The Influence of Phonemic Awareness Development on Acoustic Cue Weighting Strategies in Children’s Speech Perception

Catherine Mayo, James M. Scobbie, Nigel Hewlett and Daphne Waters
Speech and Language Sciences, Queen Margaret University College, Edinburgh
Clerwood Terrace, Edinburgh, EH12 8TS, UK
short title: Phonemic awareness and cue weighting

1Correspondence should be addressed to Catherine Mayo, who is now at the Department of Theoretical & Applied Linguistics, University of Edinburgh, George Square, Edinburgh, EH8 9LL, UK; e-mail: catherin@ling.ed.ac.uk
Abstract

The patterns of weighting that children give to different acoustic characteristics of speech (their cue weighting) appear to change with increased linguistic experience. Previous speech perception research has found a correlation between children’s cue weighting strategies and their ability to consciously think about and manipulate segment–sized units (phonemic awareness). That research was not, however, able to determine whether the relationship is in any way causal, and if so whether it is phonemic awareness development that impacts on cue weighting strategies, or if changes in cue weighting allow for the later development of phonemic awareness. The aim of this study was to follow the development of these two processes longitudinally in order to determine which of the above two possibilities was more likely. Five–year–old children were tested 3 times in 7 months on their cue weighting strategies for a /so/–/ʃo/ contrast, in which the two cues being manipulated were the frequency of fricative spectrum and the frequency of vowel onset formant transitions. The children were also tested at the same time on their phoneme segmentation and phoneme blending skills. Results showed that phonemic awareness skills tended to improve before cue weighting changed, and that early phonemic awareness ability predicted later cue weighting strategies. These results suggest that the linguistic experience which has been suggested is the catalyst for cue weighting changes must include meta–phonemic development.
1 Introduction

The purpose of this study is to examine the relationship between weighting of acoustic cues in speech perception, and the development of meta-phonemic awareness. It is well established that speech contrasts are signalled, or “cued”, by means of multiple different characteristics of the acoustic signal—for instance, a /da/-/ta/ contrast can be signalled by both the duration of voice onset time and the onset frequency of the following vowel formants (e.g. Liberman, Dellatre & Cooper 1952, Liberman, Dellatre, Cooper & Gerstman 1954, Liberman, Dellatre, Gerstman & Cooper 1956, Liberman, Dellatre & Cooper 1958, Repp, Liberman, Eccardt & Pesetsky 1978). A number of studies have found that different cues to a speech contrast are not always equivalent in the relative role that they play in signalling that contrast. That is, in determining what speech sound they have heard, listeners do not always give equal importance, or weight, to all of the cues available to them (e.g. Dorman, Studdert-Kennedy & Raphael 1977, Ohde & Haley 1997, Walley & Carrell 1983, Wardrip-Fruin 1982, Wardrip-Fruin 1985, Whalen 1991).

There is also some evidence that patterns of cue weighting are not fixed developmentally: studies have shown that for certain contrasts adults and children weight acoustic cues differently. Nittrouer and colleagues, for example, have consistently found that in identifying /s/vowel–/ʃ/vowel contrasts children seem to give more weight to vowel onset formant transitions compared to adults, and relatively less weight than adults to the spectral characteristics of the fricative noise (e.g. Nittrouer 1992, Nittrouer 1996a, Nittrouer & Miller 1997, Nittrouer & Studdert-Kennedy 1987). Other studies have found further differences between children and adults in their relative weighting of acoustic cues (e.g. Greenlee 1980, Krause 1982, Lacerda 1992, Morrongiello, Robson, Best & Clifton 1984, Ohde & Haley 1997, Parnell & Amerman 1978, Watson 1997, Wardrip-Fruin & Peach 1984).
Nittrouer and colleagues have proposed a theory to explain this apparent developmental change in cue weighting strategy (e.g. Nittrouer, Manning & Meyer 1993, Nittrouer & Miller 1997). The hypothesis, called the Developmental Weighting Shift (DWS), asserts that “the weights assigned to various acoustic speech parameters change as the child gains experience with a native language, and that this developmental weighting shift is related to developmental increases in sensitivity to phonetic structure” (Nittrouer 1996b, pp. 1060–1061).

This hypothesis is rooted in theories that propose that the level of detail required to represent lexical items changes with lexical growth (see e.g. Jusczyk & Derrah 1987, Studdert-Kennedy 1987). Within this type of framework, a small lexicon, such as an infant or child would have, requires only a global or gross–grained level of detail to adequately accommodate and differentiate between all items stored within it. A small lexicon can therefore be represented in terms of syllables or mono–syllabic words. As the lexicon grows, a more fine–grained level of detail is required in order to differentiate between all the items. A larger lexicon (such as an older child or adult would have) must therefore be represented in terms of much smaller units, such as segments or possibly features. Studies of children’s speech production provide support for this view. Patterns in early child utterances seem to indicate that for children “the word is an entity, stored and accessed as a block” (Menn 1971, p. 247; see also Ferguson & Farwell 1975, Menn 1983, Nittrouer, Studdert-Kennedy & McGowan 1989, Studdert-Kennedy 1987). Children also appear to be more coarticulatory in their production than adults, suggesting that they organize their speech over larger frames (Nittrouer et al. 1989, Nittrouer 1993, Nittrouer 1995, Nittrouer, Studdert-Kennedy & Neely 1996). Studies that follow children’s production behaviour over time, on the other hand, “reveal a gradual qualitative shift from a predominance of processes affecting the structure of whole words (consonant harmony, reduplication, final consonant deletion) to those affecting specific segments or classes of segments (stopping of fricatives, gliding
Relating this framework to her perceptual findings, Nittrouer suggests that “perceptual strategies for speech may depend largely on the linguistic decision to be made” (Nittrouer 1996a, p. 295). That is, listeners’ perceptual cue weighting strategies are governed by the level of linguistic information they are trying to recover. If they are trying to recover global or gross–grained levels of detail—as it is proposed that young children do—then they should give more weight to acoustic cues primarily associated with more global characteristics of speech. If they are trying to recover more detailed phonetic structure—as it is proposed that adults do—then they should give more perceptual weight to acoustic cues primarily associated with fine–grained characteristics of speech (Nittrouer 1996a).

Nittrouer and colleagues go on to propose that the physical correlates of global or gross–grained characteristics are vowel onset formant transitions, as these are “perceptually salient and delimit signal portions corresponding to syllables” (Nittrouer, Miller, Crowther & Manhart 2000, p. 268): characteristics of transitions within a CV syllable depend on the place and manner of articulation of both the consonant and the vowel. Physical correlates of fine–grained characteristics, on the other hand, are suggested to be “all the language–specific properties that acoustic/phonetic research has found over the years to correspond to perceived phonetic units” (Nittrouer et al. 2000, p. 268). Therefore, the DWS theory essentially states that the reason that young children make relatively more use of transitional information than adults is that children perceive/process speech more globally, while adults make relatively less use of transitional information because they perceive/process speech more analytically.

Thus the “linguistic experience” referred to by Nittrouer and colleagues in their definitions of the DWS is taken to mean the lexical growth that occurs as a result of increased exposure to a native language. But, the DWS theory does not rule out the possibility
that “increased sensitivity to phonetic structure” and cue weighting changes could additionally be related to other linguistic experiences. In fact, Nittrouer (1996b) investigated the possibility of a relationship between cue weighting strategies and phonemic awareness skills. Phonemic awareness (which is closely tied to alphabetic literacy, see e.g. Morais 1991) is the ability to think about and manipulate phonemic segments, e.g. the ability to say that “phone” is made up of the sounds [f], [o] and [n] in that order. Someone who has not yet developed this skill will have the ability to think about larger units, such as syllables and onset–rimes (e.g., for the word “glad,” the onset is [gl], the rime is [æd]), but will not be able to access single segments. That is, they will be able to say that “please” has one syllable and rhymes with “sneeze”, but will say that the first sound in “please” is [pl] rather than [p]. The development of phonemic awareness, therefore, can be seen as a process of becoming able to consciously recover fairly detailed segmental information about speech. In this sense it appears to parallel the changes proposed to take place in perception from more global to more analytical levels of attention.

Nittrouer’s (1996b) study found that in 8–year–old children with various different social and linguistic histories (low or mid socio–economic backgrounds; with and without significant histories of otitis media) poor phonemic awareness skills correlated strongly with global cue weighting strategies, while good phonemic awareness correlated strongly with analytical cue weighting.

This result would suggest that cue weighting changes do not just parallel the development of phonemic awareness but are related to it. What Nittrouer’s study was unable to make clear, however, is the exact nature of the relationship. As Nittrouer points out, the cross–sectional design of the study meant that it could say “little about the direction of causality between the development of these processes” (Nittrouer 1996b, p. 1067). It is possible, therefore, that the development of phonemic awareness should be consid-
ered to be one of the linguistic experiences that influences changes in cue weighting. Equally, it could be the case that analytical cue weighting strategies develop as a result of other, non–metalinguistic, experiences, and the relationship observed is simply due to the fact that analytical cue weighting strategies allow for phonemic awareness to develop.

There are, however, some data from Nittrouer’s study which could indicate a possible causal direction in the relationship between cue weighting and phonemic awareness. For the most part the results of the study were markedly bimodal: children with good phonemic awareness skills had very analytical cue weighting strategies, and children with poor phonemic awareness skills had very global cue weighting strategies. However, Nittrouer highlights two children who fell outside these two groupings, with good phonemic awareness scores but global cue weighting. Additionally there were no subjects who showed poor phonemic awareness but analytical cue weighting. One explanation for this is that these two children represent an intermediate stage of development between the two larger groups of responses. If this is the case then it is possible that “discovering syllable–internal structure [i.e. the development of phonemic awareness] may actually create pressure to develop the most effective processing strategies for providing access to that structure” (Nittrouer 1996b, pp. 1067–1068).

Unfortunately this evidence is too minimal to allow any firm conclusions to be drawn. Additionally, the behaviour of the two subjects highlighted by Nittrouer is contradicted by the results of a number of longitudinal studies carried out by McBride-Chang and colleagues (e.g. Manis, McBride-Chang, Seidenberg, Keating, Doi, Munson & Petersen 1997, McBride-Chang 1995a, McBride-Chang 1996, McBride-Chang, Chang & Wagner 1997). These studies were not specifically designed to examine cue weighting. However, they do show that speech perception skill, as measured by performance on a categorical labelling task, together with memory and cognitive ability, predict
later phonemic awareness ability. If cue weighting strategies (that is, the pattern of weight given to available cues to a contrast) and categorical perception (the perception of a continuum of noise in terms of a limited number of phonemic categories) are both reflective of the same aspect of perceptual ability, then McBride–Chang and colleagues’ results would suggest that changes in cue weighting strategy might simply allow for the later development of phonemic awareness. It should, however, be noted that it has been shown for children older than those studied by Nittrouer (aged 6–years through 11–years) that while categoriality of responses (that is, how steeply categorical the responses were) continues to increase with age, cue use does not seem to change (Hazan & Barrett 2000). The fact that these two factors separate out in development suggests that cue weighting and categorical perception may in fact be different aspects of perception. This leaves the issue of a possible causal direction in the relationship between cue weighting and phonemic awareness still to be addressed.

The question underlying this investigation as a whole, therefore, is whether the relationship between cue weighting and phonemic awareness is in any way causal, and if so, in which direction the causality moves. That is, we are interested in finding out whether the development of phonemic awareness impacts in some way on listeners’ cue weighting strategies, or if changes in cue weighting are simply a prerequisite for the later development of phonemic awareness.

The first aim of the current study is to determine whether the behaviour of the two children who showed good phonemic awareness but global cue weighting can be replicated. The main aim, however, is to determine whether children who display this pattern of results are indeed at an intermediate stage of development between poor phonemic awareness/global cue weighting, and good phonemic awareness/analytical cue weighting. The study will therefore examine phonemic awareness and acoustic cue weighting within the framework of a longitudinal study. Both the relative speed at
which the two processes develop, and any predictive relationship(s) between the processes should provide evidence to constrain claims regarding the nature and direction of the relationship.

2 Method

2.1 Subjects

Eighteen children participated in this study: 8 female and 10 male. All 18 children were tested at Time 1 and 2 of the study; only 15 were available to be tested at Time 3. At the beginning of the study, the children ranged in age from 5;2 to 6;0 with an average age of 5;8; at the end of the study the average age was 6;3. All the children were in their first year of full–time primary education at schools in Edinburgh (UK), and all had undergone approximately 6–7 months of reading/reading–readiness training. All of the children were native speakers of Scottish Standard English (SSE, see e.g. Stuart-Smith 1999). Parental questionnaires determined that none of the children had a history of chronic otitis media (defined, by Nittrouer 1996b, as more than 3 ear infections in the first three years of life and/or the implantation of myringotomy tubes), and that none of the children or their siblings had ever been referred for speech and/or language therapy. All of the children had been tested at school for hearing deficits and all were reported to have hearing within a normal range.

In order to establish cue weighting norms for literate adults for the contrast used in this study, 8 adult listeners (4 female, 4 male) were assessed on their cue weighting

\(^2\)Six of the 18 spoke a second language in addition to English, to differing degrees of bilingualism, as reported by parents. However, no significant differences in performance were found between the bilingual and the monolingual children for any of the tests carried out in the study.
strategies. The adults ranged in age from 21 years to 52 years, with an average age of 27 years. All of the adult listeners were native speakers of English, and all had lived in the Edinburgh area for at least one year at the time of testing (average number of years: 12). None of the adults reported having hearing deficits or histories of chronic otitis media, and none had ever received therapy for expressive language disorders.

2.2 Test Stimuli

2.2.1 Cue weighting

The stimuli used for the current cue weighting tests were /so/–/jo/ (“sew”, “show”) stimuli that varied in terms of (i) the frequency of the fricative spectrum, and (ii) the frequency of the formants at vowel onset, both of which are fairly strong cues to the identity of the fricative (see Figure 1).

[Figure 1 about here]

The creation of the stimuli followed the basic design used by Nittrouer in most of her studies of /s/–/ʃ/ contrasts (e.g. Nittrouer 1992, Nittrouer 1996b): a modified categorical perception test based on a trading relations paradigm (Fitch, Halwes, Erickson & Liberman 1980). In this type of study, one of the two cues to the contrast is made to vary along a continuum. For Nittrouer’s studies, this cue was generally the frequency of the fricative spectrum, which varied from a frequency appropriate for /s/ to a frequency appropriate for /ʃ/. The other cue to the contrast takes one of two forms. In Nittrouer’s studies this second cue was the frequency of the vowel onset formant transitions, and the two forms were (i) onset transitions appropriate for a vowel following /s/, and (ii) onset transitions appropriate for a vowel following /ʃ/.

Combining each of the two vowels with each point on the fricative continuum results
in two /s/-vowel to /ʃ/-vowel continua, each with identical fricative noises, but with different vowel onset transitions, as illustrated in Figure 2. The premise of this methodology is that if a listener’s perception is strongly influenced by the vowel onset transitional information, then their category boundaries between /s/ and /ʃ/ should differ depending on the transitions (see response curves in Figure 2). If, on the other hand, a listener’s perception is predominantly not influenced by the vowel onset transitions, then their category boundaries should be the same, regardless of the transitions.

[Figure 2 about here]

The vowel context used in the current study was the Scottish monophthongal back vowel /o/. This context is comparable, in terms of the extent of the vowel onset transitions, to the North American English /u/ used by Nittrouer in her studies (e.g. Nittrouer 1992, Nittrouer & Miller 1997). This context was used because /u/ as a back vowel does not exist in SSE (Stuart-Smith 1999).

The stimuli used in this study were synthetic CV syllables. These syllables were created using a method known as copy synthesis (e.g. Hazan & Rosen 1991), in which highly detailed acoustic analyses are made of natural speech, and the resulting values used to synthesize the stimuli. An adult male SSE speaker recorded 10 repetitions each of the two target words (/so/, /ʃo/) in random order. The natural tokens were recorded onto DAT (Sony DTC–60ES) via microphone (Sony ECM–77B) and amplifier (Alice PAK2), and were transferred to computer for analysis. The speech was downsampled to 16kHz at this point.

All acoustic analysis was carried out using Entropics Waves+ software. The analysis consisted of durational measurements, spectral analysis of the fricative noise, spectral analysis of the vowel formants (8 measurements each were taken from F1, F2 and F3: 4 from the formant transitions, and 4 from the vowel target), and spectral analysis of F0 (3 measurements were made: onset of voicing, midpoint, offset of voicing).
The synthetic stimuli were created using *SenSyn*, Sensimetrics' cascade/parallel formant synthesizer (based on Klatt 1980). The overall duration of each synthetic stimulus was 480ms, with 230ms of fricative noise (see Nittrouer 1992), and 250ms of vowel. The main change in formant frequencies associated with the transitions occupied the first 75ms of each vowel, with any residual frequency change generally over by 125ms into the vowel.

Nine different fricative noises were designed. Each noise consisted of a single pole of aperiodic noise, varying along a continuum in 200Hz steps from 2.2kHz (most /j/-like) to 3.8kHz (most /s/-like). These endpoint values are consistent with those described in Nittrouer (1992). The amplitude of frication rose from 0dB at 0ms to 60dB at 90ms, staying at 60dB until 180ms and falling again to 30dB from 180ms to 230ms.

Measurements from 10 of the 20 natural tokens (5 each of /o/ and /so/) were chosen to model 10 synthetic vowels. This follows a strategy adopted by Nittrouer (1992) to ensure that listeners’ responses would not be influenced by idiosyncrasies in any one vowel utterance. Each set of 5 natural tokens was chosen based on the similarity of vowel formant frequencies and length of transitions. In addition, all 10 tokens were selected based on similarity of vowel target frequencies.

For the 5 /s/-transition vowels the average vowel onset formant frequencies were F1: 387Hz, F2: 1220Hz, F3: 2319Hz, and the average offset frequencies were F1: 387Hz, F2: 827Hz, F3: 2442Hz. For the 5 /j/-transition vowels the average onset frequencies were F1: 387Hz, F2: 1359Hz, F3: 1982Hz, and the average offset frequencies were F1: 387Hz, F2: 846Hz, F3: 2388Hz. The amplitude of voicing for all stimuli was 60dB from the beginning of the vowel for 185ms, falling gradually to 0dB from 185ms to 250ms (see Klatt 1980).

Each of the 10 synthetic vowels was combined with each of the 9 fricative noises, resulting in 90 different stimuli. F0 for each token began at 160Hz at 230ms (the onset
of voicing), rose to 180Hz at 355ms and then fell to 100Hz at 480ms. Figure 3 shows spectrograms of four of the 90 test stimuli.

[Figure 3 about here]

### 2.2.2 Phonemic awareness

Phonemic awareness was tested by means of two tasks: phoneme segmentation, in which the subject is required to divide a specified word into a sequence of separate phonemes, e.g. “phone” is [f], [o], [n], and phoneme blending, in which the subject is required to re-synthesize a number of separate phonemes into a single word, e.g. [f], [o], [n] is “phone”.

All of the words selected for the phonemic awareness tests were real words that had appeared five or more times in the CHILDES database (MacWhinney 1995). A number of factors have been shown to affect performance on phonemic awareness tests, namely: the type of phonemes to be manipulated (i.e. stop, fricative, nasal, vowel), placement of these different types of phoneme within words (i.e. word-initial, –medial, –final), total number of phonemes in a word, number of consonants in a consonant cluster, placement of consonant clusters, and possible phoneme–morpheme confounds for /t/, /d/, /s/, and /z/ in word–final position (McBride-Chang 1995b). Therefore, to the extent possible within the constraints of the database, the test items were chosen to be balanced within tests for the type and placement of phonemes. No words with word–final /t/, /d/, /s/, or /z/ were used if there was any way in which these sounds could be mistaken for past tense or plural morphemes (e.g. a child unfamiliar with the word “cord” could construe this as the past tense of the nonsense word [kɔr]). The total number of phonemes in each test word was manipulated deliberately (i.e. was increased throughout each test) in order to maintain a reasonably high level of difficulty.
throughout the longitudinal testing period.

Both tests had 50 test items each: 20 three–phoneme words, 20 four–phoneme words (CCVC and CVCC), and 10 five–phoneme words (CCVCC and CCCVC). The four–phoneme words were balanced for the position of the clusters. The five–phoneme words were predominantly CCVCC, with only 3 words in the blending test and 4 in the segmentation test of the configuration CCCVC.

The stimuli for all three tests (including training and pre–test stimuli) were recorded for presentation to the subjects by a phonetically trained, adult male SSE speaker. The speaker was instructed to produce all words clearly, and to produce all individual unvoiced phonemes without a following schwa, i.e. [s] rather than [so].

2.3 Procedures

All test materials were presented to the subjects using a portable MiniDisk player (Sony MZ–R3) via headphones. Testing of each subject took place individually in a quiet room.

Child subjects were tested three times over the course of 7 months, with testing taking place at months 1, 4 and 7. The testing for the child subjects was spread out over two consecutive days. Testing on each day consisted of a complete set of both cue weighting and phonemic awareness tests (each complete set was different: for the cue weighting tests, each child heard a different randomisation of the 90 stimuli each day, for the phonemic awareness tests, the words presented were different each day).

The adult subjects were tested on only one occasion, and only on their acoustic cue weighting strategies. The testing for the adults took place on one day, with a short break half way through testing.
2.3.1 Cue weighting

The children were introduced to the target words “sew” and “show” by means of a short story (recorded by the same adult male SSE speaker that had made all previous recordings) and an accompanying picture book. A synthesized version of the story was also presented in order to familiarize the children with computer generated speech.

During testing, the children were shown pictures from the story corresponding to the two target words, and were told to indicate which of the words they had heard by placing a counter on the appropriate picture. The experimenter recorded the children’s responses. Before testing, the children were trained on natural tokens of the two words, presented (unrecorded) by the experimenter. The children received feedback on their performance throughout this training.

The adult subjects did not listen to either version of the story or to the unrecorded natural tokens of the two target words. Instead they were told that they would hear repetitions of the two words “sew” and “show” in random order, and were instructed to indicate which they had heard by placing a tick in a box on a form provided.

A pre–test was administered to both child and adult subjects. This test consisted of the endpoints of the fricative continua with the appropriate vowel formant transitions for each fricative, i.e. the 3.8kHz noise plus vowels with an /s/–transition (the most “sew”–like stimuli) and the 2.2kHz noise plus vowels with an /ʃ/–transition (the most “show”–like stimuli). There were 10 stimuli in the pre–test (five vowels in each transition condition) which were presented in random order. All listeners were required to correctly identify 9 of the 10 pre–test stimuli in order for their results to be included in analysis.

For the test proper, 5 different randomisations of the 90 stimuli were generated, and each randomisation was split into 9 blocks of 10 stimuli for presentation. All subjects
heard the entire 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 fricative continuum. All listeners were required to respond correctly to at least 8 of the 10 continuum endpoints (i.e. the stimuli presented in the pre–test) for their results to be included in analysis. These criteria, and those described for the pre–test were used in order to eliminate any listener who was not completely attentive to the task.

Following the procedure adopted by Walley & Carrell (1983), the interstimulus interval was not fixed for presentation of the perceptual stimuli to the children. Instead, the experimenter monitored the stimuli over headphones, and paused the presentation briefly after every stimulus to allow the children to respond and to record this response. A bell indicated the end of each block of 10, at which point the children were allowed to choose a small prize (a sticker). The interstimulus interval for presentation to the adults was fixed at 3 seconds, and the inter–block interval 5 seconds.

2.3.2 Phonemic awareness

The phoneme blending task was always presented before the phoneme segmentation task. Before testing, the children were introduced to a puppet that “says every word all broken up into little bits.” During the phoneme blending test, the children were presented with a word as the puppet would say it—i.e. segmented into phonemic units—and were asked to guess what word the puppet had said—i.e. blend the segments into a single word. A correct response was simply the correct identification of the segmented word with all the phonemes present and in the correct order.

For the phoneme segmentation test, the children were asked to say a word “all broken up into little bits” as the puppet would say it—i.e. segmented into phonemes. A correct response was one in which all phonemes were segmented from each other, and were
presented accurately and in the correct order. If the child segmented the word into units larger than the target units (e.g. onset–rime, rather than phoneme) they were encouraged to “try to break the word up into even smaller bits.” Only phoneme sounds were accepted as answers: if responses were given as letter names, the child was encouraged to respond again, using sounds only. There was some flexibility in terms of what was considered an accurate response: diphthongs were allowed to be segmented as either one phoneme or two (e.g. “mouse” = [m], [au], [s] or [m], [a], [u], [s]); dialectal variation was allowed for (e.g. “train” = [t], [r], [e], [n] or [tʃ], [r], [e], [n]; both phonetic and phonological approaches were accepted where appropriate (e.g. “space” = [s], [p], [e], [s] or [s], [b], [e], [s]).

Both phonemic awareness tasks were introduced and explained by the experimenter. Testing in each case involved the experimenter asking the child to listen to the recorded stimuli, and to perform the required manipulation (e.g. the experimenter would say “Can you break this word up into little bits?” followed by the recorded voice saying “pig”).

Before testing, the children were introduced to the concept of the relevant metalinguistic manipulation for each task. This training period began with manipulation of syllable–sized units (e.g. “cowboy” = [kaʊ], [bɔɪ] and moved to gradually smaller units (i.e. onset–rime units followed by phonemes). The children received feedback on their performance throughout this training.

In order to continue to the test proper, the children were required to perform a pre–test relevant to the task. For both tasks, the manipulation required at this pre–test level involved onset–rime awareness only. If any child manipulated all five items in either one of the pre–tests incorrectly, testing in that task was discontinued. The child was not, however, eliminated from further testing (either in other tasks at that session, or in other sessions) and their data were not eliminated from analysis. The stimuli for
both tests were split into two balanced sets. The words in each set were presented in order of increasing number of phonemes, and thus increasing difficulty (McBride-Chang 1995b). During the test proper, if any child was unable to correctly manipulate 5 out of 6 consecutive stimuli, the test was discontinued. Again, the child was not eliminated from further testing, nor were their data eliminated from analysis.

3 Results

3.1 Cue weighting

The graphs in Figure 4 show the average perceptual response curves for the 8 adults, the 18 children at Time 1 and 2, and the 15 remaining children at Time 3.

In order to better analyse these responses, the data from the cue weighting tests were normalized using a probit transformation (see e.g. Liberman 1996, Nittouer & Studdert-Kennedy 1987, Nittouer 1992, Nittouer 1996b). This transform assumes data on an S–shaped curve, from which it extracts rate–of–change information (i.e. in this case, the rate at which a listener’s responses changed from mostly /ʃ/ to mostly /s/) and gives estimates of values which can be used to describe each set of response curves (Cohen & Cohen 1983). These values are the mean: the point along the fricative continuum at which the /s/ responses reach 50%—i.e. the phoneme boundary, and the slope, or degree of categoriality, of each individual response curve. From the first value we can also calculate the degree of separation of the response curves by taking the difference between the means for the continuum with /s/-transitions and the continuum with /ʃ/-transitions. This gives a measure of the extent to which the subject’s category boundaries were shifted as a result of the presence of the two different sets of formant...
transitions. Separation of the response curves can therefore be seen as representative of the **transitional effect**—i.e. the extent to which the subject attended to, or weighted, the transitional information.

Nittrouer (see e.g. Nittrouer 1992) proposes that both separation and slope of response curves are indicative of cue weighting strategy. However, while it is reasonably straightforward to see the connection between separation of response curves and the degree to which a listener made use of transitional information, the relationship between slope (categoriality) and the influence of either of the available cues is not so clear. It is in fact quite difficult to determine to what extent an increase in categoriality of a given response is due to the increased use of the cue that changes along the continuum, and to what extent it simply reflects an increase in ability to categorize in general. Additionally, as noted above, Hazan & Barrett (2000) found that cue weighting and categorical perception do not develop completely simultaneously, suggesting that these two processes might not be reflective of the exact same aspect of perception. As Nittrouer (1996b) examined separation (rather than slope of response curves) in relation to phonemic awareness, the main focus of examination in the current study was the relationship between those two measurements. Both separation and slope were, however, examined individually and in relation to each other and phonemic awareness.

From the graphs in Figure 4 it can be seen that the children’s response curves at Time 1 were shallower and more widely separated than those of the adults. It can also be seen that the children’s response curves became slightly less shallow and less widely separated at the subsequent two sessions (Time 2 and Time 3).

ANOVA with the perceptual measures of **slope** and **separation** as dependent variables, and **age** (that is, adult or child) as the independent variable, showed that there was a significant difference in both slope \(F(1, 24) = 4.25, p = .05\) and separation \(F(1, 24) = 6.42, p = .01\) between the adults and the children at Time 1. This difference was smaller...
but still significant at Time 2: slope \( F(1, 24) = 5.92, p = .02 \); separation \( F(1, 24) = 3.95, p = .05 \). By Time 3, there was no longer any significant difference in either slope or separation between the adults and the children.

Examining just the performance of the children, ANOVAs with the perceptual measures of slope and separation as dependent variables, and Time (that is, Time 1, 2, or 3) as the independent variable, showed that there was a significant change in separation of response curves between Time 1 and Time 3 \( F(2, 48) = 3.24, p = .04 \). There was no significant difference in slope of response curves across the three sessions. Examining separation measures session by session, it can be seen that there was no significant difference in separation of response curves between Time 1 and Time 2. The difference in separation approached significance between Time 2 and Time 3 \( F(1, 31) = 3.47, p = .07 \). This suggests that the majority of the change in cue weighting strategy took place in the latter half of the study.

ANOVAs with separation as the dependent variable, and day of testing, and Time as independent variables, showed a significant difference in perceptual behaviour across the three sessions in the study, but no significant difference in behaviour across the different days of testing. This indicates that changes in perceptual behaviour were significantly accounted for by differences from session to session, and not by any day–to–day variation in behaviour.

### 3.2 Phonemic awareness

[Table 1 about here]

The mean scores for the phonemic awareness tests across all three sessions are shown in Table 1. From this table it can be seen that there was a progressive increase in phonemic awareness ability across all three sessions of the study, with the largest increase in
phonemic awareness ability occurring between Time 1 and Time 2 for both tasks. Table 1 also shows that the two measures of phonemic awareness correlated very highly with each other within all sessions.

ANOVA with both phonemic awareness measures (blending and segmentation) as dependent variables and Time as the independent variable, confirm that there was a significant change in ability for both tasks between Time 1 and Time 3: phoneme blending \( F(2, 48) = 3.983, p = .025 \), phoneme segmentation \( F(2, 48) = 6.865, p = .002 \).

An examination of the change in phonemic awareness ability across sessions shows that there was a significant difference in ability on both measures between Time 1 and Time 2 (phoneme blending \( F(1, 34) = 4.077, p = .05 \), phoneme segmentation \( F(1, 34) = 5.762, p = .02 \)). However, there was no significant difference in ability between Time 2 and Time 3 for either of the measures, suggesting that the majority of phonemic awareness development took place before Time 2.

### 3.3 Relationship between cue weighting and phonemic awareness

The graphs in Figure 5 show the children’s acoustic cue weighting in terms of separation of response curves (x-axis) and phonemic awareness in terms of raw phoneme blending score (y-axis), at Time 1, 2 and 3.

These graphs are divided into quadrants at specific points on the x– and y–axis. For cue weighting, the division was made at the point that corresponds to the most global of the adult responses: 0.13kHz separation between response curves. The “analytical” side of the graph can therefore be considered to roughly cover the range of possible adult responses to the stimuli. The division for phonemic awareness was made at the median score for all children on the phoneme blending test at Time 1 of the study. The
choice of this second point was somewhat arbitrary, but ensured that the responses of the children at Time 1 were equally distributed on either side of the division. These two division points allowed for each child to be roughly categorized in terms of their performance on each measure: ‘poor’ or ‘good’ phonemic awareness, ‘global’ or ‘analytical’ cue weighting. Additionally, the data points on the graph (each representing an individual child) were given one of four symbols—open circle, filled circle, open triangle, filled triangle—each corresponding to the position of that data point at Time 1 of the study. The symbols allocated to each child at Time 1, along with the two points used to divide the graphs into quadrants, were maintained throughout all three sessions of the study so that changes in the children’s performance could be tracked in terms of their progression from these initial starting points.

A visual examination of the graphs in Figure 5 shows that at all three sessions there were children with both good phonemic awareness and analytical cue weighting strategies (top left quadrant), and at the first two sessions there were children with both poor phonemic awareness and global cue weighting strategies (bottom right quadrant). It can also be seen that there were a number of children with very good phonemic awareness who had very global cue weighting strategies (top right quadrant), but very few who had poor phonemic awareness and very analytical cue weighting strategies (bottom left quadrant), suggesting again that phonemic awareness develops before changes in cue weighting take place.

Examining the movement of data points from Time 1 through Time 3, it can be seen that the pattern observed at each individual session of the study was also the pattern seen throughout the study. Children who had both poor phonemic awareness and global speech perception strategies at Time 1 (filled triangles) developed better phonemic awareness before developing more analytical cue weighting strategies at later sessions of the study. Children who had already developed good phonemic awareness
at Time 1 (open circles) then went on to develop more analytical cue weighting strategies. Importantly, none of the children developed strongly analytical cue weighting strategies while still having very poor phonemic awareness skills.

Multiple regression analysis shows that 72% of the variability of separation measures at Time 3 was accounted for by a combination of the separation of response curves at the previous session (Time 2) and phoneme blending skill at both previous sessions (Time 1 and Time 2) \( R^2 = .72(F = 9.497), p = .002 \). Each of these measures made a unique contribution to the 72% of variability accounted for by all three: separation at Time 2 accounted for 16.75% (of the 72%) \( Beta^2 = .1675, p = .007 \); blending at Time 2 accounted for 34.65% \( Beta^2 = .3465, p = .01 \); blending at Time 1 accounted for 48.6% \( Beta^2 = .4860, p = .003 \).

Examining the relationship in the opposite direction, multiple regression analysis shows that none of the variability of either phoneme blending or phoneme segmentation measures at Time 3 was accounted for by measures of separation at previous sessions in the study. 81% of the variability of measures of blending at Time 3 was accounted for by blending at Time 2 \( R^2 = .81(F = 57.815), p < .0001 \). When measures of response curve separation at Time 1 and Time 2 were added, however, neither made a unique contribution to phoneme blending at Time 3. 83% of the variability in segmentation measures at Time 3 was accounted for by segmentation measures at Time 2 \( (R^2 = .83(F = 63.553), p < .0001) \). Measures of response curve separation, however, did not make any contribution to the variability of phoneme segmentation skill.

Interestingly, an examination of the relationship between the slope of the children’s response curves and the phonemic awareness measure of segmentation shows a different pattern. As noted above, 83% of the variability in segmentation measures at Time 3 was accounted for by segmentation measures at Time 2. If the slope of the response curves at Time 1 and Time 2 are added, 89% of the variation in segmentation measures
at Time 3 can be accounted for ($R^2 = .89(F = 30.48), p < .0001$), with slope making a unique (though small) contribution to the variability: segmentation at Time 2 accounts for 80.71% (of the 89%); slope at Time 2 accounts for 9.17% ($Beta^2 = .0917, p = .04$); slope at Time 3 accounts for 10.12% ($Beta^2 = .1011, p = .03$).

4 Discussion

The first aim of this study was to attempt to replicate the behaviour of two subjects from Nittrouer’s (1996b) study of cue weighting and phonemic awareness. The two children in that study were found to have developed good phonemic awareness skills while still displaying global cue weighting strategies, leading Nittrouer to suggest that the development of phonemic awareness might be a factor in changes in cue weighting. The present study demonstrates that the behaviour of the two subjects from Nittrouer’s study was not anomalous. The graph in Figure 5 shows that there were a number of subjects who began the present study with good phonemic awareness skills and very global cue weighting, while there were almost none with poor phonemic awareness and analytical cue weighting (and in fact none at all with poor phonemic awareness and extremely analytical cue weighting).

The second aim of the study was to determine to what extent children who display this pattern of results are representative of an intermediate stage of development between poor phonemic awareness/global cue weighting and good phonemic awareness/analytical cue weighting. The present study shows, first, that all of the children developed in terms of both their phonemic awareness ability and their cue weighting strategies over the course of the study. The latter is important to note, as this is the first time that the change from global to analytical cue weighting strategies has been observed longitudinally. More importantly, however, is the relative time scale
at which these two processes developed. Better phonemic awareness skills tended to develop before cue weighting strategies changed. That is, it appears that the relationship between phonemic awareness and cue weighting observed in the two children from Nittrouer’s study, and in the first session of the current study, hold throughout development. Closer examination backs this up: There was a significant increase in phonemic awareness ability between Time 1 and Time 2, but no significant difference in ability between Time 2 and Time 3. For cue weighting, although differences in strategy were significant only between Time 1 and Time 3, changes in strategy approached significance more closely between Time 2 and Time 3 than between Time 1 and Time 2.

It should be pointed out that despite a relatively short time lag between sessions, there were in fact a number of children who moved directly from the poor phonemic awareness/global cue weighting quadrant at one session to the good phonemic awareness/analytical cue weighting quadrant at the next session, without appearing to move through the global/good quadrant. This could suggest that the two processes develop simultaneously. It is important to note, however, that the division of the responses into quadrants was done in order to more easily follow the progress of the subjects across time. The division points themselves were based on the performance of the subjects, and are therefore in some way arbitrary. If one were to take the view that they are representative of critical stages in development, then one might expect to see movement of each child through all three quadrants (poor/global to good/global to good/analytical). As they are not critical developmental landmarks, however, what one should hope to see is each child first improving in terms of their phonemic awareness, followed by the same child changing to a more analytical cue weighting strategy, as compared to their own results at Time 1. In general this does seem to be the pattern, particularly in the case of those children who began with above average phonemic awareness skills and relatively analytical cue weighting (children represented as filled circles in Figure 5). Despite the fact that they began the study with scores that fell
within the “adult” range for both processes, most of these children tend to show improved phonemic awareness at Time 2 compared to Time 1, and more analytical cue weighting at Time 3 than at Time 2.

The fact that overall most improvements in phonemic awareness ability took place earlier on in the study, while changes in cue weighting strategies tended to take place later on in the study, suggests a possible direction of causality in the relationship. Specifically, it is likely that the earlier developing process—phonemic awareness—has some sort of influence on the later developing process—changes in cue weighting. Statistical analyses of the predictive relationship between phonemic awareness and cue weighting showed underline this possibility: 72% of the variance seen in the subjects’ cue weighting strategies at Time 3 can be accounted for by a combination of cue weighting strategies at the Time 2 and phoneme blending ability at both Time 1 and Time 2, with all three measures making a unique contribution to the variability.

It would therefore appear that the linguistic experience proposed by Nittrouer and colleagues as the driving force behind changes in cue weighting must include the development of phonemic awareness, and indeed possibly other metalinguistic processes. It is not clear at this stage whether phonemic awareness is a necessary prerequisite for changes in cue weighting—that is, whether some cue weighting changes would take place without the development of phonemic awareness. However, pilot studies of later reading (but normally developing) children reported in Mayo (2000) provide suggestive evidence that delaying phonemic awareness acquisition results in delayed cue weighting changes. This would suggest that phonemic awareness ability plays a relatively large role in a listener’s cue weighting strategies.

It should also be noted that the results of this study seem to underline the possibility that cue weighting and categorical perception tasks access different aspects of perception. Phonemic awareness ability was found in this study to predict cue weighting
strategies in speech perception, as measured by difference in response curves due to transitional context. However, the categoriality of listeners’ responses was found to predict a very small, but significant amount of the variation in phoneme segmentation ability. This finding supports those of McBride–Chang and colleagues (Manis et al. 1997, McBride-Chang 1995a, McBride-Chang 1996, McBride-Chang et al. 1997).

The fact that two different patterns of influence were found suggests that some changes in speech perception, such as a change in categoriality of responses might reflect a more general maturation of ability to categorize sounds, which would be necessary for the development of phonemic awareness. Other changes, such as the shifts in acoustic cue weighting described by the DWS, could reflect alterations in perceptual strategy due to pressure from processes external to the development of the perceptual system itself. Future perceptual studies should investigate this apparent division.
Acknowledgements

This study was supported by a doctoral studentship to the first author from Queen Margaret University College, Edinburgh. We would like to thank the head teachers, classroom teachers, parents and children who participated in this study for their enthusiasm and interest in the research. We also thank Alice Turk for her useful comments on earlier versions of this manuscript.
Figure and table captions

Figure 1: Stylized spectrograms of /so/ and /Jo/ syllables. Frequency of the fricative spectra are highlighted with diagonal lines, vowel onset formant transitions are highlighted in grey.

Figure 2: Stylized spectrograms of Nittrouer–style /so/–/Jo/ continua. The top continuum has /s/–transition vowels, the bottom continuum has /j/–transition vowels. The dashed lines represent idealized responses (presented in terms of % /s/ responses) for a listener whose perception is influenced by the transitions.

Figure 3: Spectrograms of four ‘endpoint’ stimuli: the two most extreme fricative noises (the top stimuli have the most /s/–like noise, the bottom have the most /j/–like noise) combined with each of the two transitional contexts (the left stimuli have /s/–transitions, the right two stimuli have /j/–transitions.)

Figure 4: Adults’ and children’s responses to /Jo/–/so/ stimuli with /s/–transitions (solid line) and /j/–transitions (dashed line). The x-axis shows the continua of fricative noises, ranging from 2.2kHz (most /j/–like) to 3.8kHz (most /s/–like).

Figure 5: Relationship between cue weighting and phonemic awareness for children. The vertical division is made at the most global of the adult cue weighting responses (0.13kHz separation). The horizontal division is made at the median blending score at Time 1 (26.5). The dots correspond to placement of each child’s data point at Time 1: filled triangles began the study with global cue weighting and poor phonemic awareness, open circles began with global/good, filled circles began with analytical/good, open triangles began with analytical/poor.

Table 1: Results of the two phonemic awareness tests, and correlation between the tests (given as raw scores out of a possible 50). For the correlation measures, $p < .001$. 

29
Figure 1:
Figure 2:
Figure 4:
Figure 5:
## Table 1:

<table>
<thead>
<tr>
<th>Time</th>
<th>Blending</th>
<th>Segmentation</th>
<th>Correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>28</td>
<td>25</td>
<td>$r = .8942$</td>
</tr>
<tr>
<td>2</td>
<td>39</td>
<td>36</td>
<td>$r = .7898$</td>
</tr>
<tr>
<td>3</td>
<td>41</td>
<td>42</td>
<td>$r = .8690$</td>
</tr>
</tbody>
</table>
References


36


