PLA – WP

PURDUE LINGUISTICS ASSOCIATION WORKING PAPERS

VOLUME 1, ISSUE 1 SPRING 2008

CUE WEIGHTING AT DIFFERENT AGES: THE CASE OF /S-ʃ/

ALEJANDRINA CRISTIÀ

Purdue University

Over the last twenty years, a considerable number of studies have assessed the perception of /s-ʃ/ by participants of different ages. This work has shown differences in how children and adults tend to distribute their attention among a set of acoustic cues, although there is disagreement on the biases that contribute to the dissimilar cue weighting schemes. In this paper, I compare 9 such experiments presenting a summary of the design, participants, stimuli construction, and conclusions. I conclude that, although the finding that the same stimuli were labeled differently by these two populations is robust, we cannot conclude that this is due to a difference in cue weighting due to perceptual and phonological reasons only. On the contrary, stimuli construction and procedure have generally conspired to make the task much harder for children than for adults. Confirmation that children's and adults' cue weighting is dissimilar for truly cognitive reasons may need to await corpora studies showing that the distribution of acoustic cues in the input to children is the same as that in adults, and labeling studies were stimuli are more similar to natural tokens.

1. INTRODUCTION**.** In order to learn a spoken language one must learn the language’s phonology, its sound system. Given that languages vary in the way that they encode sound contrasts, in order to learn the phonology, the first language learner must discover which acoustic variation is most informative in their language. For example, /p/ and /b/ are produced with a complete closure on the lips in both Italian and English; however, the duration of this closure between vowels may used by English listeners to determine whether they heard /p/ or /b/ (Lisker, 1957), while that duration is not an informative cue for Italian listeners, who rely on other acoustic cues (Esposito, 2002). Unfortunately, there is little research devoted to how infants distribute their attention among multiple acoustic cues to such sound contrasts. The present paper represents a review of the literature assessing developmental differences in perception of /s/ and /ʃ/.

In a seminal paper, Lisker (1986) lists 16 acoustic correlates to voicing of stops in intervocalic position; however, subsequent experimentation suggests that listeners do not rely on all correlates equally, but allocate more or less attention to one or other correlate instead listeners. Further, these attentional schemes or cue weighting strategies have been suggested to vary depending on age. One contrast that has been targeted extensively is that between the coronal and palatal fricatives /s/ and /ʃ/ in prevocalic position, first addressed in Nittrouer and Studdert-Kennedy (1987). Since then, the contrast has been revisited mainly by Nittrouer and colleagues (Nittrouer, 1992; Nittrouer & Miller, 1997b, 1997a; Nittrouer, Miller, Crowther, &

← GO BACK TO TABLE OF CONTENTS

Manhart, 2000; Nittrouer, 2002) and recently by other researchers (Mayo, 2000; Mayo & Turk, 2004).

In this report, I focus on 9 experiments reported in 6 papers (Nittrouer & Studdert-Kennedy, 1987; Nittrouer, 1992; Nittrouer & Miller, 1997b, 1997a; Hazan & Barrett, 2000; Mayo & Turk, 2004) and 1 dissertation (Mayo, 2000, Experiment). Although this does not constitute a comprehensive review, it is representative enough to draw some general conclusions on what aspects ought to be taken into account for future studies. The rest of this paper summarizes some key aspects of the studies reviewed, while the final discussion addresses some possible general weaknesses affecting this line of research.

2. DESIGN. The goal of most of these studies was to assess differences in how adults and children weight two acoustic cues to place of articulation in the coronal fricative /s/ and the palatal fricative /ʃ/, namely frication noise and transitional cues (Nittrouer & Studdert-Kennedy, 1987; Nittrouer, 1992; Nittrouer & Miller, 1997b, 1997a; Mayo, 2000; Mayo & Turk, 2004). In order to address this question, two manipulations were done. On one hand, the frication noise was varied in steps between an /s/-like sound and an /ʃ/-like sound in a continuum. On the other, transitional information was taken from naturally produced or synthesized vowels and could have only two values, one appropriate for /s/, the other appropriate for /ʃ/. The precise number of resulting stimuli varied across studies, but most often it was 18 per contrast (9 different frication noises combined with /s/-appropriate transitions and with /ʃ/- appropriate transitions). The design of Hazan and Barrett (2000) was slightly different, since the goal there was to investigate the development of phonemic categorization. Nonetheless, the design is similar in that a minimal pair /su-ʃu/ (as in ‘Sue’ versus ‘shoe’) was chosen as endpoints of a continuum, and a number of steps were created between the two endpoints. In all of the studies, adults and children were tested on tokens forming a continuum using a forced choice between two words, which depended on the identity of the vowel(s) used. For example, if the vowel was /u/, the choices participants had were ’Sue’ and ’shoe’.

2.1. PARTICIPANTS. Table 1 (see APPENDIX A) summarizes some data on participant attrition and sample size. The first two columns specify the experiment to be reported, while the third indicates the average age or age group. In the column labeled “Participants”, the first number indicates the total number of participants reported, the second is participants whose data was eliminated because they could not identify endpoints correctly within a given criterion; and the third number corresponds to participants whose data was eliminated because their speech was less developed than expected (Nittrouer & Studdert-Kennedy, 1987) or because they did not pass a hearing screening (all other papers). Where only one number is reported (Nittrouer, 1992; Hazan & Barrett, 2000), this is because participant attrition was not fully explained. For example, Hazan & Barrett (2000) mention that 36 participants were excluded primarily because of failure to pass a hearing screening, but do not specify the ages of these participants. On the other hand, Nittrouer (1992) did not report subjects excluded. However, in the discussion of results she mentions that 27 3-year-olds began the study but most of them were unable to complete the /a/ vocalic context.

Aside from Nittrouer and Studdert-Kennedy (1987) and Hazan and Barrett (2000), all of the other studies specify a minimum number of correct labeling of the endpoint stimuli. For example, in order to include a participant's data, Nittrouer (2002) that required they labeled endpoint stimuli correctly at least 8 times out of 10. One problem with such a criterion is that there is no straightforward way in which to decide whether the participant is not attending to the task or whether the task or stimuli are problematic. Mayo and Turk (2004), for example, attribute their loss of adult subjects to the fact that, although their stimuli sounded good to the experimenters, it may not have been informative enough to the participants. Where it was reported (e.g. Mayo, 2000), participants' attrition rate was very high for children, and this was assumed to be due to children's general inattention. However, there is, at present, no reason to assume that a good exemplar for an adult will be a good exemplar to children.

The number of participants in each group is sometimes smaller than the number used in cue weighting studies with just adults (e.g. 32 in Experiment 1 of Whalen, 1991), though it is not uncommon to find 8 participants in studies of speech perception in children (Morrongiello, Robson, Best, & Clifton, 1984; Ohde, Haley, & McMahon, 1996). The fact that all of these studies replicate basically the same phenomenon may indicate that this phenomenon is robust enough to be shown even with a small sample size.

2.2. STIMULI CONSTRUCTION. As mentioned in the Design section, all of the reported studies have made used of synthetic continua between two endpoints /s-ʃ/. Specific details on the values of parameters for frication and vocalic portions are summarized in Tables 2 and 3.

All of the studies have used productions from a male speaker to base their stimuli on. Except for Hazan and Barrett (2000), all of the studies summarized here processed fricative and vocalic portions in slightly different ways. In these studies, vocalic portions were completely natural (Nittrouer & Studdert-Kennedy, 1987; Nittrouer, 1992; Nittrouer & Miller, 1997b), stylized syntheses based on measurements (Nittrouer & Miller, 1997b, 1997a), or copy-synthesized (Mayo, 2000; Mayo & Turk, 2004). Frication was synthesized using a single pole (Nittrouer, 1992; Nittrouer & Miller, 1997b; Mayo, 2000), a pole and a zero (Nittrouer & Studdert-Kennedy, 1987), or by combining the two frication noises at different amplitudes (Nittrouer & Miller, 1997a).

As a result, most of these studies have used somewhat simplified frication noises, given that there are many other differences in acoustic parameters between English /s/ and /ʃ/ (Jongman, Wayland, & Wong, 2000). They also differ very little in the values for the parameters of frication noise, in spite of the fact that they have based the pole values on measurements of at least two different talkers and three different vocalic context, both values that are known to affect peak location (Hughes & Halle, 1956; Soli, 1981). Most experiments used a center frequency of 3.8 for /s/ and 2.2 for /ʃ/ for the single pole (Nittrouer & Studdert-Kennedy, 1987; Nittrouer, 1992; Nittrouer & Miller, 1997b, 1997a; Mayo, 2000). These values are surprising in view of literature describing the acoustic characteristics of these sibilants. For example, according to Jongman et al. (2000), /s/ has its most prominent peak at 4-5 kHz and /ʃ/ at 2.5-3 kHz. This may have been an artifact of using a low sampling frequency in the earliest studies. The values reported in Mayo and Turk (2004) lie closer to what has been found in the literature, although higher than those, with the pole for the /s/-appropriate frication located at 5.8 kHz and at 3.1 kHz for the /ʃ/.

A different method was used by Hazan and Barrett (2000), who copy-synthesized the whole syllables /su/ and /ʃu/. The frication was made up of noise excitation passing through five formants and was therefore more complex than the fricatives used in other studies. Both amplitude and frequency were varied along the continuum. If we take the formant with the highest amplitude to parallel the single pole used in the other studies, the characteristic frequency of both sibilants are even higher than that of Mayo and Turk (2004) at 6.8 kHz for /s/ and 4 kHz for /ʃ/. Another difference with the studies summarized above lies in how stimuli were presented.

In Hazan and Barrett (2000) there were three kinds of stimuli: combined cue, where both transitions and frications cued the same place of articulation; frication-cue, where transitions where fixed at a midpoint between the extreme values but frication varied in six steps between the two endpoints; and transition-cue, where frication was kept at a midpoint (no information is given as to the values of this midpoint) and transitions were varied in six steps along the continuum.[1] These many differences in stimuli, procedure and goals do not allow easy comparisons between this paper and others, although the fact that children were also able to use the frication continuum in labeling suggests that, at least when presented with more informative frications, children are able to make use of this dimension in speech processing.

In short summary, most of the literature on /s-ʃ/ perception has used rather simplified frications that assume that the salient cue in this portion is the peak location. Although their success in revealing developmental differences speaks in favor of this method, it is not granted to conclude that young children rely on dynamic aspects more than static aspects (cf. Nittrouer, 2002). It could be the case that children are more biased than adults to attend to the most informative aspects of the signal within a given setup. Thus, in these studies were frication has been simplified to a great extent but vocalic transitions have not, children would be biased to attend to the latter more than adults are. I will return to this point in the general discussion.

3. ANALYSES. In earlier work (Nittrouer & Studdert-Kennedy, 1987; Nittrouer, 1992; Nittrouer & Miller, 1997b, 1997a), two measures are used in order to estimate the impact that each cue has on a participant group: the slope of the identification function and the placement of the phoneme boundary.

The slope is defined as the mean rate of change in the response introduced by a change along the fricative continuum, while the phoneme boundary is found through the intersection of the identification function with the 50% response line. Since there is one identification function for each transitional context, each one will have a slope and define a phoneme boundary. Figure 1 shows an example of the different identification functions found for children and adults.

Notice that the functions are shallower for children (hence the slopes are different from those of adults) and that the intersection with the 50% line (5 responses) of each of the two identification functions is more separated for the children than for the adults. While Nittrouer argues that the slope of the identification function is indicative of the weight assigned to

GO BACK TO TABLE OF CONTENTS

frication, Mayo (2000) relates differences in slope across age groups to the fact that children tend to have more categorical responses as they age. This resonates with the view from adult cue weighting that analyzes slope as a sign of how certain participants are of their responses (e.g. Allen & Miller, 2004) and with work showing that phonemic categorization is still developing into adolescence (Hazan & Barrett, 2000). On the other hand, Mayo (2000) and Mayo and Turk (2004) agree with Nittrouer in the use of the difference between the two phoneme boundaries to describe reliance on transitions.

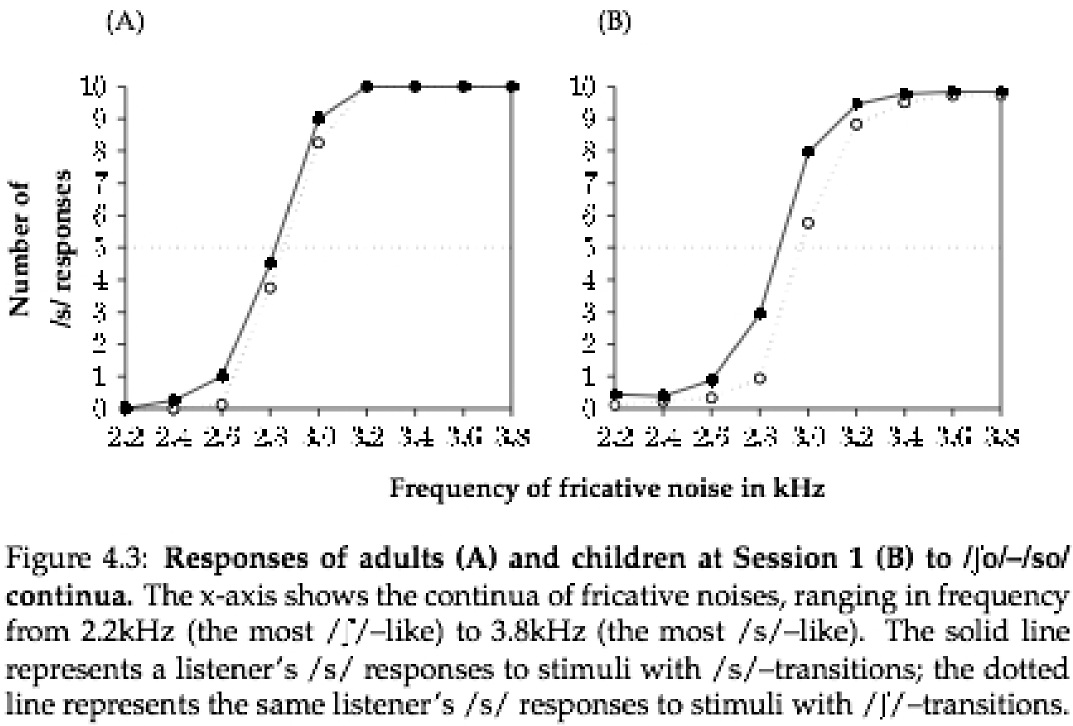

FIGURE 1. Sample identification functions from Mayo (2002). Reproduced with permission of the author.

Since change in the phoneme boundary is a comparative measure, it is useful to estimate the impact of transitional information across ages. However, it does not translate easily into the weight given to this dimension. Later work by Nittrouer (Nittrouer et al., 2000; Nittrouer, 2002) estimates the weight assigned to the two dimensions by doing regression analyses with proportion of responses for a category as dependent measure and the two dimensions as independent variables. They report the mean partial correlation coefficients as indicative of the weight given to each dimension.

Finally, Hazan and Barrett (2000) compared performance across the three types of stimuli presented (combined cue, frication-only and transitions-only). In this case, cues are not pitted against each other, but neutralized. Therefore, this comparison shows whether a cue may be sufficient, but does not reveal whether listeners will pay more attention to one or the other dimension.

Although using one of the above methods would allow comparisons with previous literature, other options that might be borne in mind are similarity metrics and response times. Similarity judgments require participants to make an overt response by deciding how similar two tokens may be after they are presented one after the other with a short interstimulus interval (used in e.g. Johnson, 2004). Response time measured from onset of stimuli to the moment in which the participant enters a response or label has been successfully used for native and non-native contrasts (see, for example, McGuire, 2007), although this is unlikely to be useful with very young children. Another alternative is not to require an overt response but instead to present the labels (e.g. Sue and shoe) as pictures on a screen and then measure the time it takes participants to fixate on a label after presentation of the verbal stimuli (as in Eberhart et al., 1995) and take overall looking time to target as a measure of accuracy.

4. DISCUSSION**.** The work reviewed in this report has focused mainly on whether children and adults differ in their perception of /s-ʃ/. Overall this work can be read as a positive answer to this question. Children and adults do differ in that the former tend to show a larger separation between the labeling functions of stimuli with transitions appropriate to the two fricatives. Despite the unanimity with regards this conclusion, the reasons behind this change are still a source of debate.

Specifically, Nittrouer and colleagues (Nittrouer & Studdert-Kennedy, 1987; Nittrouer, 1992; Nittrouer & Miller, 1997b, 1997a; Nittrouer et al., 2000; Nittrouer, 2002) attribute the change that children must undergo in order to gain adult-like cue weighting schemes to a "Developmental Weighting Shift" (DWS), which has been explained as follows:

> [T]he DWS proposes that young children initially focus their perceptual attention on general movements of the vocal tract, as conveyed by patterns of changing formant frequencies. This perceptual strategy would meet the needs of novice language users who are just

> learning the fundamentals of how to move their own vocal tracts for the purposes of communication.
> (Nittrouer, 2002, p.718)

Alternatively, the change in weighting scheme would not follow from the more extensive experience in producing sounds, but from a development of the perceptual system. According to Sussman (2001) children initially focus on cues that are easier to perceive by virtue of being longer, louder or more contextually informative, but as their central auditory system develops, children can converge on the weighting scheme that is appropriate to their language.

Both explanations rely on the sole difference between children and adults being their stage in development, either in production or perception. After assessing differences in labeling across age groups in different consonantal and vocalic contexts, Mayo and Turk (2005) conclude that neither hypothesis explains all of the differences in cue weighting. A third account that has been put forward is the increased phonemic awareness children gain as they receive reading instruction. Mayo (2000) shows that phonemic awareness in a group of children beginning reading instruction predicts most of the variance in the separation between the two labeling functions corresponding to the vocalic portions from the two fricative contexts.

The studies reported in Mayo (2000) make us a great deal wiser with respect to changes between childhood and adulthood. In particular, they suggest that in order to investigate cue weighting in early language acquisition, the end state should not be the adult cue weighting, given that it may be influenced by phonemic awareness. Furthermore, Mayo and Turk (2005) shows that no hypothesis put forward up to now can explain how children allocate their attention among different cues. Therefore, if positing children's cue weighting as the end state of acquisition, one should be very cautious as to what are the reasons underlying this weighting.

Finally, one potentially important factor for investigating cue weighting across development is the possible differences in the input different age groups are exposed to. Infants', children's and adults' input may not be the same, for several reasons. First, while both adults and children hear themselves, adults' productions are probably within the same parameters of the general population, while even in typical children production of /s-ʃ/ is still developing at seven years of age (Nittrouer, 1995). Given that clarity in production and discrimination abilities for this contrast have been found to be correlated (Perkell et al., 2004), young children who are already attempting these sounds may be influenced by their own (possibly off-target) productions.

Second, the average control adult in these studies (generally, college students) has been exposed to a few hundreds or thousands of tokens by many different speakers of both sexes, while infants and children, especially young ones, may be primarily exposed to child-directed speech from a reduced number of speakers. This is particularly problematic given that /s-ʃ/ vary a great deal across speakers, and to a certain extent within speakers. This variability affects listeners' ability to perceive the difference between the two sounds, such that when presented with a speaker for whom the categories are overlapping, listeners' responses are significantly slower (Newman, Clouse, & Burnham, 2001). Therefore, children who are primarily exposed to the speech of a small group of people may have difficulty adapting to the novel speaker presented to them in testing. Due also to this adaptation, it could be the case that infants are more heterogeneous as a group, with each of them having learned to rely on the specific cues which separate the /s-ʃ/ categories in their caregiver's speech.

Even beyond the question of how adapted to a single speaker infants may be, it is not a trivial assumption that children and adults are exposed to the same cue distributions. Unfortunately, the question of whether acoustic cue reliability in child-directed speech (CDS) differ from adult-directed speech (ADS) has yet to be addressed in the literature, and up to now relatively little is known about cue distribution in CDS. Nonetheless, studies on the general characteristics of these registers, together with evidence regarding the effect of other factors on cue distribution, suggest that there may be some differences in a cue's strength in CDS and ADS. The following are just two ways in which these registers differ, and some potential acoustic consequences for cue distributions. The examples come from voicing, which is a much more studied aspect than place of articulation in fricatives, although similar findings are predicted in this domain. For example, CDS consists of shorter utterances with a lower speech rate (Gallaway & Richards, 1994). The fact that utterances are shorter may imply that children's segmental representations of, for example, word-final /p/ will include more /p/ tokens that come from a boundary edge. A wealth of findings point to strengthening of acoustic cues at prosodic boundaries (Fougeron & Keating, 1997; Cho & Keating, 2001; Tabain, 2003), and to coarticulation patterns being affected by the presence of prosodic boundaries as well as speech rate (Hertrich & Ackermann, 1995; Cho, 2004), which may affect segmental representations. Second, vowels are significantly longer in Infant-Directed Speech and CDS as compared to ADS even when speech rate is taken into account (Kuhl et al., 1997; Englund, 2005; Englund & Behne, 2006). Further, the vowel space is expanded in those registers (Kuhl et al., 1997). Both factors would make transitional cues more salient in speech to young children. In short, with these many differences between the two registers, it is uncertain whether one can safely assume that infants and young children will have been exposed to the same cue distributions as adults.

To conclude, the studies summarized and compared above show a robust finding: when provided with the same stimuli, adults and children label them differently. A host of reasons may underlie this finding. As noticed in previous literature, three important changes could be development in production (which would help children focus on more static states; Nittrouer, 2002), a shift towards an input-based strategy (rather than a salience-based one; Sussman, 2001); and phonemic awareness (Mayo, 2000). I pointed out two other possibilities. First, the stimuli generation methods in many of these papers are problematic for two reasons: they make assumptions of what listeners are relying on within the vocalic and frication dimensions; and they assume that children will be as able as adults to deal with synthetic speech that has incongruent cues. Second, the acoustic cue distributions may be different in the input to these two populations. Both of these problems conspire to minimize the chances that the developing group will succeed, given that they entail the reduction of a dimension, e.g. frication, to a single acoustic cue, e.g. a single pole; and this pole will be at a frequency that is measured from male, adult-directed speech, which may not be the most frequent realization in the child's input. Future research using more natural stimuli and independent corroboration that cue distributions and values are similar in the children's and adults' input is necessary before concluding that, in the case of the /s-ʃ/ contrast, children cannot rely on the frication portion as much as adults do.

REFERENCES

Allen, J. S., & Miller, J. L. 2004. Listener sensitivity to individual talker differences in voice-onset-time. *Journal of the Acoustical Society of America, 115*, 3171–3183.

Cho, T. 2004. Prosodically conditioned strengthening and vowel-to-vowel coarticulation in English. *Journal of Phonetics, 32*, 141–176.

Cho, T., & Keating, P. A. 2001. Articulatory and acoustic studies on domain-initial strengthening in Korean. *Journal of Phonetics, 29*, 155–190.

Eberhart, K., Spivey-Knowlton, M.J., Sedivy, J.C., and Tanenhaus, M.K. 1995. ‘Eye movements as a window into real-time spoken language comprehension in natural contexts.' *Journal of Psycholinguistic Research, 24*, 409–436.

Englund, K. 2005. Voice onset time in infant directed speech over the first six months. *First Language, 25*, 219-234.

Englund, K., & Behne, D. 2006. Changes in infant directed speech in the first six months. *Infant and Child Development, 15*, 139-160.

Fougeron, C., & Keating, P. A. 1997. Articulatory strengthening at edges of prosodic domains. *Journal of the Acoustical Society of America, 101*, 3728–3740.

Gallaway, C., & Richards, B. J. 1994. *Input and interaction in language acquisition*. Cambridge: Cambridge University Press.

Hazan, V., & Barrett, S. 2000. The development of phonemic categorization in children aged 6–12. *Journal of Phonetics, 28*, 377–396.

Hertrich, I., & Ackermann, H. 1995. Coarticulation in slow speech: durational and spectral analysis. *Language and Speech, 38*, 159–187.

Hughes, G. W., & Halle, M. 1956. Spectral properties of fricative consonants. *Journal of the Acoustical Society of America, 28*, 303–310.

Johnson, K. 2004. Crosslinguistic perceptual differences emerge from the lexicon. in Agwuele, A., Warren, W., & Park, S.-H. Eds. *Proceedings of the 2003 Texas Linguistics Society Conference* pp. 26-41. Somerville, MA: Cascadilla Proceedings Project.

Jongman, A., Wayland, R., & Wong, S. 2000. Acoustic characteristics of English fricatives. *Journal of the Acoustical Society of America, 10*8, 1252–1263.

Kuhl, P. K., Andruski, J. E., Chistovich, I. A., Kozhevnikova, E. V., Ryskina, V. L., Stolyarova, E. I., et al. 1997. Cross-language analysis of phonetic units in language addressed to infants. *Science, 277*(5326), 684-686.

Lisker, L. I. 1986. 'Voicing' in English: A catalogue of acoustic features signaling /b/ versus /p/ in trochees. *Language and Speech, 29*, 3-11.

Mayo, C. 2000. *The relationship between phonemic awareness and cue weighting in speech perception: Longitudinal and cross-sectional child studies*. Unpublished doctoral dissertation, Queen Margaret University College.

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(6), 3184-3194.

Mayo, C., & Turk, A. 2005. No available theories currently explain all adult-child cue weighting differences. *PSP2005*, 126–129.

McGuire, G. 2007. *Phonetic category learning*. Unpublished doctoral dissertation, Ohio State University.

Morrongiello, B. A., Robson, R. C., Best, C. T., & Clifton, R. K. (1984). Trading relations in the perception of speech by five-year-old children. *Journal of Experimental Child Psychology, 37*, 231–250.

Newman, R. S., Clouse, S. A., & Burnham, J. L. (2001). The perceptual consequences of within-talker variability in fricative production. *Journal of the Acoustical Society of America, 109*, 1181–1196.

Nittrouer, S. 1992. Age-related differences in perceptual effects of formant transitions within syllables and across syllable boundaries. *Journal of Phonetics, 20*, 351–382.

Nittrouer, S. 1995. Children learn separate aspects of speech production at different rates: Evidence from spectral moments. *Journal of the Acoustical Society of America, 97*, 520–530.

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

Nittrouer, S., & Miller, M. E. 1997a. Developmental weighting shifts for noise components of fricative-vowel syllables. *Journal of the Acoustical Society of America, 102*, 572–580.

Nittrouer, S., & Miller, M. E. 1997b. Predicting developmental shifts in perceptual weighting schemes. *Journal of the Acoustical Society of America, 101*, 2253–2266.

Nittrouer, S., Miller, M. E., Crowther, C. S., & Manhart, M. J. 2000. The effect of segmental order on fricative labeling by children and adults. *Perceptual Psychophysics, 62*, 266–284.

Nittrouer, S., & Studdert-Kennedy, M. 1987. The role of coarticulary effects in the perception of fricatives by children and adults. *Journal of Speech and Hearing Research, 30*, 319–329.

Ohde, R. N., Haley, K. L., & McMahon, C. W. 1996. A developmental study of vowel perception from brief synthetic consonant-vowel syllables. *Journal of the Acoustical Society of America, 100*, 3813–3824.

Perkell, J. S., Matthies, M. L., Tiede, M., Lane, H., Zandipour, M., Marrone, N., et al. 2004. The distinctness of speakers' /s/-/ʃ/ contrast is related to their auditory discrimination and use of an articulatory saturation effect. *Journal of Speech, Language, and Hearing Research, 47*, 1259–1269.

Soli, S. D. 1981. Second formants in fricatives: Acoustic consequences of fricative-vowel coarticulation. *Journal of the Acoustical Society of America, 70*, 976–984.

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*(3), 1173-1180.

Tabain, M. 2003. Effects of prosodic boundary on /aC/ sequences: Acoustic results. *Journal of the Acoustical Society of America, 113*, 516–531.

Whalen, D. H. 1991. Perception of the English /s/-/ʃ/ distinction relies on fricative noises and transitions, not brief spectral slices. *Journal of the Acoustical Society of America, 90*, 1776–1784.

APPENDIX A

Authors	Year	Age	Number
Nittrouer and Studdert-Kennedy	1987	Adults	12-0-0
		3	4+1+3
		4	6+1+1
		5	7-0-1
		7	8-0-0
Nittrouer	1992	Adults	16
		3	9
		5	10
		7	10
Nittrouer and Miller	1997; Exp. 1	Adults	20+0+0
		4	20+3+0
		7	20+2+0
Nittrouer and Miller	1997; Exp. 2	Adults	20+0+0
		4	19+4+0
		7	22+0+0
Nittrouer and Miller	1997; Exp. 3	Adults	10+0+0
		3	14+7+0
		5	10+3+0
		7	12+0+0
Nittrouer and Miller	1997b	Adults	21+2+0
		6	21+5+0

GO BACK TO TABLE OF CONTENTS

Hazan and Barrett	2000	Adults	13
		6;6	15
		8 to 9	16
		10	19
		11	16
		12	18
Mayo	2000	Adults	8+0+0
		6 (read.inst.)	18+9+0
		7 (no read.inst.)	8+3+0
Mayo and Turk	2004	Adults	10+5+0
		3 to 4	9+2+0
		5	6+4+0
		7	8+7+0

TABLE 1. *Participants included and discarded in each age group and experiment. The first two columns specify the experiment to be reported, while the third indicates the average age or age group. In the column labeled "Participants", the first number indicates the total number of participants that were included in the analysis, the second is participants whose data was eliminated because they could not identify endpoints correctly a certain number of times; and the third number corresponds to participants whose data was eliminated because their speech or hearing was less developed than expected. More details in the main text.*

Authors	Year	General	Sampling Rate (kHz)	Frication				
				End points	Maxima (kHz)	Step Size (kHz)	Duration (ms)	Notes
Nittrouer and Studdert-Kennedy	1987	s/ + i/u synthetic fricative noise and natural vowel	10	s	3.8 2.2	0.2	210	Zero at .75 x pole value
Nittrouer	1992	s/ + u/a synthetic fricative noise and natural vowel	10	Same as previous			230	Pilot with adults showed no more than one ambiguous token
Nittrouer and Miller	1997; Exp. 1	s/ + u/a synthetic fricative noise and natural vowel	20	Same as previous				
Nittrouer and Miller	1997; Exp. 2	s/ + u/a synthetic fricative noise and synthesized vowel - F3 made uninformative	20	Same as previous				

TABLE 2. *Details on parameters used for the generation of the frication portions of the stimuli.*

Authors	Year	General	Sampling Rate (kHz)	Frication: Endpoints	Maxima (kHz)	Step Size (kHz)	Duration (ms)	Notes
Nittrouer and Miller	1997; Exp. 3	s/ + u/a synthetic fricative noise and synthesized vowel; both F2 and F3 varied	10 (Pilot shows no effect of sampling rate)	Same as previous				
Nittrouer and Miller	1997 b	s/ + u/a synthetic fricative noise and synthesized vowel	20	There are 7 steps, with endpoints: s=/s/+/ / combined at same amplitude; and =/s/+4*/ / combined at a 1/4 amplitude ratio. The major difference among resulting stimuli was the amplitude of a low frequency pole at roughly 2.2 kHz. Relative amplitude of this peak was 12 dB lower for the most s-like noise than for the -like noise.				
Hazan and Barrett	2000	s/ + u copysynthesized	16	There are 6 steps. Noise excitation produced by passing through 5 formants, three of them fixed (F3 approx 3.0 kHz, F4= 4 kHz, F6 = 6.8 kHz), although varying the amplitude of F4 (s: 45 dB to : 53 dB) and F6 (s: 58 dB to : 43 dB). Also, F2 varied in frequency (s: 1.9 to : 2.1 kHz) and amplitude (30-35dB) and F5 in frequency only (s: 5.4 to : 4 kHz)				
Mayo	2000	s/ + u copysynthesized copysynthesized	16	s	3.8 2.2	0.2	230	Single pole frication based on measurements. Amplitude rose from 0 to 60 over first 90ms, falling to 30 db over final 50 ms.
Mayo and Turk	2004	s/ + aI copysynthesized		s	5.8 3.1		155	Single pole frication noise

TABLE 2. (Continued)

Authors	Year	Vowel	Vowel Duration	Onset F1	DurF1Trans	Onset F2	DurF2Trans	Onset F3	DurF3Trans	Notes
				Transitions						
Nittrouer and Studdert-Kennedy	1987	(s)i (s)u ( )i ( )u	320-370							matched in intonational contour and duration
Nittrouer	1992	(s)a (s)u ( )a ( )u	333 347 337 348			1365 1520 1532 1706		2457 2496 2367 2288		5 tokens; when presented to adults in isolation, they performed at chance
		( )u	348			1706	Throughout vowel (to 930 Hz)	2288	100 ms (to 2320 Hz)	
Nittrouer and Miller	1997; Exp. 1	(s)a	333			1365	60-90ms (to 1309 Hz)	2457	100 ms (to 2365 Hz)	Reported as natural and 'similar' to above. Here again 5 vocalic portions - notice average values are same as above.
		(s)u	347			1520	Throughout vowel (to 930 Hz)	2496	100 ms (to 2320 Hz)	
		( )a	337			1532	60-90ms (to 1309 Hz)	2367	100 ms (to 2365 Hz)	

TABLE 3. *Details on parameters used for the generation of the vocalic portions of the stimuli.*

Authors	Year	Vowel	Vowel Duration	Onset F1	DurF1Trans	Onset F2	DurF2Trans	Onset F3	DurF3Trans	Notes
						Transitions				
Nittrouer and Miller	1997; Exp. 2	(s)a	270	450	50 ms (to 650 Hz)	1250	100 ms (to 1130 Hz)	2400	50 ms (to 2100 Hz)	F0 fell over 100 ms from 100 to 80 Hz for /a/, and throughout /u/ from 120 to 100 Hz (This was necessary to maintain similar pitches.)
		(s)u	270	250	(constant F1)	1480	Throughout vowel (to 850 Hz)	2400	50 ms (to 2100 Hz)	
		( )a	270	450	50 ms (to 650 Hz)	1570	100 ms (to 1130 Hz)	2400	50 ms (to 2100 Hz)	
		( )u	270	250	(constant F1)	1800	Throughout vowel (to 850 Hz)	2400	50 ms (to 2100 Hz)	
Nittrouer and Miller	1997; Exp. 3	(s)a	Same as previou					2300	100 ms (to 2300 Hz)	F3 onset frequencies maintain the transition-related difference between /a/ and /u/ found in Exp. 1
		(s)u						2200	130 ms (to 2100 Hz)	
		( )a						2460	100 ms (to 2300 Hz)	
		( )u						2520	130 ms (to 2100 Hz)	
Nittrouer and Miller	1997 b	Vocalic portions same as Experiment 2								
Authors	Year	Vowel	Vowel Duration	Onset F1	DurF1Trans	Onset F2	DurF2Trans	Onset F3	DurF3Trans	Notes
						Transitions				
Hazan and Barrett	2000	(s)u		363	Not reported (offset F1: 256 Hz)	1900	40 ms (to 2063 Hz - offset 1098 Hz)	3050	Not reported (final F3: 2773 Hz)	
		( )u		363	Not reported (offset F1: 256 Hz)	2108	40 ms (to 2063 Hz - offset 1098 Hz)	3050	Not reported (final F3: 2773 Hz)	
Mayo	2000	(s)o ( )o	250	The frequency values were taken from actual values of 5 tokens. These tokens were measured at 4 points in the transitions and 4 in the rest of the vowel. Duration was taken from the average of all the tokens spoken in isolation.						
Mayo and Turk	2004	(s)aI	385	537	Average transition 60 ms (to 762 Hz)	1536	Average transition 60 ms (to 1536 Hz)	2551	Average transition 60 ms (to 2551 Hz)	
		( )aI		435	Average transition 80 ms (to 762 Hz)	1574	Average transition 80 ms (to 1536 Hz)	2400	Average transition 80 ms (to 2551 Hz)	

TABLE 3. (Continued)

AUTHOR NOTE

Correspondence concerning this article should be addressed to:

Alejandrina Cristià

Linguistics

500 Oval Drive

West Lafayette, IN, 47907

[acristia@purdue.edu]

NOTES

[1] Other studies investigating fricative perception have used a rather different method (e.g. Whalen, 1991). This method consists of cross-splicing different proportions of two different fricatives or of the vowels following them. For example, a continuum in the frication noise would start with have as first token 100% of one fricative (e.g. /s/); the second token would be formed with 90% of /s/ followed by 10% of /ʃ/; and so forth until 100% of the frication noise is taken from /ʃ/. To my knowledge, this method has not be used in the literature on cue weighting at different ages, although it constitutes an interesting possibility given that it does not make any a priori assumptions on which cues listeners may be attending to within the frication portion (e.g. peak location, centroid, skewness), on one hand, and the vocalic portion, on the other.