesis that the frequency of frication noise is more important than the transition. Two strategies are typically used to estimate model fit. First, a null model is generated by taking only the intercept as independent variable plus the random effects at the subjects' level. In this way, the mean log odds are estimated across all contexts by taking into account only the variance at the subjects' level. The resulting model without linguistic factors is then compared with the full model by the likelihood ratio test (for further details, see Baayen, Davidson & Bates, 2008). We can safely conclude that the models are very well fitted for either [a] ($\chi^2 = 1902.3$, df = 11, $p < 2.2e-16$) or [u] ($\chi^2 = 872.2$, df = 8, $p < 2.2e-16$). The second strategy determines is the size of the proportion of variance accounted for by the model as estimated by R^2 from the correlation between the fitted values calculated by the model with the original data. In this case, the probability of [sa] responses was calculated for each group and condition using the model's estimates. The resulting values were then correlated with the observed [sa] responses (see Table 1). The observed responses by the adults and 4-yr-olds were more strongly correlated with the estimates than were those by the 7-year-olds. A vowel effect was also to see in the correlation coefficients, with responses in the [a] context more strongly correlated with the estimates than those in the [u] context.

TABLE 1 AND TABLE 2 ABOUT HERE

In Table 2, the first row of statistics is the reference level of the model's results, which means all other age groups were compared with the adults. The “Estimate” column presents the log odds of the [sa] responses. For adults, the log odds are -9.116 at the first level of frication noise frequency combined with a [ʃ]-transition. In the second column, the log odds are transformed into probabilities, a much more intuitive number. The adults had a very low probability of [sa] responses at that first level, which is indeed much more appropriate to [ʃa]: with $e^{-9.116} \approx 0.00011$, $p = 0.00011/(1 + 0.00011)$, i.e., 0.01%. More [sa] responses are expected when increasing the level of frication noise. The “noise frequency” row shows that the cumulative probability of [sa] responses contributed by the intercept plus each frequency step is 0.05%. The model tests how important the difference is in the noise frequency step from the cumulative probability, and the “z value” column reports the result of that test. The next column reports the p value associated with the z test. From the last column, we find that the contribution of the noise frequency is very important in explaining the variance. On the other hand, neither the formant transition nor the interaction between the noise frequency and formant transition adds any further explanation of variance in the data after the noise frequency was entered into the model. Simply stated, the formant transition was irrelevant to this model, and the increase in the adults' [sa] responses along both continua can be accounted for by change in the noise frequency level alone.

Compared with the adults, the children were more likely to answer [sa] to the first token, a difference that was greater for the 7-year-olds (0.082 to 0.001, $z = 7.674$, $p < 0.001$) than the 4-year-olds (0.043 to 0.001, $z = 6.804$, $p < 0.001$). For all age groups, there was an important effect of the noise frequency level, but the transition was only important in explaining the data variance of the 4-year-olds ($t = -2.749$, $p < .01$), and the frequency*transition interaction was even more important ($t = 3.88$, $p < 001$). In the 4-year-olds, the transition actually *decreased* the [sa] response probability from 6.4% to 2%. Another model was fitted without the frequency*transition interaction to test whether it was important for the fit of the model as a whole even when it was only important for one age group. The model with interaction was significantly better according to the likelihood ratio test ($\chi^2 = 18.314$, $df = 3$, $p < 0.0005$). In summary, the frequency level in all age groups was important to model the variation in the response behaviour from [ʃa] to [sa], while the formant transition was not important except for the 4-year-old children.

When frication noise was followed by [u], all three age groups showed a greater probability of responding to [su] at the first frequency level compared with the [a] vowel. In the model, the frequency noise level was even more important than it was for [a], except for the 7-year-olds, as seen from the z value column. Within contrast to the results for [a], the formant transition was important in modelling data of all three age groups, but no significant frequency*transition interaction was found. An age effect was found between the 7-year-old children and adults but not between the 4-year old children and adults (the difference between both groups of children was not directly tested by the model).

TABLE 3 AND TABLE 4 ABOUT HERE

As mentioned above, one way to compare the weights attributed to each cue directly is by means of the estimated coefficients (β's) in Model 2 (see Table 4 for comparison statistics for full versus null model). As calculated from the data in Table 3 (plotted in Figure 3 for convenience), in the context of vowel [a] both the 4-year-old and 7-year-old children attributed comparatively less weight to the noise frequency (.26 and .47, respectively) than to the formant transition. For [u], the adults weighted the formant transition more heavily (.58) than the noise frequency, as did the 7-year-olds (.65). Nonetheless, the 4-year-olds show an inverted pattern (.52 for noise frequency). Figure 3 shows a clear developmental trend in the vowel [a] context: as age increases, so does the importance of the noise frequency. Thus, it is observed that both the 4-year-olds and 7-year-olds children weighted more heavily the F2 transition. This pattern changes in the interval between 7-year-old children and adults where there was a reversal of the cue weighting strategies. For the adults, the F2 transition is less important than the noise frequency (.53). In contrast, this developmental trend was not observed in the vowel [u] context, where only for 4-year-old children the noise frequency seemed to be important.

FIGURE 3 ABOUT HERE

Discussion

For the classification of the [ʃ]-[s] contrast, the response curves were found to be steeper the older the participant (Nittrouer & Studdert-Kennedy, 1987; Nittrouer & Miller, 1997; Hanzan & Barret, 2000; Gerrits, 2001), a strong suggestion that the centre frequency of frication noise is more heavily weighted compared with other available cues when deciding between those sounds. Our results agree with DWS only in the context of vowel [a], for in the context of vowel [u] the younger children did not weight more heavily the dynamic cue, nor did the adults weight more heavily the static cue. In classifying the continuum within the context of the minimal pair [a'ʃa]:[a'sa], the 4-year-olds appear to have been more influenced by the F2 transition, as previously described (Nittrouer & Studdert-Kennedy, 1987; Nittrouer & Miller, 1997; Gerrits, 2001; Mayo & Turk, 2004). The results of the 7-year-olds also agree with the previous studies in English (Nittrouer & Miller, 1997; Nittrouer, 2002), as they weighted the F2 transition more heavily than the adults in the [a'ʃa]:[a'sa] contrast. The trends found by Gerrits (2001) in Dutch were in the opposite direction: 6-year-olds distinguished between "sjok-sok" [ʃok]-[sok] based on the centre frequency of frication noise. In our results, the 7-year-olds were significantly influenced by the F2 transition, but they also showed a slight influence of the frication noise in classifying a [ʃ-s] continuum. The adults weighted more heavily the frication noise than the F2 transition, as expected from the previous literature (Nittrouer & Studdert-Kennedy, 1987; Nittrouer & Miller, 1997; Gerrits, 2001). This finding is also in line with DWS.

In summary, in the minimal pair 4-year-old children were indeed more influenced by the dynamic cue of the F2 transition compared with the other age groups, as in English and Dutch. As expected, the adults here showed a reverse pattern of results. A remarkable difference between adults and 4-year-olds is that both cues significantly interacted for the younger children, while there was no interaction for the adults. It seems that at this age the children do not use cues independently. For the [ʃu-su] continuum, our results differ from those obtained for the minimal pair [a'ʃa]:[a'sa]. Noise frequency was a significant cue only for the 4-year old children. From previous studies, this pattern of results was expected for the adults (Nittrouer & Studdert-Kennedy, 1987; Nittrouer & Miller, 1997). This group also used the F2 transition, but contrary to what happened with [a'ʃa]:[a'sa], there was no interaction between cues in this condition. Noise frequency is apparently not as important a cue for to contrasting [ʃ-s] in the context of [u] in Brazilian Portuguese as it is for English listeners. Interestingly, the 7-year-old children were the least categorical of all groups in both vowel contexts, and a developmental trend would predict such a pattern from the 4-year-olds. The 4-year-olds seem to have been confounded by the difference in formant transition, as it actually decreased the probability of responding in the same direction cued by the transition. Perhaps

more crucially from a developmental perspective is that both cues interacted in the 4-year-old group and not in the other groups, indicating that younger children may not use these cues independently of each other, as shown in Table 2.

In general, most studies on cue weighting show that neither young children nor adults change their weighting patterns for the [ʃ-s] contrast because frication noise is associated either with [a], [i], [o], [u] or [ai]. Young children weight more heavily the formant transition, while adults prefer the frication noise (Nittrouer & Studdert-Kennedy, 1987; Nittrouer & Miller, 1997; Gerrits, 2001, Mayo & Turk, 2004). For stop-consonant contrasts, however, the results provide a less clear picture and challenge the dynamic-to-static explanation of DWS (see Mayo & Turk, 2004, 2005; Gerrits, 2001). Two observations seem to fit here. First, in DWS, the preference of a static- versus a dynamic cue is related to differences in processing strategies that change along with language development, not with the phonetic details of the available cues in different vowel contexts. Thus, we would expect that different uses of specific cues by different age groups would not be affected by vowel contexts. This is not what we observed when comparing the β coefficients by vowel and age (see Figure 3). The adults weighted the F2 transition more heavily than noise frequency in the [u] vowel context. The 4-year-old children weighted the F2 transition more heavily in the [a] context – where both cues interacted significantly – but in the [u] context they weighted the noise frequency more heavily, with no interaction between cues. Acoustically, the F2 is less intense due to labialisation in [u] and it is therefore less available than in [a]. Thus, what emerges here is not a dynamic-to-static DWS-like picture but a pattern where the 4-year-olds are less stable and use whatever cue is more available in the context at hand, alone or in interaction. The adults show a more consistent independent use of both cues in different but stable levels. For the stop-consonant contrasts already tested, previous studies showed that the F2 transition was also the primary cue for adults while contrasting place of articulation.

The second observation relates to the minimal pair [a'ʃa]:[a'sa] categorisation being quite different from the non-lexicalised contrast [ʃu-su]. We recall from Figure 1 that the adults' responses were clearly categorical in the minimal pair, but not in the non-lexicalised contrast. This result reproduces the pattern described in the second experiment by Nittrouer & Miller (1997, p.2259) with a different relationship between vowel and lexical status: while "sue-shoe" is a minimal pair in English, "sa-sha" is not. Thus, it is again more likely that the categorical nature of the responses comes from the lexical status rather than the vowel context, even though a direct test would only be possible within the same language. Such a test, to the best of our knowledge, has not been pursued. Mayo and Turk (2004, 2005) used only non-words, Gerrits (2001) used only words, and most studies have used words or non-words as we did, and thus, they cannot be compared directly (Nittrouer, 1992; 1996; Nittrouer & Miller, 1997; Nittrouer & Studdert-Kennedy, 1987). Categorical perception was shown to be affected by linguistic knowledge as an effect of a labelling strategy that changes bias towards one response in tasks that do not require the exclusive use of auditory cues (Gerrits & Schouten, 2004; Kingston, 2005). Because phonological knowledge from classification (or “identification”) and the lexicon has proven to be affected by phonological knowledge, it is not a surprise to find categorical responses affected by lexical status.

Conclusions

In summary, the results reported here do not entirely agree with the developmental pattern of cue weighting predicted by DWS (Nittrouer, Manning & Meyer, 1993). The cue weighting by the 4-year-olds and the adults changed as a function of vowel and/or lexical status, which was not observed among the 7-years-olds. For the minimal pair [a'ʃa]:[a'sa], the observed results agree to DWS predictions. They were not, however, in line with the classification results for [ʃu-su]. Mayo and Turk (2004, 2005) consistently showed that neither the predictions by DWS nor those by Sussman (2001) are sustained after manipulating many different segmental and vocalic contexts. However, no other systematic explanation was offered. Because we compared cue weighting in children and adults by means of β coefficients estimated from a mixed model, a picture emerged in which 4-year old children are less consistent than adults in their relative use of available cues, despite the lexical status of the sequence in which the contrasted segments were inserted. It remains an open question as to whether the use of the F2 transition and noise frequency and the differences from children to adults reported here are only to be found in Brazilian Portuguese.

References

André C., Ghio, A., Cavé, C., & Teston, B. (2003). PERCEVAL: a Computer-Driven System for Experimentation on Auditory and Visual Perception. Proceedings of XVth ICPhS (pp. 1421-24). Barcelona, Spain.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of memory and language*, *59*, 390-412.

Bates, D., Maechler, M., & Bolker, B. (2011). lme4: Linear mixed-effects models using S4 classes. Retrieved from: < http://cran.R-project.org/package=lme4>. R package version 0.999375-42.

Boersma, P., & Weenink, D. (2011). *Praat*: *doing phonetics by computer*. Version 5.2.25. Retrieved from: <http://www.praat.org>.

Davis, H., & Silverman, S. R. (*1970). Auditory Test Hearing Aids*. In Davis, H. & Silverman, S (Eds.), *Hearing and Deafness.* Holt: Rinehart and Winston.

Escudero, P., Benders, T., & Lipski, S. C. (2009). Native, non-native and L2 perceptual cue weighting for Dutch vowels: The case of Dutch, German, and Spanish listeners. *Journal of Phonetics*, *37*, 452-65.

Escudero, P., Boersma, P., Rauber, A. S., & Bion, R. A. (2009). A cross-dialect acoustic description of vowels: Brazilian and European Portuguese. *Journal of the Acoustical Society of America*, 126, 1379-93.

Gerrits, E. (2001). *The categorization of speech sounds by adults and children. A study of the categorical perception hypothesis and the developmental weighting of acoustic speech cues* (PhD thesis). Utrecht University, The Netherlands, LOT series.

Hanzan, V., & Barret, S. (2000). The development of phonemic categorization in children aged 6-12. *Journal of Phonetics*, 28, 377-96.

Holt, L. L., & Lotto, A. J. (2006). Cue weighting in auditory categorization: Implications for first and second language acquisition. *Journal of the Acoustical Society of America*, 119, 3059-71.

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of memory and language*, *59*, 434-46.

Lindau, M., & Ladefoged, P. (1986). Variability of feature specifications. In J. Perkell & D. Klatt (Eds.), *Invariance and variability in speech processes*. Hillsdale, NJ: Lawrence Erlbaum, 464-78.

Lindblom, B., MacNeilage, P., & Studdert-Kennedy, M. (1983). Self-organizing processes and the explanation of phonological universals. *Linguistics*, 21, 181-204.

Mayo, C., & Turk, A. (2004). Adult–child differences in acoustic cue weighting are influenced by segmental context: Children are not always perceptually biased toward transitions. *Journal of the Acoustical Society of America*, 115, 3184-94.

Mayo, C., & Turk, A. (2005). The influence of spectral distinctiveness on acoustic cue weighting in children's and adult's speech perception. *Journal of the Acoustical Society of America*, 118, 1730-41.

Morrison, G. S. (2005). An appropriate metric for cue weighting in L2 speech perception. *Studies in Second Language Acquisition*, 27, 597-606.

Morrison, G. S. (2007). Logistic regression modelling for first- and second- language perception data. In M. Solé, P. Prieto, & J. Mascaró (Eds.), *Segmental and prosodic issues in Romance phonology*. Amsterdam: John Benjamins, 219–36.

Nittrouer, S. (1992). Age-related differences in perceptual effects of formant transitions within syllables and across syllable boundaries. *Journal of Phonetics*, 20, 351-82.

Nittrouer, S. (1996). Discriminability and perceptual weighting of some acoustic cues to speech perception by three-year-olds. *Journal of Speech, Language and Hearing Research*, 39, 278-97.

Nittrouer, S., Miller, M.E., Crowther, C.S., & Manhart, M.J. (2000). The effect of segmental order on fricative labelling by children and adults. *Perception & Psychophysics*, 62, 266-84.

Nittrouer, S., Manning, C., & Meyer, G. (1993). The perceptual weighting of acoustic cues changes with linguistic experience. *Journal of the Acoustical Society of America*, 94, 1865.

Nittrouer, S., & Studdert-Kennedy, M. (1987). The role of coarticulatory effects on the perception of fricatives by children and adults. *Journal of Speech, Language and Hearing Research*, 30, pp. 319-29.

Nittrouer, S., & Miller, M. E. (1997). Predicting developmental shifts in perceptual weighting schemes. *Journal of the Acoustical Society of America*, 101, 2253-66.

Northern, J. L., & Downs, M. P. (1989). *Audição em crianças.* São Paulo: Manole.

Oliveira, C.C. (2002). *Aquisição das fricativas /f/, /v/, /š/ e /ž/ do Português Brasileiro*. (Master's Thesis). Faculdade de Letras, PUCRS, Porto Alegre.

Quené, H., & Van den Bergh, H. (2004). On multi-level modeling of data from repeated measures designs: A tutorial. *Speech Communication*, *43*, 103-21.

Quené, H., & Van Den Bergh, H. (2008). Examples of mixed-effects modeling with crossed random effects and with binomial data. *Journal of Memory and Language*, *59*, 413-25.

R Development Core Team. (2012). R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing.

Santos, M. T. M. (1987). *Uma análise espectrográfica dos sons fricativos surdos e sonoros do Português Brasileiro.* Escola Paulista de Medicina, São Paulo. (manuscript).

Sávio, C. B. (2001). *Aquisição das fricativas /s/ e /z/ do Português Brasileiro.* (Master's Thesis). Faculdade de Letras, PUCRS, Porto Alegre.

Sussman, J. E. (2001). Vowel perception by adults and children with normal language and specific language impairment: Based on steady states or transitions? *Journal of the Acoustical Society of America*, 109, 1173-80.

Wright, R. (2001). Perceptual cues in contrast maintenance. In E. Hume & K. Johnson (Eds.), *The role of speech perception in phonology* (pp. 251-78). San Diego: Academic Press.

Yavas, M., Hernandorena, C. L. M., & Lamprecht, R. R. (2002). *Avaliação Fonológica da Criança: Reeducação e Terapia*. Porto Alegre: Artmed.

Table 1

R^2 between fitted and observed probability values

Contrast	Group					
	Adult		7 year		4 year	
	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2
[(ʃ)a]	.996	.996	.948	.949	.978	.978
[(s)a]	.98	.988	.833	.864	.985	.986
[(ʃ)u]	.96	.96	.749	.749	.927	.927
[(s)u]	.948	.948	.756	.756	.957	.957

Table 2

Model 1- Fixed Effects for Group Response Probabilities by Noise Frequency and Transition

Group [a]	Factor	Estimate	Calculated Prob.	Std. Error	z value	Pr(>|z|)
Adults	(Intercept)	-9.116	0.0001	0.818	-11.147	< 2e-16 ***
	Noise Frequency	1.528	0.0005	0.13	11.730	< 2e-16 ***
	Transition	1.762	0.0029	0.987	1.785	0.074
	Frequency * Transition	-0.081	0.0027	0.172	-0.471	0.638
4 year	Age Group	6.016	0.0431	0.884	6.804	1.02e-11 ***
	Noise Frequency	0.411	0.0636	0.04	10.374	< 2e-16 ***
	Transition	-1.171	0.0206	0.426	-2.749	0.006 **

	Frequency * Transition	0.248	0.0262	0.064	3.880	0.0001 ***
7 year	Age Group	6.7	0.082	0.873	7.674	1.67e-14 ***
	Noise Frequency	0.397	0.1175	0.037	10.729	< 2e-16 ***
	Transition	0.03	0.1206	0.327	0.091	0.927
	Frequency * Transition	0.083	0.1298	0.054	1.533	0.125
[u]						
Adults	(Intercept)	-2.781	0.058	0.2205	-12.61	< 2e-16 ***
	Noise Frequency	0.519	0.094	0.0304	17.04	< 2e-16 ***
	Transition	0.712	0.175	0.1457	4.88	1.0e-06 ***
4 year	Age Group	0.311	0.078	0.3030	1.03	0.305
	Noise Frequency	0.412	0.113	0.0264	15.62	< 2e-16 ***
	Transition	0.383	0.158	0.1343	2.85	0.004 **
7 year	Age Group	1.727	0.258	0.2867	6.02	1.7e-09 ***
	Noise Frequency	0.208	0.3	0.0223	9.30	< 2e-16 ***
	Transition	0.392	0.388	0.1242	3.16	0.002 **

Note. Significance of the codes: ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05.

Table 3

Model 2 - Fixed Effects for Group Response Probabilities by Noise Frequency and Transition

Group	Factor	Estimate	Calculated Prob.	Std. Error	z value	Pr(>|z|)
[a]						
Adults	(Intercept)	-9.12	0.0001	0.659	-13.834	< 2e-16 ***
	Noise Frequency	1.529	0.0005	0.095	16.052	< 2e-16 ***
	Transition	1.35	0.0029	0.24	5.602	2.12e-08***
4 year	(Intercept)	-3.08	0.0439	0.317	-9.73	< 2e-16 ***
	Noise Frequency	0.409	0.0647	0.039	10.359	< 2e-16 ***
	Transition	-1.16	0.0212	0.425	-2.742	0.006116**

	Frequency * Transition	0.246	0.027	0.06	3.87	0.000109***
7 year	(Intercept)	-2.62	0.0679	0.23253	-11.281	< 2e-16 ***
	Noise Frequency	0.433	0.101	0.02731	15.877	< 2e-16 ***
	Transition	0.48	0.154	0.13728	3.509	0.00045***
[u]						
Adults	(Intercept)	-2.774	0.0587	0.2147	-12.91	< 2e-16 ***
	Noise Frequency	0.517	0.0948	0.0303	17.04	< 2e-16 ***
	Transition	0.709	0.176	0.1455	4.88	1.07e-06 ***
4 year	(Intercept)	-2.449	0.0795	0.1777	-13.78	< 2e-16 ***
	Noise Frequency	0.408	0.115	0.0262	15.56	< 2e-16 ***
	Transition	0.38	0.16	0.134	2.84	0.00457 **
7 year	(Intercept)	-1.069	0.256	0.2239	-4.775	1.8e-06 ***
	Noise Frequency	0.2109	0.298	0.0225	9.36	< 2e-16 ***
	Transition	0.397	0.387	0.1251	3.177	0.00149 **

Note. Significance of the codes: ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05.

Table 4

Comparison statistics for full versus *null models by age**

Vowel	Group	χ^2	Df
[a]	Adults	1124.6	2
	4-yrs.	424.05	3
	7-yrs.	342.4	2
[u]	Adults	441.14	2
	4-yrs.	320.63	2

	7-yrs.	103.28	2

* Models are highly significant at p< 2.2e-16.

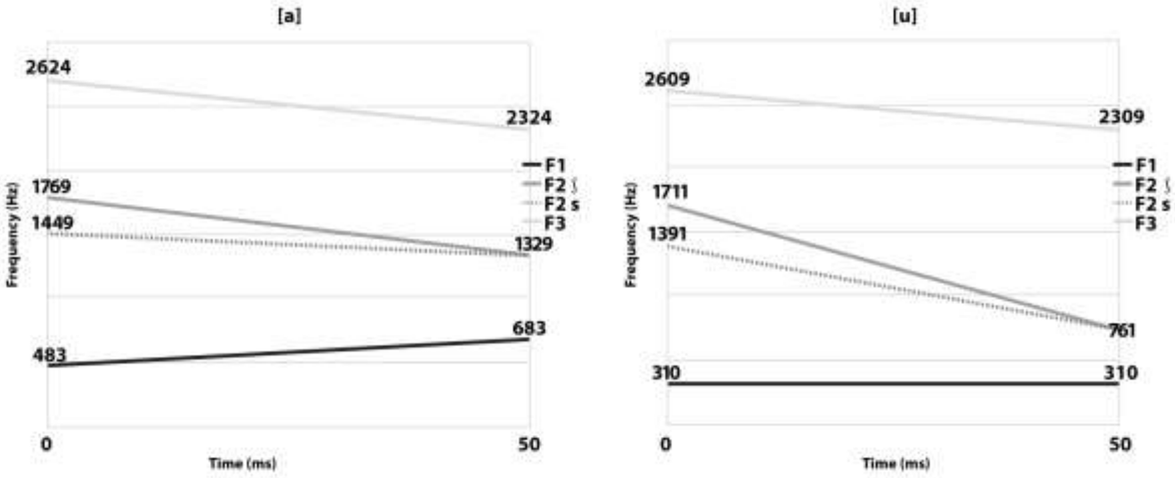

Figure 1. Initial and final values of the vowel formant frequencies (Hz).

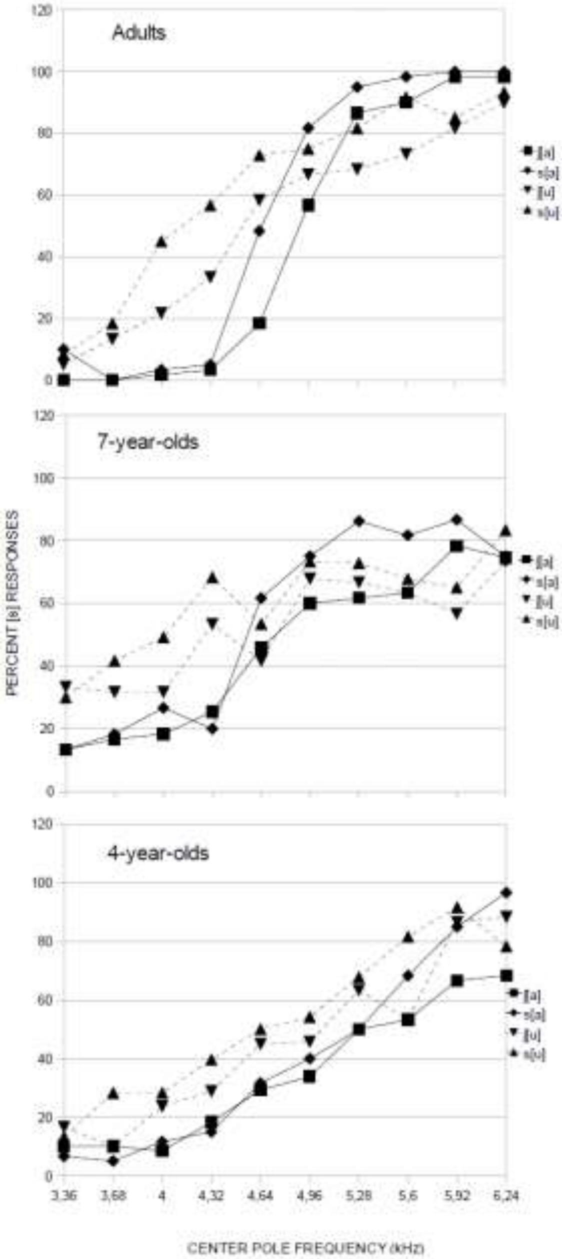

Figure 2. Percent “s” responses as a function of noise frequency compared by vowel context, formant transition and age group.

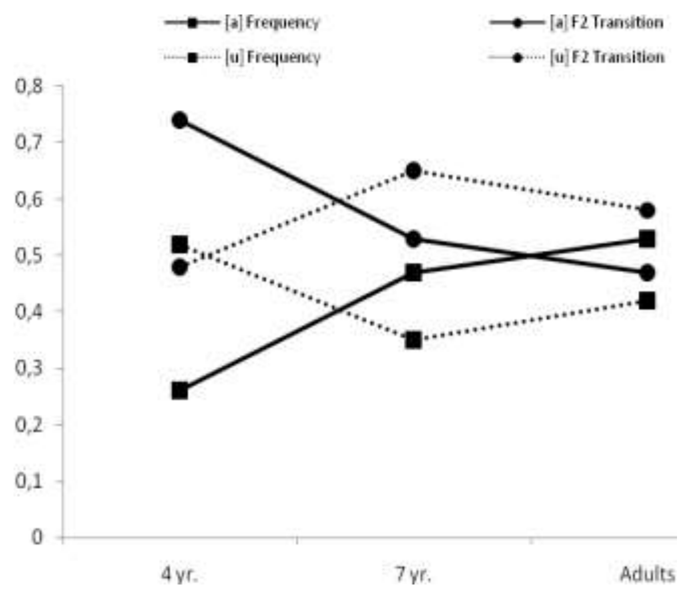

Figure 3. Cue weighting as a function of age (β calculated from Table 3).