Adult–child differences in acoustic cue weighting are influenced by segmental context: Children are not always perceptually biased toward transitions

Catherine Mayo[a] and Alice Turk
Theoretical & Applied Linguistics, University of Edinburgh, Adam Ferguson Building, 40 George Square, Edinburgh, EH8 9LL, United Kingdom

(Received 18 November 2003; revised 16 March 2004; accepted 21 March 2004)

It has been proposed that young children may have a perceptual preference for transitional cues [Nittrouer, S. (**2002**). J. Acoust. Soc. Am. **112**, 711–719]. According to this proposal, this preference can manifest itself either as heavier weighting of transitional cues by children than by adults, or as heavier weighting of transitional cues than of other, more static, cues by children. This study tested this hypothesis by examining adults' and children's cue weighting for the contrasts /saɪ/-/ʃaɪ/, /de/-/be/, /ta/-/da/, and /ti/-/di/. Children were found to weight transitions more heavily than did adults for the fricative contrast /saɪ/-/ʃaɪ/, and were found to weight transitional cues more heavily than nontransitional cues for the voice-onset-time contrast /ta/-/da/. However, these two patterns of cue weighting were not found to hold for the contrasts /de/-/be/ and /ti/-/di/. Consistent with several studies in the literature, results suggest that children do not always show a bias towards vowel–formant transitions, but that cue weighting can differ according to segmental context, and possibly the physical distinctiveness of available acoustic cues.
[DOI: 10.1121/1.1738838]

PACS numbers: 43.71.Ft [RLD] Pages: 3184–3194

I. INTRODUCTION

It is well established that listeners do not pay equal attention to all acoustic information available to them in the speech stream. Instead, when perceiving speech, listeners give some acoustic cues more attention than others (Dorman *et al.*, 1977; Ohde and Haley, 1997; Walley and Carrell, 1983; Wardrip-Fruin, 1982, 1985; Whalen, 1991). These patterns of acoustic cue weighting appear to change developmentally: Children have been found to show different cue weighting strategies from adults. Nittrouer and colleagues, for example, have found that in identifying fricative contrasts (e.g., /su/-/ʃu/) based on frequency of frication noise and vowel-onset formant transition cues, young children give more attention or weight to the formant transitions than do older children and adults (Nittrouer and Studdert-Kennedy, 1987). Other researchers have also found differences between children and adults in their relative weighting of acoustic cues (Krause, 1982; Lacerda, 1992; Mayo *et al.*, 2003; Morrongiello *et al.*, 1984; Ohde and Haley, 1997; Parnell and Amerman, 1978; Watson, 1997; Wardrip-Fruin and Peach, 1984).

Nittrouer and colleagues (Nittrouer, 1993; Nittrouer and Miller, 1997) have suggested that these differences in cue weighting between adults and children are not random, but are related to developmental changes in the way in which listeners process speech. This hypothesis, called the Developmental Weighting Shift (DWS) theory, is based on the premise that children start out processing speech globally, in terms of large units such as syllables or monosyllabic words. With development, it is proposed, processing becomes more analytical, such that adults parse speech in terms of smaller units (e.g., Jusczyk and Derrah, 1987; Menn, 1971; Studdert-Kennedy, 1987, although note that contrasting views exist, see, e.g., Dollaghan, 1994; Gerken *et al.*, 1995). The DWS proposes that this change in processing impacts on speech perception development, and that it does so in terms of the acoustic cues that listeners attend to or weight most heavily. That is, children and adults show different patterns of acoustic cue weighting because they process speech in terms of different sized units.

Nittrouer has gone on to suggest that the acoustic correlates of more global speech perception could be syllable-internal formant transitions, because these cues are "perceptually salient and delimit signal portions corresponding to syllables" (Nittrouer *et al.*, 2000, p. 268). In other words, children, as more global perceivers, should be perceptually biased toward making heavier use of vowel-formant transitional cues. How this bias is manifested is not entirely straightforward. Initially, based on the results of the /s-vowel/-/ʃ-vowel/ study noted earlier, it was proposed that children give more weight to transitional cues than do adults. This was indeed found to be the case for a number of other /s-vowel/-/ʃ-vowel/ studies (Nittrouer, 1992; Nittrouer and Miller, 1997; Mayo *et al.*, 2003; Watson, 1997), but a later study of /f-vowel/-/θ-vowel/ perception (Nittrouer, 2002) showed no significant difference between adults and children in their weighting of transitional cues. This /f-vowel/-/θ-vowel/ study did show, though, that for this contrast children (and adults) gave more weight to transitional cues than to the other available cue, frequency of frication noise. In light of this, Nittrouer (2002) suggested that for those contrasts in which nontransitional cues are particularly uninformative to both adults and children (such as in /f/-/θ/, see, e.g., Harris,

[a]Electronic mail: catherin@ling.ed.ac.uk

 0001-4966/2004/115(6)/3184/11/$20.00

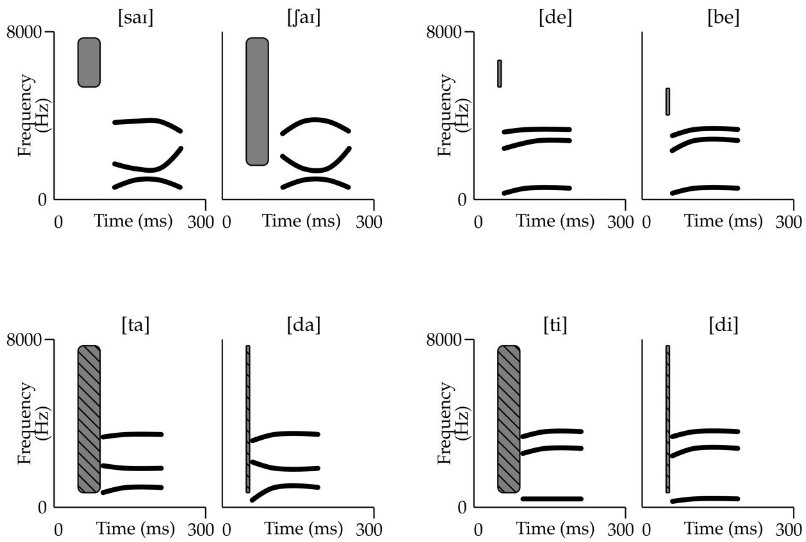

FIG. 1. Stylized spectrograms of prototypical tokens of the contrasts used in this study. The gray boxes represent frication noise (/saɪ/-/ʃaɪ/), stop burst (/de/-/be/), and burst+aspiration (/ta/-/da/,/ti/-/di/). The black lines represent vowel formants (F1,F2,F3).

1958) children's preference for transitional information could operate acoustically, leading children to weight transitional cues more heavily than other acoustic cues to the same contrast. In summary, according to this theory, children should weight transitions either (i) more heavily than do adults (what we will call the *developmental transitional bias hypothesis*), or (ii) more heavily than they weight other acoustic cues (what we will call the *acoustic transitional bias hypothesis*). Furthermore, there should not exist a contrast for which children are *not* biased in one of these two ways toward vowel-formant transitional cues.

Evidence for the two aspects of this theory is equivocal. Supporting the developmental transitional bias hypothesis, studies have shown that children are more influenced by transitional cues than are adults in determining voicing in final stops (Krause, 1982; Wardrip-Fruin and Peach, 1984), in identifying place of articulation of some initial stops (particularly /g/, Ohde *et al.*, 1996; Ohde and Haley, 1997), and in determining voicing for some initial stops (/biz/-/piz/ with long VOT, and /bat/-/pat/ with both long and short VOT, Howell *et al.*, 1992). However, children appear to be less influenced by transitions than older listeners for some VOT contrasts (/got/-/kot/, Simon and Fourcin, 1978, /biz/-/piz/ with short VOT, Howell *et al.*, 1992), and for identifying vowels (Malech and Ohde, 2003; Sussman, 2001). Furthermore, while the acoustic transitional bias hypothesis would predict that in these latter cases children should give more weight to transitional than to nontransitional cues, this does not appear to be the case: In the same studies, children gave *less* weight to transitional than to nontransitional cues. Children were found to give more weight to vowel duration (Ohde *et al.*, 1996; Ohde and Haley, 1997) or to steady-state formant frequencies (Sussman, 2001) than to formant transitions when identifying vowel contrasts. In identifying voicing in initial stops, children have been consistently found to weight VOT more heavily than vowel-onset transitions (Simon and Fourcin, 1978; Howell *et al.*, 1992).

It appears from these studies that children may not always attend more closely to transitional cues. The studies suggest that adult–child differences in cue weighting may change both with the segmental context of the contrast, and with the acoustic characteristics of the cues available to signal the contrast. Based on evidence such as this, a number of researchers have suggested alternatives to the DWS, most notably an account based on general auditory processing differences between adults and children (e.g., Elliott *et al.*, 1981; Elliott and Busse, 1984; Eisenberg *et al.*, 2000; Sussman, 1993, 2001). This type of explanation generally proposes that because children have less well developed auditory systems than do adults, they will have trouble processing acoustic cues that are not physically distinct. However, there are problems with dismissing the two transitional bias hypotheses based only on evidence from the above-noted studies. In particular, the variety of different methods and stimulus types (natural, modified-natural, or synthetic speech) used across these studies makes direct comparisons between the results complicated at best.

It therefore remains unclear to what extent the transitional bias hypothesis holds for different segmental contexts if experimental conditions remain constant. Our goal, therefore, was to examine cue weighting strategies across multiple contrasts using the same methodology for each. Two different place of articulation contrasts were used in the study: The fricative contrast /saɪ/-/ʃaɪ/, and the stop burst contrast /de/-/be/. Two voice-onset time (VOT) contrasts were chosen, due to the contradictory evidence found previously for VOT: /ta/-/da/ and /ti/-/di/.[1] Figure 1 shows stylized spectrograms of all four contrasts. For each contrast, we tested adults' and children's weighting of (i) a vowel-onset formant transition cue, and (ii) one of the following nontransitional cues: Frequency of frication noise for /saɪ/-/ʃaɪ/, frequency of stop burst for /de/-/be/, and duration of VOT for /ta/-/da/ and /ti/-/di/.

Two types of analysis were carried out on the results. First, adults' and children's cue weighting strategies were compared for each of the different contrasts. This allowed for an examination of the possible role of segmental context in adult–child cue weighting differences. This also enabled us to determine the scope of children's possible transitional bias as it relates to adults' cue weighting patterns. Second, for a set of contrasts which did not support the developmental transitional bias hypothesis, logistic regression was employed to compare the weight given to transitional cues and the weight given to nontransitional cues. This allowed us to test the acoustic version of the transitional bias hypothesis.

It should be noted that varying contrasts also means varying the physical distinctiveness of the available transitional and nontransitional acoustic cues in those contrasts. For example, in the contrast /su/-/ʃu/, the difference in F2 frequency at vowel onset is relatively larger (more distinct) than the difference in F2 frequency at vowel onset in the contrast /si/-/ʃi/. Nittrouer (1992) has shown that this type of difference in physical distinctiveness impacts on listeners' perception for /s-vowel/-/ʃ-vowel/ contrasts. Both adults and children were found to weight transitional cues less when perceiving /si/-/ʃi/ than when perceiving /su/-/ʃu/ (although note that children were consistent in weighting transitional cues more heavily than adults for both contrasts). This suggests that the transitions in the /si/-/ʃi/ contrast provided less useful information overall. Similar differences in physical distinctiveness, and thus presumably in informativeness, should be expected for the transitional cues manipulated in this study. Additionally, the nontransitional cues used in the current study (frication noise, stop burst, voice onset time) are also likely to differ in their perceptual informativeness, although quantifying the difference in informativeness between three different types of cue is more complex than comparing the informativeness of two different sets of one type of cue. Therefore, while this is not the focus of the current study, the fact that informativeness often co-varies with segmental context will need to be borne in mind when analyzing the results.

II. METHOD

A. Participants

For the /saɪ/-/ʃaɪ/ contrast, 10 adults (age range of 20–35 years, average age 26 years) out of 15 adults tested,[2] 8 seven-year-olds (age range 7;3–7;10 [year;month], average age 7;6) out of 15 seven-year-olds tested, 6 five-year-olds (age range 5;1–5;10, average age 5;4) out of 10 five-year-olds tested, and 9 three- to four-year-olds (age range 3;7–4;9, average age 4;3) out of 11 three- to four-year-olds tested met the testing criterion (described below).

For the /de/-/be/ contrast, 7 out of 7 adults tested (age range 21–33 years, average age 26 years) met the testing criterion, but only 2 seven-year-olds (both 7;5) out of 10 seven-year-olds tested, 2 five-year-olds (5;5 and 5;7) out of 7 five-year-olds tested, and 1 three- to four-year-old (4;0) out of 8 three- to four-year-olds tested met the criterion.

For the /ta/-/da/ contrast, 8 out of 8 adults tested (age range 21–49 years, average age 33 years), 10 seven-year-olds (age range 7;0–7;11, average age 7;7) out of 11 seven-year-olds tested, 9 five-year-olds (age range 5;1–5;8, average age 5;5) out of 10 five-year-olds tested and 12 three- to four-year-olds (age range 3;0–4;11, average age 4;0) out of 18 three- to four-year-olds tested met the testing criterion.

For the /ti/-/di/ contrast, 8 out of 8 adults tested (age range 21–49 years, average age 33 years), 10 seven-year-olds (age range 7;0–7;11, average age 7;7) out of 11 seven-year-olds tested, 9 five-year-olds (age range 5;1–5;8, average age 5;5) out of 10 five-year-olds tested and 9 three- to four-year-olds (age range 3;7–4;11, average age 4;1) out of 18 three- to four-year-olds tested met the testing criterion.

All of the five- and seven-year-olds were in full-time primary education (first and third year) in Edinburgh (Scotland). The three- to four-year-old children were selected from independent and school-associated nursery (pre-school) classes. All of the children were monolingual native speakers of Scottish Standard English (SSE), and all performed appropriately for their age on standardized tests of reading (Schonell Graded Word Reading Test, Schonell and Goodacre, 1971) and receptive vocabulary (BPVS, Dunn *et al.*, 1997). Parental questionnaires determined that all of the children and their siblings were free from speech/language disorders, hearing deficits and histories of chronic otitis media (defined as more than three ear infections between birth and 3;0, and/or the implantation of myringotomy tubes, see Nittrouer, 1996). No child was tested if he or she was suffering from, or had suffered from at any point in the weeks preceding the test session, any upper respiratory infection.

All adults were monolingual native speakers of English living in Edinburgh (average duration of time in Scotland: 12 years). All of the adults reported themselves as being free from speech/language disorders, hearing deficits, and histories of chronic otitis media. Again, no adult was tested if he or she was suffering from, or had suffered from at any point in the weeks preceding the test session, any upper respiratory infection.

B. Stimuli

The contrasts used in this study were /saɪ/-/ʃaɪ/, /de/-/be/, /ta/-/da/ and /ti/-/di/.[3] Continua of synthetic speech sounds were created for each contrast. The end points of the synthetic continua were copy-synthesized versions of the above-noted syllables based on detailed acoustic analysis of natural tokens spoken by a male native speaker of SSE (aged 39 years, with normal speech, language and hearing). The stimuli were created using SENSYN (Sensimetrics Org.), a cascade/parallel formant synthesizer based on Klatt (1980).

In order to make direct comparisons between the results of this study and those of Nittrouer and colleagues, the design of the stimuli followed the modified trading relations paradigm used by Nittrouer in most of her studies of /s/-/ʃ/ contrasts (e.g., Nittrouer and Studdert-Kennedy, 1987; Nittrouer, 1992). In this paradigm, two continua of speech sounds are created in which (at least) two acoustic cues are manipulated. One of the two cues is varied *along* both continua. The two continua are therefore identical in terms of this cue. The other cue is varied *across* the two continua. The two continua therefore differ in terms of this cue. In most of Nittrouer and colleagues' studies of cue weighting for /s/-/ʃ/ contrasts (see, e.g., Nittrouer, 1992), the two cues manipulated were frequency of frication noise, and vowel-onset formant transitions. The frication pole varied along both continua from a frequency appropriate for /s/ to one appropriate for /ʃ/. The vowel-onset formant transitions varied across the two continua: One continuum had transitions that were appropriate for a preceding /s/, the other continuum had transitions that were appropriate for a preceding /ʃ/ (see Fig. 2).

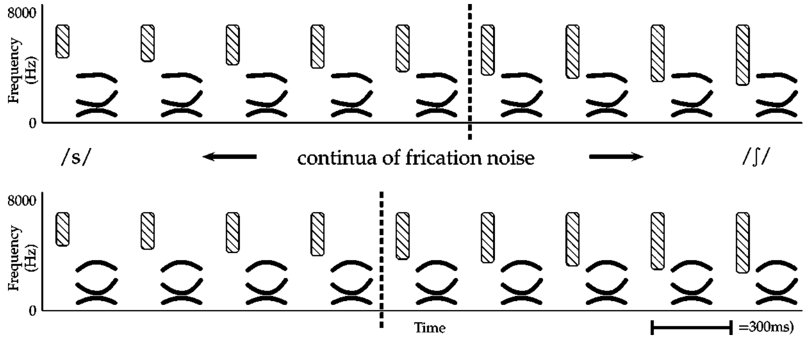

FIG. 2. Stylized spectrograms of two /saɪ/-/ʃaɪ/ continua. The top continuum has /s/-transition vowels, the bottom continuum has /ʃ/-transition vowels. The dashed lines represent hypothetical category boundaries for a listener whose perception is influenced by the transitions.

This type of design allows for an investigation of the perceptual effect of the two cues. A listener who is not influenced by the cue that changes across the continua (e.g., the transitional cue in the /s/-/ʃ/ example above) will perceive the two continua as the same. In contrast, a listener who *is* influenced by the cue that changes across the continua will perceive the two continua differently (see the category boundaries marked in Fig. 2).

The current study followed the general design of those carried out by Nittrouer and colleagues by manipulating one nontransitional cue and one transitional cue for each contrast. The nontransitional cues were: Frequency of frication noise for /saɪ/-/ʃaɪ/, frequency of stop burst for /de/-/be/, and duration of VOT for /ta/-/da/ and /ti/-/di/. The transitional cues were the frequency and time-varying properties of vowel-onset formant transitions in two conditions: (i) appropriate for having followed the first consonant in the contrast, or (ii) appropriate for having followed the second consonant in the contrast. As these transitional cues were closely modeled on natural speech they had gradually changing formant frequencies rather than the straight line transitional slopes typical of some synthetic stylisations. Two 9-point continua were created for every contrast, in which the nontransitional cue varied along both continua and the transitional cue varied across the two continua. Following Nittrouer (1992), five different repetitions of the same vowel were synthesized for each transition condition. This was done to enhance the naturalness of the synthetic speech by capturing a small amount of within-speaker variability. Subsequent examination of the results showed no obvious influence of any one vowel token on listeners' responses. Each of these vowels was combined with the 9 continuum values, resulting in 90 stimuli per contrast. In the following are details of the parameters manipulated for each contrast.

1. /saɪ/-/ʃaɪ/

Nine different single-pole frication noises were synthesized, ranging from 3100 Hz (most /ʃ/-like) to 5800 Hz (most /s/-like). Two sets of /aɪ/ vowels were created, one with onset frequencies appropriate for a preceding /s/ and one with onset frequencies appropriate for a preceding /ʃ/. The average /ʃ/-transition formant onset frequencies were F1: 435 Hz, F2: 1574 Hz, F3: 2400 Hz; the average /s/-transition formant onset frequencies were F1: 537 Hz, F2: 1536 Hz, F3: 2551 Hz. The average vowel target values for the /a/ portion of the diphthong for all 10 synthetic vowels were F1: 762 Hz, F2: 1184 Hz, F3: 2784 Hz. The vowel target values for the /ɪ/ portion of the diphthong for all 10 synthetic vowels were F1: 448 Hz, F2: 1958 Hz, F3: 2419 Hz (see Appendix A for all values used for synthetic stimuli).

The total duration of each syllable was 540 ms, with 155 ms of frication noise and 385 ms of vowel. The average duration of vowel formant transitions as measured from vowel onset to vowel steady state was 60 ms for /s/-transition stimuli and 80 ms for /ʃ/-transition stimuli.[4] F0 for each complete syllable began at 140 Hz at onset of voicing, rose to 150 Hz 110 ms after onset of voicing, and fell to 90 Hz at vowel offset.

2. /de/-/be/

Nine different complex bursts were synthesized. The spectral shape of the bursts was modeled by means of three spectral peaks, one of which was designed to model a cavity at the front of the mouth.[5] The amplitude of these bursts ranged from 54 dB at 5550 Hz, 36 dB at 2700 Hz, and 20 dB at front cavity peak (most /d/-like) to 6 dB at 4500 Hz, 0 dB at 2100 Hz, and 50 dB at front cavity peak (most /b/-like). Two sets of /e/ vowels were created, one with onset frequencies appropriate for having followed /d/ and one with onset frequencies appropriate for having followed /b/. The average /d/-transition formant onset frequencies were F1: 220 Hz, F2: 1809 Hz, F3: 2446 Hz; the average /b/-transition formant onset frequencies were F1: 257 Hz, F2: 1694 Hz, F3: 2247 Hz. The average vowel target values for all 10 synthetic vowels were F1: 428 Hz, F2: 2116 Hz, F3: 2539 Hz.

The total duration of each syllable was 400 ms, with 15 ms of burst, and 385 ms of vowel. The average duration of vowel formant transitions as measured from vowel onset to vowel steady state was 100 ms for /d/-transition stimuli and 110 ms for /b/-transition stimuli. F0 for each complete syllable began at 140 Hz at onset of voicing, rose to 150 Hz 110 ms after onset of voicing, and fell to 90 Hz at vowel offset.

3. /ta/-/da/

Nine different VOT values were synthesized, varying in 5 ms steps from 40 ms (most /t/-like) to 0 ms (most /d/-like). Two sets of /a/ vowels were created, one with onset frequencies appropriate for having followed /t/ and one with onset frequencies appropriate for having followed /d/. The average /ta/-transition formant onset frequencies were F1: 537 Hz, F2: 1536 Hz, F3: 2551 Hz; the average /da/-transition formant onset frequencies were F1: 261 Hz, F2: 1642 Hz, F3:

2472 Hz. The average vowel target values for all 10 synthetic vowels were F1: 711 Hz, F2: 1433 Hz, F3: 2665 Hz.

The total duration of each syllable ranged from 315 ms for the shortest VOT to 355 ms for the longest VOT, with 315 ms of vowel. The average duration of vowel formant transitions as measured from vowel onset to vowel steady state was 55 ms for /t/-transition stimuli and 85 ms for /d/-transition stimuli. F0 for each complete syllable began at 124 Hz at onset of voicing, rose to 130 Hz 90 ms after onset of voicing, and fell to 60 Hz at vowel offset.

4. /ti/-/di/

The VOT values used for the /ti/-/di/ contrast were the same as those used for the /ta/-/da/ contrast. These varied in 5 ms steps from 40 ms (most /t/-like) to 0 ms (most /d/-like). Two sets of /i/ vowels were created, one with onset frequencies appropriate for having followed /t/ and one with onset frequencies appropriate for having followed /d/. The average /ti/-transition formant onset frequencies were F1: 311 Hz, F2: 1924 Hz, F3: 2599 Hz; the average /di/-transition formant onset frequencies were F1: 221 Hz, F2: 1893 Hz, F3: 2569 Hz. The average vowel target values for all ten synthetic vowels were F1: 309 Hz, F2: 2183 Hz, F3: 2819 Hz.

The total duration of each syllable ranged from 315 ms for the shortest VOT to 355 ms for the longest VOT, with 315 ms of vowel. The average duration of vowel formant transitions as measured from vowel onset to vowel steady state was 110 ms for /t/-transition stimuli and 105 ms for /d/-transition stimuli. F0 for each complete syllable began at 124 Hz at onset of voicing, rose to 130 Hz 90 ms after onset of voicing, and fell to 60 Hz at vowel offset.

C. General procedure

With the exception of the subjects who heard the VOT contrasts, each subject was asked to listen to only one set of contrasts; those subjects who heard the VOT contrasts listened to both the /ta/-/da/ and the /ti/-/di/ contrasts. All subjects were tested individually in a quiet room. The stimuli were presented over headphones (Sennheiser HD 490, frequency response 17–22000 Hz) via a CD player. Volume was set at a comfortable listening level. Each subject was asked to indicate that the level was both comfortable and audible (for the child subjects, the signal was split to two headphones and the chosen listening level was monitored by the experimenter); very few adjustments to the level were made by the subjects. No adjustments to listening level were made within the presentation of a single contrast. Testing for the child subjects took place over two or three days. Testing for the adult subjects took place on one day, with a short break half-way through testing.

All subjects were introduced to the target words for their contrast. The child subjects were also familiarized with pictures that corresponded to each word in their contrast (see Appendix B for a description of the pictures used in the study). During testing, the children indicated which word they had heard by saying the word aloud, and by placing a counter on the relevant picture. Before testing, the children were given an opportunity to practice responding to natural productions of the target words. This ensured that the children were able to identify the targets in natural speech, and that they clearly associated each picture with the relevant target. The children received feedback throughout this practice, and did not proceed to the pretest with synthetic stimuli until they had, unprompted, correctly identified a complete set of 10 randomly presented natural stimuli (5 of each CV syllable).

A pretest was administered to both child and adult subjects to ensure that they understood the task. This test consisted of the congruent end points of the continua, that is, the end point values of the nontransitional cue followed by the congruent vowel-formant transitions for each nontransitional cue condition. For example, the congruent end points for the /ta/-/da/ contrast were the 40 ms VOT plus vowels with /t/-transitions (the most /ta/-like stimuli) and the 0 ms VOT plus vowels with /d/-transitions (the most /da/-like stimuli). There were 10 stimuli in the pretest (5 per congruent end point), presented in random order. No feedback was given during this pretest.

For each contrast, five different random orders of the 90 stimuli were created. During the main test, the five-year-old, seven-year-old, and adult subjects heard a complete set of 90 stimuli twice, in two different random orders, resulting in 180 responses per subject and 10 responses per transition type for each point on the continuum. The three- to four-year-old subjects heard a complete set of 90 stimuli only once, resulting in 5 responses per transition type for each point on the continuum for this group. Although this smaller number of presentations may have lead to noisier data than if 10 responses per transition type had been collected, it was only practical to test a smaller number of responses for this age group because of limitations on the children's attention span. Subsequent examination of the data showed that the results from the three- to four-year-olds were not qualitatively different from those of the other child subjects. Each randomization was split into blocks of 10 stimuli for presentation. The interstimulus interval for presentation to the adult subjects was 3 s, with an interblock interval of 10 s. Following Walley and Carrell (1983), the interstimulus interval was not fixed for presentation to the children. Instead, the presentation was paused briefly after every stimulus, allowing the children sufficient time to respond.[6] At the end of each block, the children were allowed to choose a small prize.

Here, as in the literature (see Nittrouer, 1992), only those listeners who responded correctly to 80% of the congruent continuum end points presented within the test proper were included in quantitative analysis. The purpose of this was to eliminate listeners whose responses were random or inconsistent.

III. ANALYSIS 1: ADULT VERSUS CHILD WEIGHTING OF TRANSITIONS

Each contrast engendered two sets of responses, one for each transition condition. The responses for the /saɪ/-/ʃaɪ/, /ta/-/da/ and /ti/-/di/ contrasts were normalized using a probit

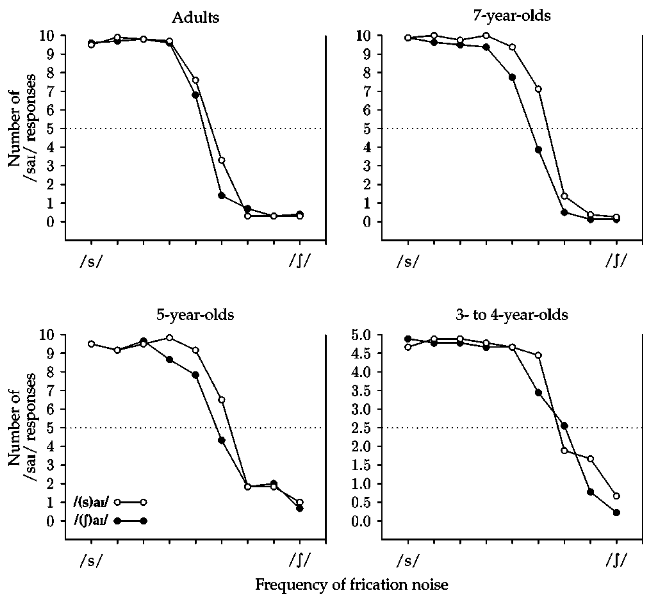

FIG. 3. Adults' and children's responses to /saɪ/-transition stimuli (open circles) and /ʃaɪ/-transition stimuli (closed circles). Responses are presented in terms of /saɪ/-responses as a function of frequency of frication noise ranging from most /s/-like (5800 Hz) to most /ʃ/-like (3100 Hz). The dotted lines indicate the 50% /saɪ/ response point. The *y*-axis range for the three- to four-year-olds is half that of the other subjects because this group heard half as many repetitions per point on the continuum.

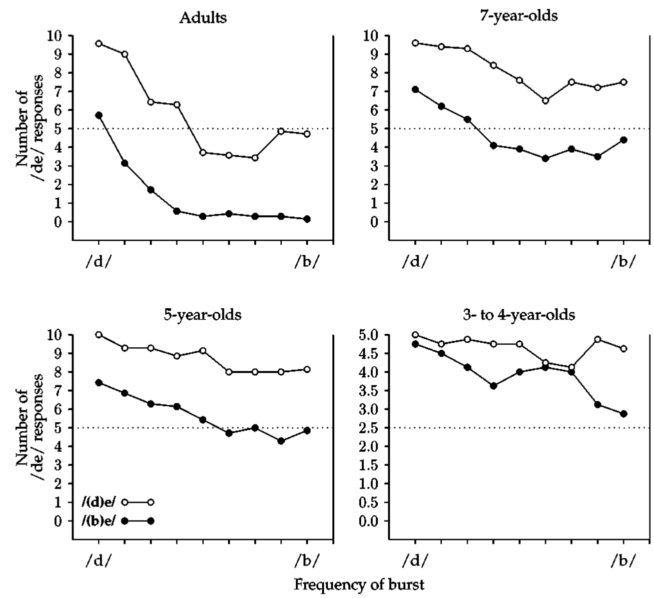

FIG. 4. Adults' and children's responses to /de/-transition stimuli (open circles) and /be/-transition stimuli (closed circles). Responses are presented in terms of /de/-responses as a function of frequency of complex stop burst ranging from most /d/-like to most /b/-like (see the text for details of frequency range). The dotted lines indicate the 50% /de/ response point. See Fig. 3 for more details.

transformation. This transform extracts rate-of-change information from data appropriately modeled with an S-shaped curve and yields estimates of the slope and the mean of the curve (Cohen and Cohen, 1983). The slope corresponds to the degree of categoriality of the responses and the mean corresponds to the point on the continuum at which the responses reach 50% (i.e., 50% /s/ or /ʃ/ responses). The degree of separation of the two response curves was calculated by taking the difference of the two means (e.g., the mean of the continuum with /s/-transitions and the mean of the continuum with /ʃ/-transitions). This gives a measure of the extent to which listeners' responses were influenced by the change in transitional information across continua.

As noted above, only five children met the testing criterion for the /de/-/be/ contrast. The responses to these contrasts were therefore not analyzed quantitatively.

A. Results

1. /saɪ/-/ʃaɪ/

The results for the /saɪ/-/ʃaɪ/ contrast are shown in Fig. 3. The listeners' responses are generally consistent with the results found previously by Nittrouer and colleagues and others for /s/-/ʃ/ contrasts: The children in the current study showed greater influence of transitional cues than adults. That is, children as a group showed a significantly greater separation of response curves than did adults in response to a change in transitional information [$F(1,31)=5.50$, $p=0.026$]. There was also a significant difference between adults and children for placement of category boundary along the frication continuum for the /saɪ/-transition response curves [$F(1,31)=10.53$, $p=0.003$] and for the /ʃaɪ/-transition response curves [$F(1,31)=6.03$, $p=0.020$]. Although Nittrouer reported differences in slope between adults and children (Nittrouer and Studdert-Kennedy, 1987; Nittrouer, 1992), which she has taken to reflect a difference between the two groups in reliance on frication noise (i.e., the cue that changes along the two continua), we did not find any significant difference between adults and children for the slope of either the /s/-transition or the /ʃ/-transition response curves. Nittrouer and colleagues also found significant differences between children of different ages for both separation and slope of response curves. However, no significant effect of age amongst the children was found for these two factors in this study.

2. /de/-/be/

Figure 4 shows the results for the /de/-/be/ contrast. Only five children met the testing criterion for this contrast. Despite this, the perceptual behavior of all of the children tested on this contrast appeared to be relatively nonrandom both along and across the continua. That is, there seemed to be a principled relationship between responses to consecutive points on each continuum, and differences between responses to points on each continuum were comparable at successive points. Therefore, for illustrative purposes, the responses of *all* subjects, including those that did not meet the testing criterion, are shown here.

Adults' response curves were both widely separated and slightly sloping, indicating that they relied on both transitional and nontransitional (stop burst) cues in making their responses. In contrast, children's response curves showed much less consistent separation and little identifiable slope;

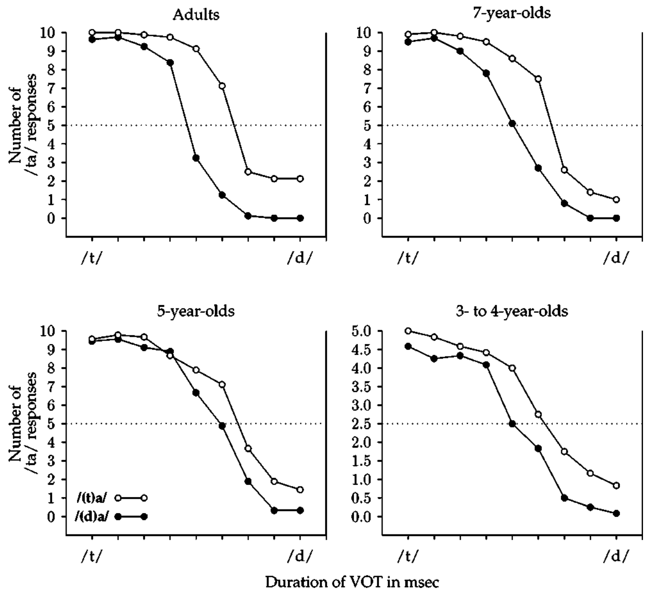

FIG. 5. Adults' and children's responses to /ta/-transition stimuli (open circles) and /da/-transition stimuli (closed circles). Responses are presented in terms of /ta/-responses as a function of VOT ranging from most /t/-like (40 ms) to most /d/-like (0 ms). The dotted lines indicate the 50% /ta/ response point. See Fig. 3 for more details.

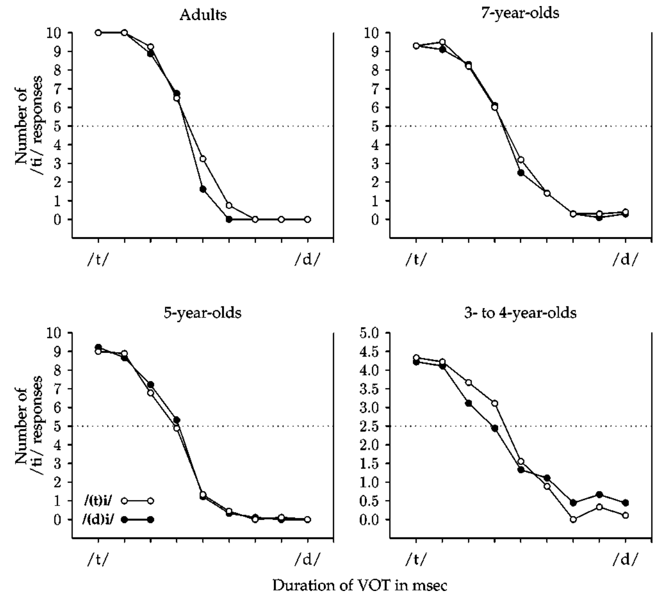

FIG. 6. Adults' and children's responses to /ti/-transition stimuli (open circles) and /di/-transition stimuli (closed circles). Responses are presented in terms of /ti/-responses as a function of VOT ranging from most /t/-like (40 ms) to most /d/-like (0 ms). The dotted lines indicate the 50% /ti/ response point. See Fig. 3 for more details.

in fact, many three- to four-year-olds appeared not to be able to consistently identify any of the /de/-/be/ stimuli as different at all. This, in addition to the fact that so few children met the testing criterion of 80% correct identification of congruent end points, suggests that children were unable to reliably make use of either the transitional or the nontransitional cues to identify this contrast.

B. /ta/-/da/ and /ti/-/di/

Figures 5 and 6 show the results for the /ta/-/da/ and /ti/-/di/ contrasts. The results appear to be qualitatively different from those seen for /saɪ/-/ʃaɪ/. Children do not appear to pay more attention to transitional information than adults in their identification of /ta/-/da/ or /ti/-/di/. For the /ti/-/di/ contrast, children give the same weight as adults to transitional cues. This is indicated by the fact that there is no significant difference between children as a group and adults in separation of response curves for this contrast. For the /ta/-/da/ contrast, children give less weight than adults to transitional cues. This is reflected in the fact that children as a group show a significantly smaller separation of response curves due to a change in transitional information than do adults [$F(1,37)=5.89$, $p=0.02$]. Children of different ages did not differ significantly from each other for this dimension for any of the three contrasts. There was no significant difference between adults and children for placement of category boundary along the VOT continuum for /ta/-transition response curves, /da/-transition response curves, /ti/-transition response curves, or /di/-transition response curves.

C. Discussion

As noted earlier, adults' and children's responses to the /saɪ/-/ʃaɪ/ contrast are consistent with previous /s/-/ʃ/ studies, as well as with the proposal that children give more weight to transitional cues than do adults. However, the findings that children and adults are equally influenced by transitions for /ti/-/di/, and that children show less influence of transitional cues than adults for /ta/-/da/ appear to contradict the developmental transitional bias hypothesis. In fact, the results from all three contrasts taken together suggest that, rather than being consistently more biased than adults toward transitions, the extent to which children make use of transitional information as compared to adults changes with segmental context.

However, there is an alternative explanation for these VOT results, which makes reference to the qualitative difference between transitions following /t/ and /d/ and those following /s/ and /ʃ/. For low or back vowels following /s/ and /ʃ/, formants differ in both direction and extent. That is, F2 transitions following /s/ are greater in extent than F2 transitions following /ʃ/ and F3 transitions following /s/ are relatively flat, while F3 transitions following /ʃ/ are rising. For vowels following /t/ and /d/, the formant transitions move in the same direction, but not to the same extent: Voiced transitions following /t/ are much less extensive than those following /d/, due to the presence of voiceless aspiration. This means that the voiced transition of a vowel following /t/ looks like a frequency-truncated version of a vowel following /d/ (for example, the frequency of F2 at the voiced onset of /ta/ is lower than the frequency of F2 at the voiced onset of /da/, see Fig. 1). One could therefore interpret the results of the two current VOT experiments as meaning that children are so much more sensitive to transitional information than adults that they are prepared to accept even "frequency-truncated" /d/-transitions (i.e., /t/-transitions) as indicating a /d/, and thus need only a short silence duration to persuade them that what they have heard is indeed /d/. This behavior would result in a smaller separation of response curves for children than for adults. This proposal was put forward by Nittrouer (1992) to explain results of a study of ‘‘say-stay’’ perception, and is consistent with the claim that children are more perceptually sensitive to transitions than adults.

There is, however, evidence against this interpretation. First, as noted earlier, a truncation explanation requires a smaller separation of response curves for children than for adults. While this is true for the /ta/-/da/ contrast, for the /ti/-/di/ contrast there was no significant difference between adults and children for separation of the /ti/ and /di/ response curves. Also, in order for a truncation explanation to be correct, the category boundaries for all of the subjects should be the same for the unambiguous, ‘‘untruncated’’ /d/-transition stimuli, but there should be a difference between children and adults for the ambiguous, ‘‘truncated’’ /t/-transition stimuli: If children require less silence to hear these as voiced, they should place their boundaries closer to the 0 ms end of the VOT continuum than do the adults (Nittrouer, 1992). However, as noted earlier, although the difference in separation of response curves between adults and children was significant for /ta/-/da/, suggesting that the (nonsignificant) shift for each type of boundary was comparable in magnitude, there was no significant effect of age on the placement of either the /ta/-transition response curves or the /da/-transition response curves. There was also no significant difference between adults and children for the placement of either the /ti/-transition response curves or the /di/-transition response curves.

Further evidence against this interpretation of the current VOT results comes from studies of fricative contrast perception. The results for the /ta/-/da/ contrast showed greater separation of response curves for adults as compared to children. If this pattern is due to the fact that /t/-transitions are simply ‘‘frequency-truncated’’ versions of /d/-transitions then the same pattern of results should be seen for all contrasts in which the transitional information in one syllable is a ‘‘frequency-truncated’’ version of the other, no matter what the segmental context. However, this is not what is found. In the /su/-/ʃu/ contrasts that have been shown to engender cue weighting differences in children and adults, F2 following /s/ is a ‘‘frequency-truncated’’ version of F2 following /ʃ/, in the sense that the onset frequency of F2 following /s/ is lower than the onset frequency of F2 following /ʃ/. Therefore, if F3 is neutralized in both of the syllables, the resulting /su/-/ʃu/ contrast will be cued by the same type of relationship between transitions as found in the /ta/-/da/ contrast above. According to the truncation view, children should accept more of these neutral-F3-/su/ stimuli as ‘‘truncated’’ versions of /ʃu/. They should therefore show a smaller separation of response curves than adults in perception of this contrast.

Nittrouer and Miller (1997) have tested children’s and adults’ cue weighting of a /su/-/ʃu/ contrast in which F3 was neutralized. However, the results of that study showed that, unlike for the /ta/-/da/ contrast, young children (4 years) showed *greater* separation of response curves than older children (7 years) and both groups of children showed *greater* separation than adults. This is in keeping with studies of /s/-/ʃ/ contrasts in which F3 was not neutral. It would therefore appear that cue weighting differences between adults and children for /ta/-/da/ are not due to /t/-transitions being ‘‘frequency-truncated’’ versions of /d/-transitions.

TABLE I. Results of logistic regression for /ta/-/da/ and /ti/-/di/.

		Children[a]	Adults[a]
/ta/-/da/	VOT	0.3944	0.2878
	Transitions	3.4064	14.1490
/ti/-/di/	VOT	0.4123	0.1409
	Transitions	Not significant	1.6334, $p=0.0282$

[a]Regression is reported in terms of exp(B), $d.f.=1$. Values are significant at $p<0.001$ unless otherwise indicated.

To summarize, the pattern of responses seen for /ta/-/da/ and /ti/-/di/ cannot be explained by the developmental transitional bias hypothesis, that is, that children are more biased toward transitional information than are adults. However, it is possible that both sets of results can be explained by the acoustic transitional bias hypothesis, that is, that children are more biased toward transitional than toward nontransitional cues. Analysis 2 will examine this possibility.

The results for the /de/-/be/ contrast should be touched on here. The fact that very few children were able to meet the testing criterion for this contrast suggests a possible inability in the children to make use of either of the available cues (stop burst or vowel formant transitions). This in itself contradicts both the developmental and acoustic transitional bias hypotheses which state that children should always be more biased toward—and thus presumably able to use—transitional cues. These results are consistent with the view that the acoustic distinctiveness or salience of cues plays a role in adult–child cue weighting differences (e.g., Sussman, 2001).

IV. ANALYSIS 2: RELATIVE WEIGHT GIVEN BY CHILDREN TO NONTRANSITIONAL CUES

A. Results and Discussion

The responses to the two VOT contrasts, /ta/-/da/ and /ti/-/di/, were subjected to logistic regression. This type of analysis is used when the dependent variable (the listeners’ response) has two possible values (e.g., the responses for each stimulus in the /ta/-/da/ contrast could be either /ta/ or /da/). It is used here to determine the amount of variance in the responses that can be explained by each of the two acoustic cues varied in each contrast, and to rank the importance of the two cues for each group of listeners: The higher the value of exp(B), the higher the relative importance of the cue.

Results and discussion. The results of the logistic regression analysis can be found in Table I. For the /ta/-/da/ contrast, transitions play a greater role than the nontransitional VOT cue for both adults and children. Therefore, although the responses to the /ta/-/da/ contrast cannot be explained by the claim that children always give more weight to transitional cues than do adults, they *can* be explained by the claim that children start out more biased toward transitions than other cues, and gradually give less or more weight (in this case, more) to these cues as the children develop perceptually. Interestingly, it was suggested earlier that the acoustic transitional bias hypothesis might only be required in order to account for perception of contrasts in which nontransitional cues are known to be uninformative to adults. This is not, however, the case for /ta/-/da/: The nontransitional VOT cue in this contrast is known to be very informative to adult listeners in the perception of voicing in CV syllables (Lisker and Abramson, 1970).

For the /ti/-/di/ contrast, on the other hand, transitions play a smaller role in perception than does the nontransitional VOT cue for children. In fact, the influence of transitional information on children's responses is so small that it is nonsignificant (it remains significant for adult listeners). This means that the responses to the /ti/-/di/ contrast cannot be explained by either the original claim of the developmental transitional bias hypothesis that children give more weight to transitional cues than do adults, or the claim of the acoustic transitional bias hypothesis that children start out more biased toward transitions than other cues.

These results show that the acoustic transitional bias hypothesis does not appear to account for all situations in which children give less or the same weight to transitional cues as adults. This in turn means that the general premise of the transitional bias theory—that children are more biased (in some way) toward vowel-formant transitional cues—cannot account for all adult–child differences in acoustic cue weighting.

V. GENERAL DISCUSSION AND CONCLUSIONS

The aims of this study were (a) to test the two nonexclusive versions of the transitional bias theory, and (b) to determine to what extent cue weighting patterns displayed by adults and children are affected by segmental context.

The Developmental Weighting Shift (DWS) theory states that differences between adults and children in the way in which they process speech—children being more "global" and adults being more "analytical" in their processing—lead to differences in perceptual cue weighting strategies between the two groups. While the results of the current study neither support nor refute this claim, the perceptual behavior observed in this study is problematic for existing interpretations of the *way* in which changes in processing might impact on speech perception. To date, Nittrouer and colleagues' tests of the DWS theory have been based on the assertion that the acoustic correlate of more global, or "child-like" speech perception is vowel formant transitions (the transitional bias hypothesis). However, the current study has shown that if children do use a more global mode of speech perception, it must be triggered by something other than, or in addition to, formant transitions. While children do appear to be more "transitional" for some contrasts, they are not consistently biased toward transitions. The first version of the transitional bias hypothesis proposes that children should weight vowel-formant transitions more heavily than should adults. The results of the current study show that, as found by Nittrouer and colleagues, children do weight vowel-onset formant transitions more heavily than do adults for fricative contrasts such as /saɪ/-/ʃaɪ/. However, children were found to give less weight than adults to transitions for the /ta/-/da/ contrast, and the same weight as adults for the /ti/-/di/ contrast. Additionally, children, but not adults, appear to have difficulty making use of transitional cues at all for the contrast /de/-/be/. The second version of the transitional bias hypothesis proposes that, where children are found not to weight transitions more than do adults, children should weight transitional cues more heavily than they weight other acoustic cues to the same contrast. However, it appears that while this may be the case for some contrasts, it is not true for all of them. While children do indeed weight vowel-formant transitions more heavily than nontransitional acoustic cues for /ta/-/da/, they weight transitions *less* heavily than nontransitional cues for /ti/-/di/. Therefore, contrary to Nittrouer's transitional bias theory, children are not always biased toward transitions, either in comparison to adults or in relation to the weight that they give to nontransitional cues.

The results of previous studies showed different types of adult–child cue weighting differences, some consistent with Nittrouer's transitional bias theory (e.g., Krause, 1982; Wardrip-Fruin and Peach, 1984), and others contradicting the theory (e.g., Howell *et al.*, 1992; Sussman, 2001). It was not clear, however, whether these contradictory results were due to variation in the segmental context being tested, or to differences in the methods used for testing. 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. Given this variation in perceptual behavior, it is not immediately clear what the underlying cause of adult–child cue weighting differences might be. The results from the /de/-/be/ contrast suggest a possible explanation. For this contrast, children appeared to have difficulty making use of the available acoustic cues. As noted earlier, poor auditory skills in children as compared to adults has been proposed as an alternative account for adult–child differences in cue weighting (e.g., Sussman, 2001). In particular, children's incomplete auditory development is said to cause them to have difficulty in making use of less acoustically distinctive cues. It does appear from the results of the /de/-/be/ contrast that there may be certain weaker cues that young children cannot use in some circumstances. However, a recent study has shown that this difference in ability to make use of less physically distinct cues cannot explain all adult–child differences in acoustic cue weighting (see Mayo and Turk, submitted). Therefore, while this and earlier studies show clear differences between adults and children in the way in which the two groups make use of acoustic cues, we are left without a satisfactory explanation for these developmental differences.

ACKNOWLEDGMENTS

This work was supported by a grant from the Wellcome Trust. The authors would like to thank Jocelynne Watson for comments on an earlier version of this paper.

APPENDIX A: VALUES FOR SYNTHETIC CV STIMULI

Vowel-formant onset and target values for synthetic /saɪ/ stimuli.

Stimulus No.	F1 onset	F1 target (/a/)	F2 onset	F2 target (/a/)	F3 onset	F3 target (/a/)
1	465	788	1444	1170	2626	2817
2	428	781	1423	1195	2620	2809
3	433	751	1464	1197	2535	2777
4	397	771	1433	1204	2518	2771
5	376	764	1408	1189	2585	2779

Vowel-formant onset and target values for synthetic /ʃaɪ/ stimuli.

Stimulus No.	F1 onset	F1 target (/a/)	F2 onset	F2 target (/a/)	F3 onset	F3 target (/a/)
1	469	796	1566	1183	2427	2799
2	489	734	1510	1172	2409	2795
3	449	761	1621	1183	2402	2775
4	411	723	1597	1185	2446	2770
5	358	753	1578	1159	2319	2750

Vowel-formant onset and target values for synthetic /de/ stimuli.

Stimulus No.	F1 onset	F1 target	F2 onset	F2 target	F3 onset	F3 target
1	219	425	1806	2096	2477	2554
2	231	436	1810	2116	2491	2542
3	221	416	1821	2093	2419	2536
4	221	443	1814	2109	2401	2544
5	206	426	1796	2112	2442	2520

Vowel-formant onset and target values for synthetic /be/ stimuli.

Stimulus No.	F1 onset	F1 target	F2 onset	F2 target	F3 onset	F3 target
1	249	420	1694	2100	2259	2535
2	247	429	1771	2114	2293	2540
3	271	427	1631	2094	2213	2524
4	247	429	1680	2117	2254	2540
5	273	429	1693	2104	2214	2553

Vowel-formant onset and target values for synthetic /ta/ stimuli.

Stimulus No.	F1 onset	F1 target	F2 onset	F2 target	F3 onset	F3 target
1	528	709	1530	1433	2560	2685
2	526	715	1524	1416	2536	2662
3	555	702	1555	1416	2513	2638
4	531	707	1541	1443	2577	2697
5	544	716	1528	1435	2564	2713

Vowel-formant onset and target values for synthetic /da/ stimuli.

Stimulus No.	F1 onset	F1 target	F2 onset	F2 target	F3 onset	F3 target
1	261	716	1631	1423	2498	2653
2	291	705	1629	1443	2490	2675
3	271	708	1667	1434	2442	2662
4	243	721	1643	1461	2496	2630
5	238	712	1639	1421	2433	2631

Vowel-formant onset and target values for synthetic /ti/ stimuli.

Stimulus No.	F1 onset	F1 target	F2 onset	F2 target	F3 onset	F3 target
1	324	324	1948	2192	2571	2831
2	306	306	1861	2171	2584	2752
3	316	316	1943	2184	2606	2832
4	300	300	1918	2153	2623	2831
5	310	310	1951	2197	2610	2856

Vowel-formant onset and target values for synthetic /di/ stimuli.

Stimulus No.	F1 onset	F1 target	F2 onset	F2 target	F3 onset	F3 target
1	225	305	1857	2202	2560	2799
2	210	302	1904	2206	2574	2837
3	199	305	1923	2189	2573	2825
4	238	305	1870	2149	2534	2812
5	235	314	1911	2186	2605	2815

APPENDIX B: PICTURES USED TO ELICIT CHILD RESPONSES

CV syllable	Picture
/saɪ/	A boy called ‘‘Si’’
/ʃaɪ/	A shy boy hiding behind a tree
/de/	A street scene with the sun coming up (‘‘day’’)
/be/	A typical Scottish bay
/ta/	A girl receiving a present (‘‘ta’’ is British English slang for ‘‘thank you’’)
/da/	A father (‘‘da’’ short for ‘‘dada’’)
/ti/	A teapot and teacup (‘‘tea’’)
/di/	A girl called ‘‘Dee’’

[1]The vowel in the /ta/-/da/ contrast, as well as the initial vowel in the diphthong in the /saɪ/-/ʃaɪ/ contrast is the Scottish low front vowel, best approximated by the IPA symbol /a/.

[2]The rejection rate for adults for this contrast may reflect the quality of the synthetic speech stimuli used for this contrast. These stimuli were judged by the two experimenters to appropriately model natural tokens of /saɪ/ and /ʃaɪ/; however, it is possible that these stimuli were more difficult to identify than others used in this and previous studies. There are, however, no data on adult rejection rate available from Nittrouer and colleagues' studies of /s-vowel/-/ʃ-vowel/ perception with which to compare these results.

[3]The data for the /de/-/be/, /ta/-/da/ and /ti/-/di/ contrasts are also featured in Mayo and Turk (submitted), where they are analyzed for a different purpose.

[4]The duration of the vowel formant transitions for all of the synthetic stimuli reflects both the steep and more gradual slope components of a natural transition.

[5]The frication excited bypass path in Sensyn's Klatt synthesiser (Sensimetrics Org.); no frequency value is given for this parameter.

[6]In general, pauses were minimal—sufficiently long for the child subject to respond aurally and by placing a marker on the correct picture. Longer pauses were taken as indicating that the child was tiring, and in this case testing was discontinued until a later time (usually only the case for the three- to four-year-old subjects).

Cohen, J., and Cohen, P. (**1983**). *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences*, 2nd ed. (LEA, Hillsdale, NJ).

Dollaghan, C. A. (**1994**). ‘‘Children's phonological neighbourhoods—half empty or half full?'' J. Child Lang **21**, 257–271.

Dorman, M. F., Studdert-Kennedy, M., and Raphael, L. J. (**1977**). ‘‘Stop-consonant recognition: Release bursts and formant transitions as functionally equivalent, context-dependent cues,'' Percept. Psychophys. **22**, 109–122.

Dunn, L. M., Dunn, L. M., Whetton, C., and Burley, J. (**1997**). *British Picture Vocabulary Test*, 2nd ed. (NFER-NELSON, Berkshire, UK).

Eisenberg, L. S., Shannon, R. V., Schaefer Martinez, A., Wygonski, J., and Boothroyd, A. (**2000**). ''Speech recognition with reduced spectral cues as a function of age,'' J. Acoust. Soc. Am. **107**, 2704–2710.

Elliott, L. L., and Busse, L. A. (**1984**). ''Syllable identification by children and adults for two task conditions,'' J. Acoust. Soc. Am. **77**, 1258–1260.

Elliott, L. L., Longinotti, C., Meyer, D., Raz, I., and Zucker, K. (**1981**). ''Developmental differences in identifying and discriminating CV syllables,'' J. Acoust. Soc. Am. **78**, 669–677.

Gerken, L., Murphy, W. D., and Aslin, R. N. (**1995**). ''3-year-olds and 4-year-olds perceptual confusions for spoken words,'' Percept. Psychophys. **57**, 475–486.

Harris, K. S. (**1958**). ''Cues for the discrimination of American English fricatives in spoken syllables,'' Lang Speech **1**, 1–7.

Howell, P., Rosen, S., Lang, H., and Sackin, S. (**1992**). ''The role of F1 transitions in the perception of voicing in initial plosives,'' in *Speech, Hearing and Language: Work in Progress* (University College, London).

Jusczyk, P. W., and Derrah, C. (**1987**). ''Representation of speech sounds by young infants,'' Dev. Psychol. **23**, 648–654.

Klatt, D. (**1980**). ''Software for a cascade/parallel formant synthesizer,'' J. Acoust. Soc. Am. **67**, 971–995.

Krause, S. E. (**1982**). ''Vowel duration as a perceptual cue to postvocalic consonant voicing in young children and adults,'' J. Acoust. Soc. Am. **71**, 990–995.

Lacerda, F. (**1992**). ''Young infants' discrimination of confusable speech signals,'' in *The Auditory Processing of Speech: From Sounds to Words*, edited by M. E. H. Schouten (Mouton de Gruyter, Berlin), pp. 229–238.

Lisker, L., and Abramson, A. (**1970**). ''The voicing dimension; some experiments in comparative phonetics,'' in *Proceedings of the Sixth International Congress of Phonetic Sciences, Prague, 1967* (Academia, Prague), pp. 563–567.

Malech, S. R., and Ohde, R. N. (**2003**). ''Cue weighting of static and dynamic vowel properties in children versus adults,'' J. Acoust. Soc. Am. **113**, 2257(A).

Mayo, C., Scobbie, J. M., Hewlett, N., and Waters, D. (**2003**). ''The influence of phonemic awareness development on acoustic cue weighting in children's speech perception,'' J. Speech Lang. Hear. Res. **46**, 1184–1196.

Mayo, C., and Turk, A. (submitted). ''The influence of cue magnitude on acoustic cue weighting in children's and adults' speech perception.''

Menn, L. (**1971**). ''Phonotactic rules in beginning speech,'' Lingua **26**, 225–241.

Morrongiello, B. A., Robson, R. C., Best, C. T., and Clifton, R. K. (**1984**). ''Trading relations in the perception of speech by five-year-old children,'' J. Exp. Child Psychol. **37**, 231–250.

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

Nittrouer, S. (**1993**). ''The emergence of mature gestural patterns is not uniform—Evidence from an acoustic study,'' J. Speech Hear. Res. **36**, 959–972.

Nittrouer, S. (**1996**). ''The relation between speech perception and phonemic awareness: Evidence from low-SES children and children with chronic OM,'' J. Speech Hear. Res. **39**, 1059–1070.

Nittrouer, S. (**2002**). ''Learning to perceive speech: How fricative perception changes, and how it stays the same,'' J. Acoust. Soc. Am. **112**, 711–719.

Nittrouer, S., and Miller, M. E. (**1997**). ''Predicting developmental shifts in perceptual weighting schemes,'' J. Acoust. Soc. Am. **101**, 2253–2266.

Nittrouer, S., Miller, M. E., Crowther, C. S., and Manhart, M. J. (**2000**). ''The effect of segmental order on fricative labeling by children and adults,'' Percept. Psychophys. **62**, 266–284.

Nittrouer, S., and Studdert-Kennedy, M. (**1987**). ''The role of coarticulatory effects in the perception of fricatives by children and adults,'' J. Speech Hear. Res. **30**, 319–329.

Ohde, R. N., and Haley, K. L. (**1997**). ''Stop-consonant and vowel perception in 3- and 4-year-old children,'' J. Acoust. Soc. Am. **102**, 3711–3722.

Ohde, R. N., Haley, K. L., and McMahon, C. W. (**1996**). ''A developmental study of vowel perception from brief synthetic consonant-vowel syllables,'' J. Acoust. Soc. Am. **100**, 3813–3824.

Parnell, M. M., and Amerman, J. D. (**1978**). ''Maturational influences on perception of coarticulatory effects,'' J. Speech Hear. Res. **21**, 682–701.

Schonell, F., and Goodacre, E. (**1971**). *The Psychology and Teaching of Reading* (Oliver and Boyd, London).

Sensimetrics Org. (n.d.). ''SenSyn: Speech Synthesizer Package,'' Cambridge, MA.

Simon, C., and Fourcin, A. J. (**1978**). ''Cross-language study of speech pattern learning,'' J. Acoust. Soc. Am. **63**, 925–935.

Studdert-Kennedy, M. (**1987**). ''The phoneme as a perceptuomotor structure,'' in *Language Perception and Production: Relationships Between Listening, Speaking, Reading and Writing*, edited by A. Allport, D. G. MacKay, W. Prinz, and E. Scheerer (Academic, London), pp. 67–84.

Sussman, J. E. (**1993**). ''Auditory processing in children's speech perception: Results of selective adaptation and discrimination tasks,'' J. Speech Hear. Res. **36**, 380–395.

Sussman, J. E. (**2001**). ''Vowel perception by adults and children with normal language and specific language impairment: Based on steady states or transitions,'' J. Acoust. Soc. Am. **109**, 1173–1180.

Walley, A. C., and Carrell, T. D. (**1983**). ''Onset spectra and formant transitions in the adult's and child's perception of place of articulation in stop consonants,'' J. Acoust. Soc. Am. **73**, 1011–1022.

Wardrip-Fruin, C. (**1982**). ''On the status of temporal cues to phonetic categories: Preceding vowel duration as a cue to voicing in final stop consonants,'' J. Acoust. Soc. Am. **71**, 187–195.

Wardrip-Fruin, C. (**1985**). ''The effect of signal degredation on the status of cues to voicing in utterance-final stop consonants,'' J. Acoust. Soc. Am. **77**, 1907–1912.

Wardrip-Fruin, C., and Peach, S. (**1984**). ''Developmental aspects of the perception of acoustic cues in determining the voicing feature of final stop consonants,'' Lang Speech **27**, 367–379.

Watson, J. (**1997**). ''Sibilant-vowel coarticulation in the perception of speech by children with phonological disorder,'' Ph.D. thesis, Queen Margaret College, Edinburgh.

Whalen, D. H. (**1991**). ''Perception of the English /s/-/ʃ/ distinction relies on fricative noises and transitions, not brief spectral slices,'' J. Acoust. Soc. Am. **90**, 1776–1784.