The influence of spectral distinctiveness on acoustic cue weighting in children's and adults' speech perception

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 1 June 2005; accepted 1 June 2005)

Children and adults appear to weight some acoustic cues differently in perceiving certain speech contrasts. One possible explanation for this difference is that children and adults make use of different strategies in the way that they process speech. An alternative explanation is that adult-child cue weighting differences are due to more general sensory (auditory) processing differences between the two groups. It has been proposed that children may be less able to deal with incomplete or insufficient acoustic information than are adults, and thus may require cues that are longer, louder, or more spectrally distinct to identify or discriminate between auditory stimuli. The current study tested this hypothesis by examining adults' and 3- to 7-year-old children's cue weighting for contrasts in which vowel-onset formant transitions varied from spectrally distinct (/no/-/mo/, /do/-/bo/, and /ta/-/da/) to spectrally similar (/ni/-/mi/, /de/-/be/, and /ti/-/di/). Spectrally distinct cues were more likely to yield different consonantal responses than were spectrally similar cues, for all listeners. Furthermore, as predicted by a sensory hypothesis, children were less likely to give different consonantal responses to stimuli distinguished by spectrally similar transitional cues than were adults. However, this pattern of behavior did not hold for all contrasts. Implications for theories of adult-child cue weighting differences are discussed. © *2005 Acoustical Society of America.* [DOI: 10.1121/1.1979451]

PACS number(s): 43.71.Ft [DOS]   Pages: 1730–1741

I. INTRODUCTION

A number of speech perception studies have shown that adults and children weight acoustic cues differently in identifying certain contrasts (e.g., Krause, 1982; Lacerda, 1992; Mayo *et al.*, 2003; Morrongiello *et al.*, 1984; Nittrouer and Studdert-Kennedy, 1987; Nittrouer and Miller, 1997; Ohde and Haley, 1997; Parnell and Amerman, 1978; Sussman, 2001; Watson, 1997; Wardrip-Fruin and Peach, 1984). For example, when identifying /s-vowel/-/ʃ-vowel/ contrasts on the basis of frequency of frication noise and vowel-onset formant transitions, children have been found to weight the transitions more heavily than do adults (e.g., Nittrouer and Studdert-Kennedy, 1987). Note that it is not simply that children and adults have different primary cues for these contrasts, as would be the case if children primarily used transitional cues and adults primarily used frication noise cues. Instead, Nittrouer and Studdert-Kennedy's (1987) study found that *both* adults and children gave more weight to frication noise than to transitional information. The difference between adults and children was in the *degree* of attention, or weight, given to different types of acoustic information available in the speech stream.

One explanation for these differences in cue weighting behavior is that adults and children make use of different strategies in the way that they process speech. One such hypothesis is Nittrouer and colleagues' Developmental Weighting Shift (DWS) theory, which proposes that children process speech in terms of large units (the size of a syllable or monosyllabic word) while adults process in terms of smaller units. This processing difference then impacts on speech perception in terms of the attention that listeners give to acoustic cues. Children should therefore attend more than adults to cues that "delimit signal portions corresponding to syllables" (Nittrouer *et al.*, 2000, p. 268).

It has been suggested that cues which could be considered to be acoustic correlates of more global, syllable-based speech perception could be something like spectrally dynamic vowel-onset/-offset formant transitions (e.g., Nittrouer *et al.*, 2000). Thus, as found by Nittrouer and Studdert-Kennedy (1987) for fricative contrasts, children should give more weight than should adults to formant transitions, while adults' attention should be focused on more spectrally static cues (see also, e.g., Nittrouer, 1992, 1996; Nittrouer and Miller, 1997). However, some studies have found children to weight transitions *less* than do adults for certain contrasts (Howell *et al.*, 1992; Malech and Ohde, 2003; Mayo and Turk, 2004; Simon and Fourcin, 1978; Sussman, 2001). For example, Sussman (2001) examined 4-year-old children's and adults' (mean age 21 years, 6 months) weighting of two cues to vowel identification in CVC syllables: (*i*) vowel-onset and -offset formant transitions and (*ii*) vowel target formant frequencies (so-called vowel "steady-state" cues). The results of the study showed that although both adults and children were able to identify the vowels when only the transitional information was given to them, the children were not as successful as adults in doing so. Additionally, in identifying stimuli for which transitional information and static target frequency information were both available, but were in

---

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

conflict, the children made more use of the target information than did the adults.

Sussman (2001) proposes that these, and all, adult-child cue weighting differences are due to general sensory processing differences between adults and children rather than to strategy differences between the two groups (see also, e.g., Eisenberg *et al.*, 2000; Elliott, 1979; Elliott *et al.*, 1979, 1981). Under this explanation, the maturation of some aspect(s) of the central auditory system cause(s) adult-child speech perception differences as well as causing normal-hearing children to show poorer *nonspeech* perception abilities than normal-hearing adults (e.g., Allen and Wightman, 1994; Berg and Boswell, 2000; Elliott and Katz, 1980; Jensen and Neff, 1993; Maxon and Hochberg, 1982; Schneider and Trehub, 1992; Trehub *et al.*, 1988, 1995; Werner *et al.*, 1992).[1] Proponents of this theory suggest that "as auditory perceptual abilities mature over the first 10–12 years, central pattern recognition for speech may be much less robust to sensory distortion [than] that observed for adults" (Eisenberg *et al.*, 2000, p. 2705). In other words, children may be unable to cope, perceptually, with incomplete or insufficient acoustic information, and may thus need greater amounts of, or more distinct, acoustic cues than adults to identify or discriminate between auditory stimuli.

Sussman (2001) summarizes a number of these general acoustic theories in a proposal that states that children have a perceptual need for "louder or longer duration speech cues…or greater amounts of spectral information" (p. 1173). The acoustic speech cues used in Sussman's (2001) study described above were characterized in terms of this hypothetical need: The vowel-onset/-offset cues were described as shorter in duration, with "less overall energy" (p. 1174), while the steady-state cues were characterized as longer in duration and "more powerful" (p. 1174). Therefore, the fact that the children in the study were found to give more perceptual weight than were the adults to the vowel target cues is taken as evidence that "children rely more on the longer, louder or more acoustically salient cue" (p. 1179). This, in turn, is taken as support for a sensory-based, rather than a strategy-based, explanation for cue weighting behavior.

However, there are some aspects of Sussman's (2001) study that make it difficult to evaluate this claim. First, Sussman's study made use of a different paradigm from that used by Nittrouer and colleagues in the studies that formed the basis of the DWS (e.g., Nittrouer and Studdert-Kennedy, 1987; Nittrouer, 1996; Nittrouer and Miller, 1997). This makes it impossible to make a direct comparison between Sussman's study and those of Nittrouer and colleagues. Second, the primary evidence for the DWS theory comes from studies of perception of fricative contrasts (particularly /s-vowel/-/ʃ-vowel/ contrasts) while Sussman's study examined adults' and children's identification of vowel contrasts. This too makes comparison between the studies difficult: While it is possible that salience may play a role in adult-child cue weighting differences, it may (as in fact noted by Sussman 2001) simply be the case that children use different strategies for perceiving different segmental contrasts, relying more heavily on vowel formant transitions for the identification of fricative contrasts, and more heavily on vowel target cues for the identification of vowels.

Finally, there is a significant problem with the assumption that the children in Sussman's (2001) study gave more weight to the vowel target cues than to the transitional cues because of the greater length and loudness of the target cues, rather than because of their cue type (i.e., vowel *targets* versus vowel *transitions*). First, the overall dB values for the two cues used in the study were not reported, making it impossible to evaluate the actual influence of the loudness of the cues. Additionally, the possible effect of formant frequency change in the transitions (as compared to the static vowel targets) is not taken into account. Such a frequency change could have had a subtle influence on the transitional cues' loudness (Fletcher and Munson, 1933) and could also impact for reasons other than loudness on the transitions' salience. Most critically, however, Sussman (2001) did not systematically manipulate cue salience in a way that can be readily interpreted. The study did not, for example, compare children's and adults' cue weighting patterns for long, loud vowel targets with their weighting patterns for short, soft vowel targets (paired with identical transitions in both cases). It is therefore unclear whether it was in fact solely the "loudness" of the target cue that influenced the weight given to it by children.

The aim of the current study, therefore, was to undertake a more systematic test of a general sensory explanation of adult-child cue weighting differences by addressing these issues. First, the study made use of the same methodology as that used by Nittrouer and colleagues, allowing direct comparisons between the current study and those that underpin the DWS theory. In addition, a perceptual paradigm was adopted which allowed both segmental contrast and cue type to be kept constant, and only cue salience, as characterized by loudness, length, or degree of spectral information, to be varied systematically.

The paradigm that allowed for the systematic manipulation of salience was modeled on Nittrouer and Studdert-Kennedy's original cue weighting research (Nittrouer and Studdert-Kennedy, 1987). That study examined adults' and children's weighting of vowel-onset formant transitions (as compared to frequency of frication noise) for two fricative contrasts: (i) /su/-/ʃu/ ("sue"-"shoe") and (ii) /si/-/ʃi/ ("sea"-"she"). The key difference between these two contrasts, with regard to the current study, is in the *spectral informativeness* of the vowel-onset formant transitions—one of the acoustic features that Sussman (2001) suggested should affect the perceived salience of a cue for children. Specifically, the onset frequency, extent, direction, and duration of the vowel-onset formant transitions in /su/ differ substantially from the onset frequency, extent, direction, and duration of the vowel-onset formant transitions in /ʃu/. This makes formant transitions a particularly informative cue to fricative identify for the /su/-/ʃu/ contrast. For the /si/-/ʃi/ contrast, on the other hand, there is much less difference in onset frequency, extent, direction, and duration of the vowel formant transitions across the two syllables, making transitions much less spectrally informative in this contrast. The results of Nittrouer and

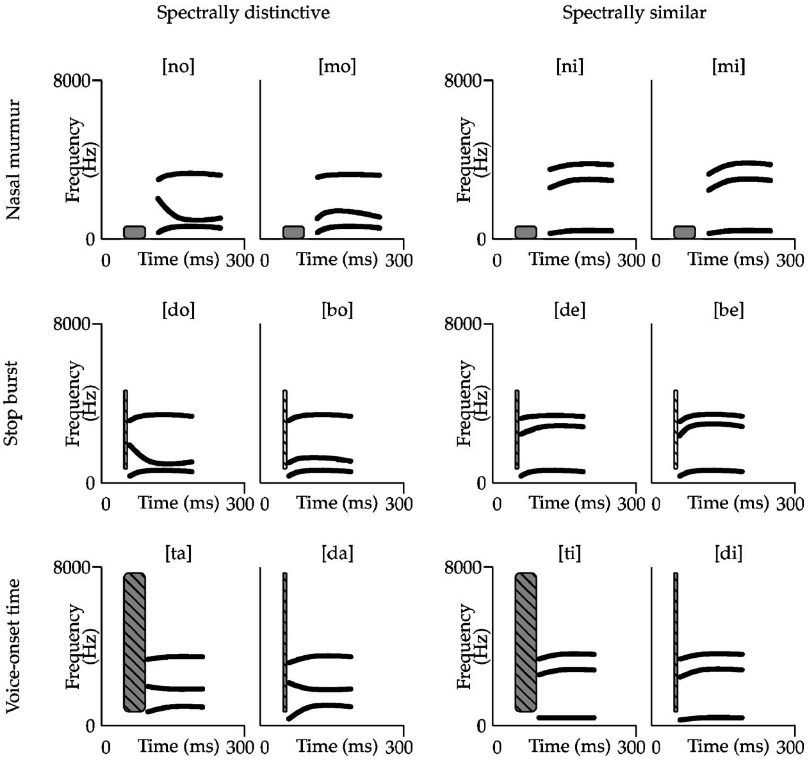

FIG. 1. Stylized spectrograms of prototypical tokens of the contrasts used in this study. The boxes represent: nasal murmur (/no/-/mo/, /ni/-/mi/), stop burst (/do/-/bo/, /de/-/be/), and burst+aspiration (/ta/-/da/, /ti/-/di/). The black lines represent vowel formants (F1, F2, F3).

Studdert-Kennedy's (1987) study showed that despite this difference in the spectral informativeness of the vowel formant transitions across the two contrasts, children gave more weight than did adults to the transitional cue for *both* /su/-/ʃu/ and /si/-/ʃi/.

It should be noted that the results of Nittrouer and Studdert-Kennedy (1987) do appear to favor a strategy-based explanation of adult-child cue weighting differences: The difference in spectral informativeness between the /su/-/ʃu/ contrast and the /si/-/ʃi/ contrast does not appear to have influenced children's cue weighting patterns. However, it remains possible that evidence for the influence of cue salience on adult-child differences in cue weighting in speech perception may emerge if a number of different segmental contrasts are examined. The current study therefore examined adults' and children's cue weighting for a range of consonantal contrasts.

A. The current study

The current study investigated adults' and children's (ages 3–4 years, 5 years, and 7 years) weighting of dynamic vowel-onset formant transition cues (compared to static steady-state cues) to one of three consonant contrasts: (*i*) the nasal contrast /n/-/m/, (*ii*) the stop burst contrast /d/-/b/, and (*iii*) the voice onset time (VOT) contrast /t/-/d/.

The focus in this study was on possible changes in listeners' transitional cue weighting patterns due to a change in degree of spectral information *across contrasts*. This is in contrast to Sussman's (2001) study, in which the focus was on differences in listeners' cue weighting patterns for two cues that appeared *within the same contrast*. This means that we did not specifically examine adults' or children's weighting of the steady-state, nontransitional cues to each contrast, despite the fact that, for the identification of some contrasts, steady-state cues may be more important to both adult and child listeners than transitional cues. However, it is only listeners' behavior in response to the change in spectral information in the transitional cues that allows us to differentiate between the DWS theory and a more general auditory theory.

The study followed the methodology of Nittrouer and Studdert-Kennedy's (1987) /su/-/ʃu/ versus /si/-/ʃi/ study in that, for each listener, the segmental contrast (/n/-/m/ or /d/-/b/ or /t/-/d/) and the type of the cue of interest—here, vowel-onset formant transitions—were held constant, and only the salience of the cue of interest was varied systematically. Salience was varied by manipulating the transitions' spectral informativeness, by pairing each consonant with two different vowels, as in Nittrouer and Studdert-Kennedy's (1987) study. For example, the /n/-/m/ contrast was paired with the vowels /o/ and /i/ to give the contrasts /no/-/mo/ and /ni/-/mi/. The vowel-onset formant transitions in /no/ are spectrally quite distinct from those in /mo/ (particularly F2, which differs in both direction and extent of movement across the two syllables), making transitions spectrally informative in determining consonant identity for this contrast. The vowel-onset formant transitions in /ni/ are quite similar to those in /mi/, making transitions much less informative in identifying the consonants in this contrast. Each listener in the current experiment thus listened to two CV contrasts with identical consonants, but with different degrees of spectral information available in the transitional cues.

Figure 1 shows stylized spectrograms of the three contrast pairs used in this study: (*i*) /no/-/mo/ versus /ni/-/mi/, (*ii*) /do/-/bo/ versus /de/-/be/, and (*iii*) /ta/-/da/ versus /ti/-/di/.[2] It should be noted that when we refer to transitional cues for the two /t/-/d/ contrasts, we are in fact referring to *voiced* transitions: Unlike many studies of perception of voicing contrasts (see, e.g., Kuhl and Miller, 1978), the synthetic stimuli used in the current study did not contain changing formant transitions for formants higher than F1 during aspiration. Instead, changing formant transitions for all three first formants began at voicing onset. Thus, although these stimuli sounded authentic, they were technically somewhat impoverished in comparison to natural speech.

In summary, this study aims to test the viability of Sussman's (2001) sensory explanation as an alternative, or an addition, to the DWS theory. Sussman's claim proposes that adult-child cue weighting differences are caused by children's heavier reliance on cues that are louder, longer, or more spectrally informative. The current study tests this claim by examining children's and adults' weighting of transitional cues in *both* a spectrally informative condition *and* a spectrally less informative condition. If the current study finds that children consistently give comparatively more weight than adults to the more spectrally informative transitions, and less weight than adults to the less spectrally informative transitions, then Sussman's claim will be supported.

It should be noted that Sussman's claim stems from a theory that children tend to rely on more informative cues due to some sort of sensory immaturity, possibly in terms of their central auditory processing. The current study did not attempt to address the issue of whether children and adults do in fact differ in their central auditory processing abilities. Additionally, this study did not directly test any sensory explanation for cue weighting behavior, but simply asks whether adult-child differences based on spectral distinctiveness or informativeness exist. Nevertheless, a positive result in the current study would be consistent with a sensory explanation of adult-child cue weighting differences.

II. METHOD

A. Participants

For the two /n/-/m/ contrasts, our testing criterion (described in Sec. II D) was met by ten out of ten adults tested (age range 21–29 years, average age 22 years), nine 7-year-olds (age range 7;1–7;11, average age 7;5 [year;month]) out of ten 7-year-olds tested, ten 5-year-olds (age range 5;0–5;9, average age 5;5) out of ten 5-year-olds tested, and four 3- to 4-year-olds (age range 4;4–4;10, average age 4;8) out of seven 3- to 4-year-olds tested.

For the two /d/-/b/ contrasts, our testing criterion was met by seven out of seven adults tested (age range 21–33 years, average age 26 years), nine 7-year-olds (age range 6;11–7;9, average age 7;5) out of ten 7-year-olds tested, seven 5-year-olds (age range 5;1–5;11, average age 5;5) out of seven 5-year-olds tested, and six 3- to 4-year-olds (age range 3;5–4;6, average age 3;11) out of eight 3- to 4-year-olds tested.

For the two /t/-/d/ contrasts, our testing criterion was met by eight out of eight adults tested (age range 21–49 years, average age 33 years), ten 7-year-olds (age range 7;0–7;11, average age 7;7) out of 11 7-year-olds tested, nine 5-year-olds (age range 5;1–5;8, average age 5;5) out of ten 5-year-olds tested, and nine 3- to 4-year-olds (age range 3;7–4;11, average age 4;1) out of 18 3- to 4-year-olds tested.

All children were in full-time preschool or primary (first and third year) education in Edinburgh (Scotland) and all were monolingual native speakers of Scottish Standard English (SSE). All of the children 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 in the first 3 years, 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 who had lived in the Edinburgh area for an average of 14 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

A trading-relations design was used for this study. For each individual contrast, two acoustic cues were manipulated: One vowel-onset transition cue and one nontransitional cue. The nontransitional cues were: Nasal murmur for /no/-/mo/ and /ni/-/mi/, frequency of stop burst for /do/-/bo/ and /de/-/be/, and voice onset time (VOT) for /ta/-/da/ and /ti/-/di/. The transitional cues were the changing formant frequencies (F1, F2, and F3) at the onset of the vowel (or at the onset of voicing in the case of the two /t/-/d/ contrasts). Two nine-point continua were created for each contrast by manipulating the two relevant acoustic cues for the contrast in question. The nontransitional cue was varied *along* the two continua from a value appropriate for the first consonant in the contrast to a value appropriate for the second consonant in the contrast. For example, for the /no/-/mo/ contrast, the nasal murmur was varied from a value appropriate for /no/ to one appropriate for /mo/ along both continua. The transitional cue, on the other hand, was varied *across* the two continua, with a value on one continuum that was appropriate for having followed the first consonant in the contrast, and a value on the other continuum that was appropriate for having followed the second consonant in the contrast. Thus, for the same /no/-/mo/ contrast, the value of the vowel formants at vowel onset were appropriate for having followed /n/ for the first continuum and appropriate for having followed /m/ for the second continuum. This design results in two continua which are identical in terms of the cue that varies along the continua and differ in terms of the cue that varies across the continua. This in turn allows for an investigation of the perceptual effect of the manipulated cues. A listener who is not influenced by the cue that changes across the continua—the vowel-onset transition information—will perceive the two continua as identical. On the other hand, a listener who *is* influenced by the cue that changes across the continua will perceive the two continua differently. Following Nittrouer (1992), five different repetitions of the same vowel were created for each transition condition. Each of these was combined with each of the nine continuum values, resulting in 90 different stimuli per consonant contrast.

The stimuli used in this study were copy-synthesized versions of /no/-/mo/, /ni/-/mi/, /do/-/bo/, /de/-/be/, /ta/-/da/ and /ti/-/di/, created using the Sensyn (Sensimetrics Corp., n.d.) version of the Klatt cascade/parallel synthesizer (cascade for voiced and aspirated sounds, parallel for fricative and plosive burst sounds, Klatt, 1980). Copy synthesis is a method in which the values used to synthesize stimuli are closely modeled on detailed acoustic analysis of natural tokens of the target syllables (e.g., Hazan and Rosen, 1991). In this case, the targets were recorded by a male native speaker of Scottish Standard English, and acoustic formant measurements were taken at 10-ms intervals. It should be noted that as a result of this synthesis method, the change in formant frequencies in the synthetic vowel-onset formant transitions is neither stylized nor interpolated as a straight line change between vowel onset and vowel target values. Instead, the formants undergo a fairly rapid change in frequency immediately following the consonant (completed in 25–35 ms for bilabials, 30–40 ms for alveolars), followed by a slower change as the frequencies gradually reach the vowel target value. This gives formant transitions which resemble exponential curves. Additionally, because the durations of the transitional portions of the synthetic vowels are based on measurements of the whole natural vowel portion that changed in frequency over time (including the portion leading up to the vowel target which often contained more gradual changes), the durations of the synthetic transitions differ depending on the syllable being modeled. In particular, as will be noted below, contrasts with more extensively changing transitions (for example those in /no/-/mo/) will often show greater differences in total transition duration between the members of the contrast than will contrasts with less extensively changing transitions (such as those in /ni/-/mi/). A more detailed description of the parameters manipulated for each contrast follows, presented in pairs based on the initial consonants of the contrasts. All other parameters were left at the default values provided by the Klatt synthesizer (Sensimetrics Corp., n.d.). Note that the values used to synthesize the nontransitional cues were identical across consonant pairs (e.g., the nasal murmur values were the same for the /no/-/mo/ continua and the /ni/-/mi/ continua). Vowel-formant onset and target values for each individual vowel repetition can be found in Appendix A.

1. /ni/-/mi/ and /no/-/mo/ contrasts

Nasalization in the initial consonant and the first 25 ms of the following vowel was modeled by adding a single pole-zero pair to vowel-like spectra. Nine nasal portions were synthesized, ranging from the most /m/-like (F1: 207 Hz, F2: 1150 Hz, F3: 2200 Hz, nasal pole: 900 Hz, nasal zero: 1200 Hz) to the most /n/-like (F1: 207 Hz, F2: 1750 Hz, F3: 2832 Hz, nasal pole: 1000 Hz, nasal zero: 1800 Hz). For all the syllables, both the pole and the zero fell from their respective values to 500 Hz over 20 ms after the onset of the vowel (at which point they merged, thus canceling each other out). The vowel from this point onward was non-nasal.

Two sets of vowels were created for each contrast, one with onset frequencies appropriate for a preceding /n/ and one with onset frequencies appropriate for a preceding /m/. The average /ni/-transition formant onset frequencies were F1: 252 Hz, F2: 1942 Hz, F3: 2632 Hz; the average /mi/-transition formant onset frequencies were F1: 262 Hz, F2: 1828 Hz, F3: 2494 Hz. The average vowel target values for all ten /ni/-/mi/ vowels were F1: 345 Hz, F2: 2228 Hz, F3: 2849 Hz. The average /no/-transition formant onset frequencies were F1: 267 Hz, F2: 1619 Hz, F3: 2299 Hz; the average /mo/-transition formant onset frequencies were F1: 257 Hz, F2: 801 Hz, F3: 2363 Hz. The average vowel target values for all ten /no/-/mo/ vowels were F1: 427 Hz, F2: 824 Hz, F3: 2421 Hz.

The total duration of each syllable was 485 ms, with 95 ms of nasal murmur, 25 ms of nasalized vowel, and 365 ms of oral vowel. The average duration of vowel formant transitions (including both rapidly and gradually changing portions) as measured from vowel onset to vowel steady state was 112 ms for /no/-transition stimuli, 74 ms for /mo/-transition stimuli, 50 ms for the /ni/-transition stimuli, and 58 ms for the /mi/-transition stimuli. F0 for each complete syllable began at 140 Hz at voicing onset, rose to 166 Hz 140 ms after onset, and fell to 97 Hz at vowel offset.

2. /de/-/be/ and /do/-/bo/ contrasts

Nine different complex bursts were synthesized. The spectral shape of the bursts was modeled by means of three spectral peaks, determined by the shape of the vocal tract at consonant release. The amplitude of these peaks ranged from 54 dB at 5550 Hz, 36 dB at 2700 Hz, and 20 dB at front cavity peak[3] (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 vowels were created for each contrast, one with onset frequencies appropriate for having followed /d/ and one with onset frequencies appropriate for having followed /b/. The average /de/-transition formant onset frequencies were F1: 220 Hz, F2: 1809 Hz, F3: 2446 Hz; the average /be/-transition formant onset frequencies were F1: 257 Hz, F2: 1694 Hz, F3: 2247 Hz. The average vowel target values for all ten /de/-/be/ vowels were F1: 428 Hz, F2: 2116 Hz, F3: 2539 Hz. The average /do/-transition formant onset frequencies were F1: 240 Hz, F2: 1558 Hz, F3: 2389 Hz; the average /bo/-transition formant onset frequencies were F1: 236 Hz, F2: 722 Hz, F3: 2378 Hz. The average vowel target values for all ten /do/-/bo/ vowels were F1: 422 Hz, F2: 855 Hz, F3: 2469 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 (including both rapidly and gradually changing portions) as measured from vowel onset to vowel steady state was 107 ms for /do/-transition stimuli, 37 ms for /bo/-transition stimuli, 100 ms for /de/-transition stimuli, and 110 ms for /be/-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/ and /ti/-/di/ contrasts

Nine different VOT values were synthesized, varying in 5-ms steps from 40 ms of aspiration (as generated by the AH parameter in the Klatt synthesiser, Sensimetrics Corp., n.d.) (most /t/-like) to 0 ms of aspiration (most /d/-like).

Two sets of vowels were created for each contrast, one with voiced onset frequencies appropriate for having followed /t/ and one with voiced 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 ten /ta/-/da/ vowels were F1: 711 Hz, F2: 1433 Hz, F3: 2665 Hz. 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 /ti/-/di/ 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 voiced vowel in all cases. The average duration of vowel formant transitions (including both rapidly and gradually changing portions) as measured from voiced vowel onset to vowel steady state was 55 ms for /ta/-transition stimuli, 85 ms for /da/-transition stimuli, 110 ms for /ti/-transition stimuli, and 105 ms for /di/-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. Procedure

Each participant listened to, and identified, stimuli from both the spectrally distinct and the spectrally similar conditions of a consonant contrast. Thus, listeners who listened to the /no/-/mo/ contrast, for example, also listened to the /ni/-/mi/ contrast. All participants were tested individually in a quiet room. The stimuli were presented over headphones (Sennheiser HD 490, frequency response 17–22 000 Hz), via a CD player. Volume was set at a comfortable listening level. Each participant was asked to indicate that the level was both comfortable and audible (for the children, 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 participants. No adjustments to listening level were made within the presentation of a single contrast. Testing for the children took place over two or three days. Testing for the adults took place on one day, with a short break half-way through testing.

The listener's task was to identify individual stimuli as either one or the other half of a given contrast (e.g., as either "no" or "mow" for the contrast /no/-/mo/). The adult participants performed the task alone, by entering their responses on a form. The child participants provided their responses to the experimenter by saying the word aloud, and by placing a counter on a picture corresponding to the relevant word (see Appendix B for a description of the pictures used in the study).

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 ten randomly presented natural stimuli (five of each CV syllable).

A pretest was administered to both child and adult participants to ensure that they understood the task. This test consisted of the congruent endpoints of the continua, that is, the most extreme values of the nontransitional cue combined with the appropriate vowel-onset frequencies. For example, for the /ta/-/da/ contrast, the congruent endpoints were the 0-ms VOT plus vowels with /d/-transitions (the most /da/-like stimuli) and the 40-ms VOT plus vowels with /t/-transitions (the most /ta/-like stimuli). There were ten stimuli in the pretest (five per congruent endpoint), presented in random order. No feedback was given during this pretest.

Five different random orders of the 90 stimuli were created for each contrast. During the main test, the 5-year-old, 7-year-old, and adult participants heard the entire set of 90 stimuli twice, in two different random orders, resulting in 180 responses per participant and ten responses per vowel-onset type for each point on the continuum. The 3- to 4-year-old participants heard the set of 90 stimuli only once, resulting in five responses per vowel-onset type for each point on the continuum for this group. Although this smaller number of presentations may have led to noisier data than if ten 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 3-to 4-year-olds were not qualitatively different from those of the other child participants. Each randomization was split into blocks of ten stimuli for presentation. The interstimulus interval for the adults was 3 s, with an interblock interval of 10 s. Following Walley and Carrell (1983), there was no fixed interstimulus interval for the child participants. Instead, the presentation was paused briefly after every stimulus, allowing the children sufficient time to respond. At the end of each block, the children were allowed to choose a small prize (a sticker).

D. Analysis

Each consonant contrast pair had two sets of responses, one for each vowel-onset transition condition—i.e., spectrally similar or spectrally distinctive. It was important to maintain consistency of participants across the spectrally distinctive and spectrally similar conditions (that is, to ensure that if data from one participant were analyzed for the spectrally distinctive condition, data for that participant were also analyzed for the spectrally similar condition). Therefore, only data from those listeners who responded correctly to 80% of the within-test congruent endpoints from the *spectrally distinctive* condition were included in analysis. Thus, if a participant was able to correctly identify 80% of the /no/-/mo/ congruent endpoints, for example, the responses of that participant for both the /no/-/mo/ and the /ni/-/mi/ contrasts were analyzed.

The responses for the two /t/-/d/ contrasts were normalized using a probit 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 is assumed to correspond 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% /t/ or /d/ responses). The responses for the two /n/-/m/ and two /d/-/b/ contrasts (which did not engender S-shaped response curves) were subjected to linear regression analysis. This yielded estimates of the slope and the *y* intercept of each regression line.

For each of the six individual contrasts (two per consonant contrast pair), the *degree of separation* of the two response curves/regression lines was calculated. For the probit-transformed responses, this was done by taking the difference of the two means; for the remaining responses, separation was calculated by taking the difference of the two *y* intercepts. Separation is thus a measure of the extent to which responses were affected by the change in vowel-onset transitional information across the two continua. Similarly, the *slope* of the response curves can be regarded as a measure of the degree to which listeners' responses were affected by the change in nontransitional information along the continua. Analyses of variance (ANOVAs) were calculated with either separation or slope of response curves as independent variables.

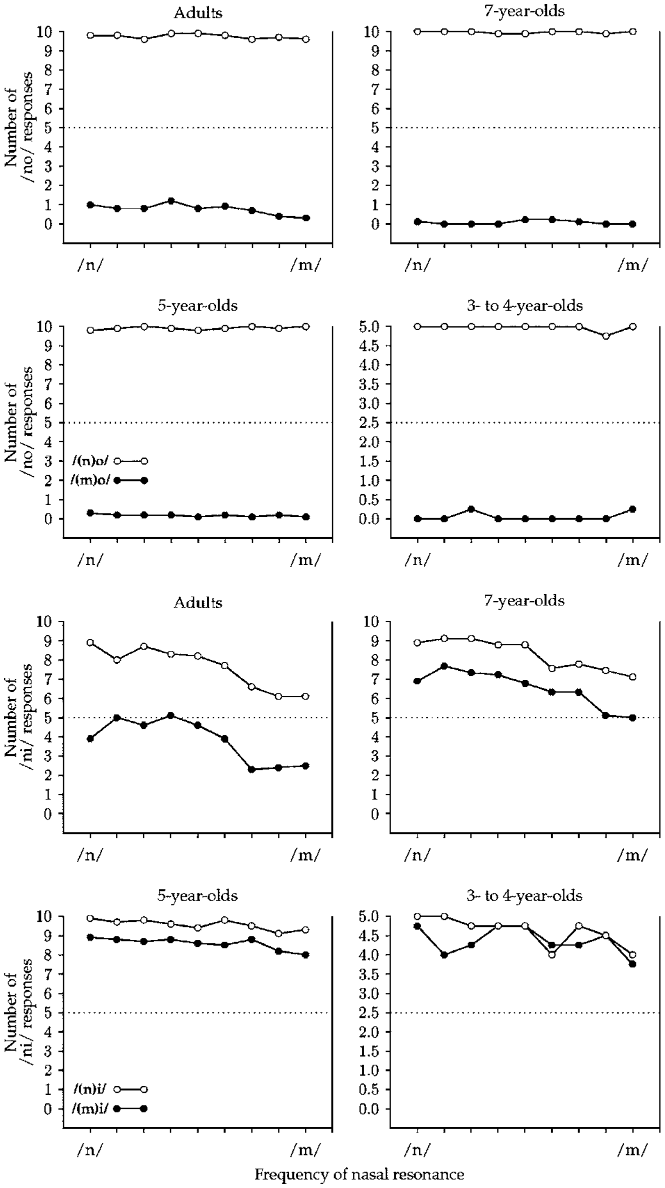

FIG. 2. Adults' and children's responses to /n/-transition stimuli (open circles) and /m/-transition stimuli (filled circles). The top four graphs represent responses to the /no/-/mo/ contrast; the bottom four graphs represent responses to the /ni/-/mi/ contrast. Responses are presented in terms of /n/-responses as a function of frequency of complex nasal murmur ranging from most /n/-like to most /m/-like (see text for details of frequency range). The dotted lines indicate the 50% /n/ response point. The *y* axis range for the 3- to 4-year-olds is half that of the other participants because this group heard half as many repetitions per point on the continuum.

III. RESULTS

This section presents results for pairs of contrasts. As noted above, we did not examine the relative weighting of the two cues (transitional versus nontransitional) available within a single contrast (e.g., /no/-/mo/), but rather examined weighting of transitional cues *across* pairs of contrasts (e.g., transitional cues in /no/-/mo/ versus transitional cues in /ni/-/mi/). While it may be the case that the steady-state nontransitional cues are more or less heavily weighted than the transitional cues in any one individual contrast, our main interest was to identify any possible changes in weighting of transitions due to the change in spectral distinctiveness of that cue across the pairs of contrasts. The above comparison allows us to determine the influence of cue distinctiveness/informativeness on perceptual weighting.

Figures 2–4 show the average responses per age group for each contrast. It is clear from these graphs that there is an overall influence of degree of spectral informativeness on both adults' and children's responses. All listeners showed larger separation of response curves (indicating heavy weighting of the cue that changed across the continua, namely the transitional cue) for those contrasts in which transitions were spectrally distinctive (/no/-/mo/, /do/-/bo/, /ta/-/da/) compared to those where transitions were spectrally similar (/ni/-/mi/, /de/-/be/, /ti/-/di/). ANOVAs with separation of response curves as the independent variables showed that these apparent differences were significant: /no/-/mo/ versus /ni/-/mi/ $[F(1,64)=163.00, p<0.001]$; /do/-/bo/ versus /de/-/be/ $[F(1,56)=85.00, p<0.001]$; /ta/-/da/ versus /ti/-/di/ $[F(1,70)=72.76, p<0.001]$. It should also be noted (although it is not the focus of this study) that there is some influence of context (specifically, the degree of spectral informativeness of the transitional cue) on adults' and children's weighting of the *non*transitional cues, particularly for

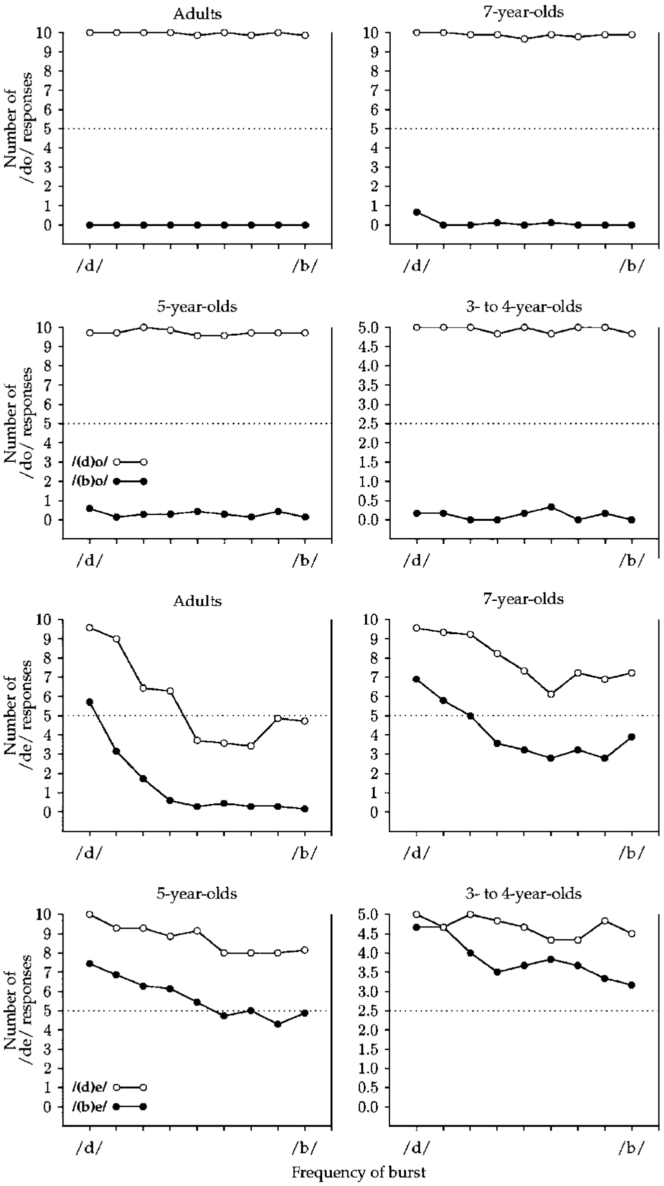

FIG. 3. Adults'and children's responses to /d/-transition stimuli (open circles) and /b/-transition stimuli (filled circles). The top four graphs represent responses to the /do/-/bo/ contrast; the bottom four graphs represent responses to the /de/-/be/ contrast. Responses are presented in terms of /d/-responses as a function of frequency of complex stop burst ranging from most /d/-like to most /b/-like (see text for details of frequency range). The dotted lines indicate the 50% /d/-response point. See Fig. 2 for more details.

the /n/-/m/ and /d/-/b/ contrasts. In spite of the fact that the acoustic characteristics of the nasal murmur continua were identical in the /no/-/mo/ and /ni/-/mi/ contrasts, and likewise the stop burst continua were identical in the /do/-/bo/ and /de/-/be/ contrasts, subjects were not able to use these cues when they were paired with the relatively distinctive /o/ transitions and steady states (as indicated by the lack of change in listeners' responses along the continua of steady-state cues). Given that the /ni/-/mi/ and /de/-/be/ results show that these nontransitional cues can be successfully used in categorization by all adult listeners and many older child listeners, it is clear that listeners' lack of successful use of these cues in the /no/-/mo/ and /do/-/bo/ contrasts is not simply due to their acoustic composition.

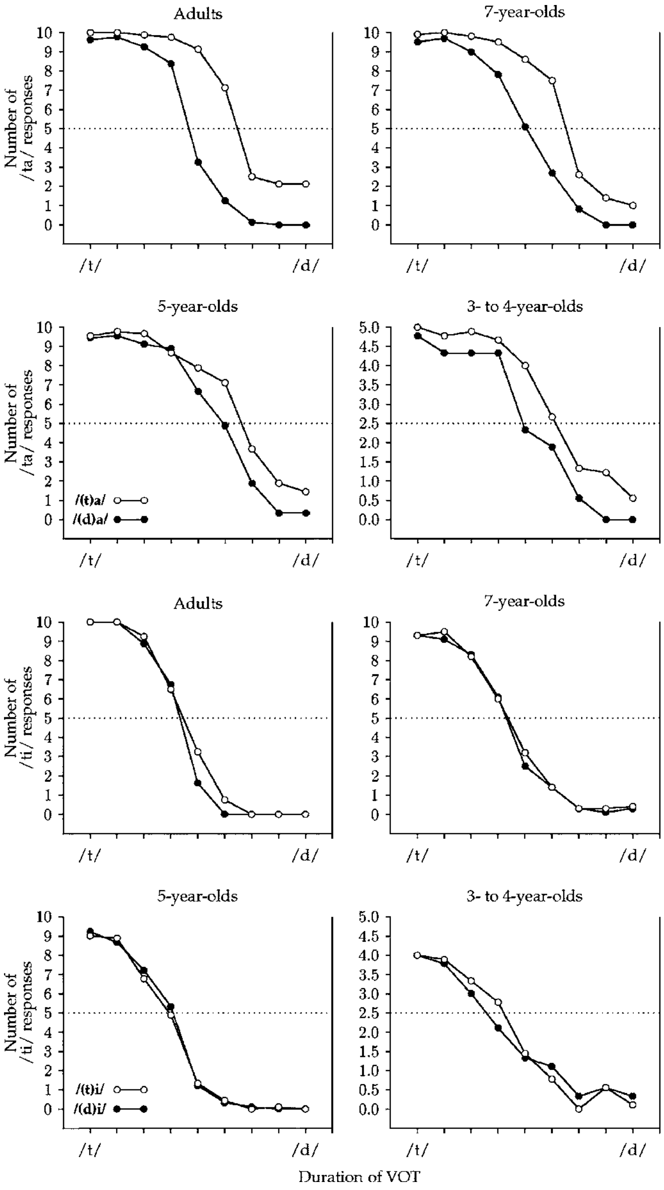

FIG. 4. Adults' and children's responses to /t/-transition stimuli (open circles) and /d/-transition stimuli (filled circles). The top four graphs represent responses to the /ta/-/da/ contrast; the bottom four graphs represent responses to the /ti/-/di/ contrast. Responses are presented in terms of /t/-responses as a function of VOT ranging from most /t/-like (40 ms) to most /d/-like (0 ms). The dotted lines indicate the 50% /t/ response point. See Fig. 2 for more details.

Turning to a comparison of adult and child transitional cue weighting behavior, the graphs indicate that for some consonantal contrasts, specifically /n/-/m/ and /d/-/b/, a change in spectral distinctiveness of transitional cues from more distinctive to less distinctive had a greater impact on children than on adults. For the spectrally distinctive /no/-/mo/ contrast, adults and children both had very widely separated response curves (again indicating heavy weighting of the formant transition cue). However, when the spectral distinctiveness of the transition was reduced in the /ni/-/mi/ contrast, children paid much less attention to the transitional cue than did adults. This is reflected in the fact that the separation of children's response curves was much smaller than that of adults for /ni/-/mi/. In fact the transitions in the

/ni/-/mi/ contrast did not cue a reliable difference between the two members of this contrast for most children. Similarly, for the /do/-/bo/ versus /de/-/be/ contrast pair, adults and children both showed widely separated curves in response to the spectrally distinctive transitions in /do/-/bo/, but children showed a smaller separation of curves than adults in response to the spectrally less distinct transitions in /de/-/be/. ANOVAs with separation of response curves as the independent variable again support these observations. There was no significant difference between adults and children in response curve separation for /no/-/mo/ but children showed significantly smaller response curve separation than adults for /ni/-/mi/ $[F(1,31)=16.85, p<0.001]$. A small difference was found between adults and children in response curve separation for /do/-/bo/ $[F(1,27)=4.67, p=0.04]$, however both groups appeared to make almost exclusive use of transitional cues in identifying this contrast. For /de/-/be/, the 7-year-olds patterned with the adults (there was no significant difference in response curve separation between these two groups) but the two younger groups of children (3- to 4-year-olds and 5-year-olds) again showed significantly smaller separation of response curves than the older listeners $[F(1,27)=6.40, p=0.02]$. These results thus lend support to Sussman's suggestion that children may have difficulty making use of acoustic cues—in this case transitional cues—that are less spectrally informative.

However, this pattern of children making less use than adults of less distinctive acoustic information did not hold for the two /t/-/d/ contrasts. First, there was *no* significant difference between adults and children in response curve separation for the contrast with spectrally similar transitions, /ti/-/di/. This seems inconsistent with Sussman's claim which would have predicted that children should have been less able to make use of this less informative cue than adults. However, it should be noted that the response curve graphs indicate that both adults and children appeared to make almost exclusive use of the nontransitional cue, VOT, and almost no use of the transitional cue. This exclusive use of VOT by both groups could explain a lack of difference in transitional cue use between the two groups. This could also suggest that, for this contrast, VOT was more informative to listeners than vowel formant transitions, although we cannot definitively establish degree of "informativeness" in this case, since VOT and vowel formant transitions are different types of acoustic cue. If VOT were the more informative cue, then we might expect to see a difference between adults and children in their weighting of this cue, as Sussman claims that children should make more use than should adults of more informative cues. However, an examination of the average slope of adults' and children's response curves (which reflects degree of weight given to the cue that changes *along* the continua, here VOT) for the /ti/-/di/ contrast shows that children's curves were less steep than were those of adults: An ANOVA with average response curve slope as the independent variable shows a significant difference between adults and children for this measure $[F(1,34)=29.53, p<0.001]$. This indicates that children gave *less* weight than adults to the VOT cue, counter to Sussman's predictions.^{4}

Furthermore, this finding of adults giving more weight than children to a more informative cue is repeated in the results of the /ta/-/da/ contrast. Here, again counter to Sussman's predictions, adults were found to have response curves that were more separated than those of the children, indicating that they gave more weight than children to the spectrally distinctive transitional cue: An ANOVA with response curve separation as the independent variable shows a significant difference between adults and children for this measure $[F(1,34)=6.00, p=0.02]$. No difference was found between adults and children in weight given to VOT for this contrast. The results for the two conditions of the /t/-/d/ contrast are clearly problematic for Sussman's hypothesis, since adults, rather than children, were shown to give most weight to a spectrally more informative cue.

IV. SUMMARY AND DISCUSSION

The goal of the current study was to determine to what extent the spectral distinctiveness or informativeness of a given cue influences listeners' cue weighting patterns. In particular, this study aimed to test the effect of changing the spectral distinctiveness of vowel-onset formant transition cues on cue weighting in 3- to 7-year-old children as compared to adults. The motivation behind this test was a claim that possible sensory immaturities in their central auditory processing would cause children to be less able to use quieter, shorter, and less spectrally informative cues, and thus more likely to make heavier use than adults of louder, longer, and more spectrally informative cues (Sussman, 2001).

The results of the current study provide evidence that spectral informativeness of transitional cues (i.e., the degree of difference in spectral extent and direction of vowel-onset formant transitions) plays a role in speech perception tasks in general. All listeners, both adult and child, were more influenced by the spectrally distinctive vowel-onset formant transitions found in /no/-/mo/, /do/-/bo/, and /ta/-/da/ than they were by the spectrally similar transitions found in /ni/-/mi/, /de/-/be/, and /ti/-/di/.

This study also provides evidence that, for some consonantal contrasts, adults and children do differ in the degree to which their perception is affected by a change from spectrally distinctive to spectrally similar information. Although children and adults did not differ in their weighting of spectrally distinctive transitional information for /no/-/mo/, young children gave much less weight than did older children and adults to the spectrally similar transitional information in /ni/-/mi/. Similarly, children and adults differed only slightly in their weighting of spectrally distinctive transitions for /do/-/bo/, but children gave much less weight than did adults to the spectrally similar transitions in /de/-/be/. This provides some support for Sussman's claim that adult-child cue weighting differences could be caused by an inability in children to make use of shorter, quieter, or less informative cues.

However, the results for the /t/-/d/ contrasts, although providing further evidence of adult-child differences in cue weighting, do not support Sussman's sensory hypothesis. First, no difference was found between adults and children in the weight given to the spectrally similar transitional cue in /ti/-/di/. A theory based on general auditory differences between adults and children, however, would predict that children should give less weight to this cue than adults. While this result could possibly be explained by the fact that it appears that *both* adults and children made very little use of the transitional cue for this contrast, adults were found to have made *more* use than children of the other available cue to the contrast, VOT. Again, this runs counter to a general auditory theory, which would predict that children should rely more heavily than adults on the more acoustically salient cue. Finally, children were also found to weight the spectrally distinctive transitions in /ta/-/da/ less heavily than adults; Sussman's sensory hypothesis predicts that children should weight more salient cues more heavily than should adults. The conclusion that can be drawn from the current study, therefore, is that while at least some adult-child cue weighting differences could be explained by a sensory hypothesis, this is not the case for all adult-child cue weighting differences.

Interestingly, the results for the two /t/-/d/ contrasts are also not predicted by Nittrouer's more strategy-based DWS hypothesis. Children were *not* found to weight transitional information more heavily than adults for either /ti/-/di/ or /ta/-/da/ (see also Mayo and Turk, 2004). It would seem, therefore, that any explanation for adult-child cue weighting which is based solely on sensory differences, or on a strong, transition-based interpretation of strategy differences, between adults and children, will not be able to account for all cue weighting differences between the two groups. Further research will be required to determine whether a combination of sensory and strategy explanations can account for all cue weighting differences, or whether these differences could be due to other general cognitive or perceptual factors.

ACKNOWLEDGMENTS

This work was supported by a grant from the Wellcome Trust. The authors would like to thank all of the adults and children who participated in the study. We are also very grateful to the parents, teachers, and schools of the child participants for their generous help in making the study possible. Thanks go to 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 /ni/ stimuli.

Stimulus no.	F1 onset	F1 target	F2 onset	F2 target	F3 onset	F3 target
1	272	363	1935	2207	2614	2805
2	261	352	1931	2231	2632	2831
3	245	336	1938	2261	2636	2882
4	232	335	1922	2258	2632	2877
5	248	366	1986	2261	2645	2849

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

Stimulus no.	F1 onset	F1 target	F2 onset	F2 target	F3 onset	F3 target
1	255	344	1808	2220	2482	2837
2	276	342	1818	2223	2492	2868
3	238	357	1851	2208	2492	2829
4	283	334	1816	2202	2508	2872
5	260	325	1847	2211	2496	2835

Vowel-formant onset and target values for synthetic /no/ stimuli.

Stimulus no.	F1 onset	F1 target	F2 onset	F2 target	F3 onset	F3 target
1	270	412	1623	811	2293	2396
2	269	410	1579	808	2311	2388
3	246	414	1620	803	2217	2398
4	279	400	1614	849	2379	2440
5	270	438	1661	850	2293	2396

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

Stimulus no.	F1 onset	F1 target	F2 onset	F2 target	F3 onset	F3 target
1	259	428	890	805	2337	2415
2	234	443	757	833	2436	2462
3	261	432	765	837	2422	2461
4	259	440	792	842	2320	2424
5	271	452	802	801	2300	2429

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 /do/ stimuli.

Stimulus no.	F1 onset	F1 target	F2 onset	F2 target	F3 onset	F3 target
1	258	423	1522	877	2426	2468
2	244	420	1545	876	2355	2453
3	219	427	1587	825	2400	2470
4	275	419	1560	878	2360	2470
5	206	424	1576	840	2403	2497

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

Stimulus no.	F1 onset	F1 target	F2 onset	F2 target	F3 onset	F3 target
1	258	422	724	880	2407	2498
2	231	409	729	836	2353	2469
3	219	422	725	837	2396	2473
4	244	429	720	854	2369	2442
5	228	421	721	847	2359	2451

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	Description of picture
[ni]	a knee
[mi]	a girl pointing to herself (“me”)
[no]	a woman wagging her finger to say “no”
[mo]	a lawn-mower mowing a lawn
[de]	a street scene with the sun coming up (“day”)
[be]	a typical Scottish bay
[do]	a person making cookie dough
[bo]	a ribbon tied in a bow
[ti]	a teapot and teacup (“tea”)
[di]	a girl called “Dee”
[ta]	a girl receiving a present (“ta” is British English slang for “thank you”)
[da]	a father (“da” short for “dada”)

---

[1] The exact physical location of the maturation has not yet been definitively established. However, most researchers accept that it is more likely to be in the central auditory system, as the auditory cortex continues to mature after birth (Eisenberg *et al.*, 2000), than in the peripheral auditory system, as the cochlea is fully developed at birth (Schneider and Trehub, 1992).

[2] The data for the /de/-/be/, /ta/-/da/, and /ti/-/di/ contrasts are also featured in Mayo and Turk (2004), where they are analyzed for a different purpose.

[3] This front cavity peak is modeled by the *frication excited bypass path* in Sensyn’s Klatt synthesiser (Sensimetrics Corp.); no frequency value is given by Klatt for this parameter.

[4] The results seen for the /t/-/d/ contrasts could be due in part to the way in which the stimuli were constructed, namely that the aspiration noise did not contain transitional information. However, Howell *et al.* (1992) found similar results for edited natural /b/-/p/ stimuli (rather than wholly synthetic as in the current study). Additionally, although some of our adult listeners were native speakers of non-Scottish varieties of English, all of the adult listeners for the two /t/-/d/ contrasts were native speakers of SSE. Therefore, these results, which are inconsistent with both the DWS and a sensory hypothesis, cannot be due to any dialect differences between the adult and child listeners.

Allen, P., and Wightman, F. (**1994**). “Psychometric functions for children’s detection of tones in noise,” J. Speech Hear. Res. **37**, 205–215.

Berg, K. M., and Boswell, A. E. (**2000**). “Noise increment detection in children 1 to 3 years of age,” Percept. Psychophys. **62**, 868–873.

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

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. (**1979**). “Performance of children aged 9–17 years on test of speech intelligibility in noise using sentence material with controlled word predictability,” J. Acoust. Soc. Am. **66**, 651–653.

Elliott, L. L., and Katz, D. (**1980**). “Children’s pure-tone detection,” J. Acoust. Soc. Am. **67**, 343–344.

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

Elliott, L. L., Connors, S., Kille, E., Levin, S., Ball, K., and Katz, D. (**1979**). "Children's understanding of monosyllabic nouns in quiet and in noise," J. Acoust. Soc. Am. **66**, 12–21.

Fletcher, H., and Munson, W. A. (**1933**). "Loudness, its definition, measurement and calculation," J. Acoust. Soc. Am. **5**, 82–108.

Hazan, V., and Rosen, S. (**1991**). "Individual variability in the perception of cues to place contrasts in initial stops," Percept. Psychophys. **49**(2), 187–200.

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).

Jensen, J., and Neff, D. (**1993**). "Development of basic auditory discrimination in preschool children," Psychol. Sci. **4**, 104–107.

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.

Kuhl, P. K., and Miller, J. D. (**1978**). "Speech perception by the chinchilla: Identification functions for synthetic VOT stimuli," J. Acoust. Soc. Am. **63**, 905–917.

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's-Gravenhage), pp. 229–238.

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).

Maxon, A. B., and Hochberg, I. (**1982**). "Development of psychoacoustic behaviour: Sensitivity and discrimination," Ear Hear. **3**, 301–308.

Mayo, C., and Turk, A. (**2004**). "Adult-child differences in acoustic cue weighting are influenced by segmental context: Children are not always perceptually biased toward transitions," J. Acoust. Soc. Am. **115**, 3184–3194.

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.

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. (**1996**). "The relation between speech perception and phonemic awareness: Evidence from low-SES children and children with chronic OM," J. Speech Hear. Res. **39**(5), 1059–1070.

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

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.

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**(2), 266–284.

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.

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

Schneider, B., and Trehub, S. (**1992**). "Sources of developmental change in auditory sensitivity," in *Developmental Psychoacoustics*, edited by L. Werner and E. Rubel (American Psychological Association, Washington, DC).

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

Sensimetrics Corp. (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.

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.

Trehub, S., Schneider, B., Morrongiello, B., and Thorpe, L. (**1988**). "Auditory sensitivity in school-age children," J. Exp. Child Psychol. **46**, 273–285.

Trehub, S. E., Schneider, B. A., and L., H. J. (**1995**). "Gap detection in infants, children, and adults," J. Acoust. Soc. Am. **98**, 2532–2541.

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., and Peach, S. (**1984**). "Developmental aspects of the perception of acoustic cues in determining the voicing feature of final stop consonants," Lang Speech **27**(4), 367–379.

Watson, J. (**1997**). "Sibilant-Vowel Coarticulation In The Perception Of Speech By Children With Phonological Disorder," PhD thesis, Queen Margaret College, Edinburgh.

Werner, L. A., Marean, G. C., Halpin, C. F., Spetner, N. B., and Gillenwater, J. M. (**1992**). "Infant auditory temporal acuity-gap detection," Child Dev. **63**, 260–272.