The Relationship between Phonemic Awareness and Cue Weighting in Speech Perception: Longitudinal and Cross–Sectional Child Studies

Catherine Mayo

A thesis submitted in fulfilment of requirements for the degree of Doctor of Philosophy

to

The Department of Speech and Language Sciences,
Queen Margaret University College, Edinburgh

February 2000

Declaration

I hereby declare that this thesis is of my own composition, and that it contains no material previously submitted for the award of any other degree. The work reported in this thesis has been executed by myself, except where due acknowledgement is made in the text.

Catherine Mayo

Abstract

Studies have shown a potential relationship between speech perception, and the development of both alphabetic literacy and phonemic awareness. Most of these studies have assumed that all perceptual changes are developmental, with literacy and phonemic awareness building on, but not affecting perception. However, a study by Nittrouer (1996*b*) of the relationship between phonemic awareness, and changes in perceptual cue weighting (also called the Developmental Weighting Shift) brings into question this assumption. Nittrouer showed that, while early perceptual development might affect later metaphonemic skills, it is equally possible that the development of phonemic awareness could impact on apparently developmental perceptual changes.

The studies in this thesis aimed to determine which of these two possible causal directions is more likely. Experiment 1 was a longitudinal study of 18 beginning–reading children (average age at the beginning of the study: 5;8). These children were tested on their phonemic awareness, and their perceptual weighting of two cues to a /ʃ/–/s/ contrast. This study showed that changes in acoustic cue weighting always follow the development of good phonemic awareness skills, and furthermore, that early phonemic awareness scores predict later cue weighting strategies. Experiment 2 was a cross–sectional study of 8 normally developing older children, who had not begun literacy or pre–literacy training (average age: 7;3). These children were also tested on their phonemic awareness skills and cue weighting strategies. This second study showed that in the absence of phonemic awareness development, changes in cue weighting strategy do not take place.

These two studies show that changes in cue weighting strategy are affected by the development of phonemic awareness. This indicates that Nittrouer's Developmental Weighting Shift model does not in fact describe a developmental process. These results also suggest that, when studying the development of low level linguistic behavior, the effect of higher order cognitive processes must be taken into account.

Acknowledgements

I would like to begin by thanking my three supervisors—Nigel Hewlett, Jim Scobbie, and Daphne Waters—for their input and advice. I know this has been a roller–coaster ride of a PhD, and I commend all three for having survived, with wits and patience intact! I owe a particular debt of gratitude to Jim, without whom I never would have known about this PhD, let alone started it (yes, Jim, it's all your own fault!), and without whose encouragement I probably would have packed it in numerous times. Thank you, Jim, for sharing your colourful (and slightly scatalogical) sense of humour, your even more colourful (but hopefully less scatalogical) socks, your friendship, and above all your boundless enthusiasm for research.

I should like to thank the schools, teachers, parents, and especially children who participated in this study, for kindly letting me take up so much of their time, and for all interest that they expressed in my work.

I also owe a great deal of thanks to a number of people in the Linguistics Department at the University of Edinburgh. First, to some fellow North Americans: Alice Turk, Ellen Bard, and Bob Ladd. My thanks to Alice, for helping me to get this thesis started by introducing me to Susan Nittrouer's work, and by encouraging me to ask the right questions about it. Thanks, too, for patiently reading and re–reading innumerable grant proposal drafts. A *very* big thank you to Ellen, for proposing the study that makes up Experiment 2—this thesis would not have been half of what it is without it! Thanks also for providing much needed statistical advice, and much welcome chats about (among other things) Edinburgh's singular lack of decent snow. My thanks also to Bob, for the lecturing and tutoring opportunities, and for the food–y reminders of home, from pancakes and brownies, to the ever–expanding Linguistics Thanksgiving Dinner!

I am also indebted to a number of other staff at Edinburgh, in particular Ethel, Irene, Cedric, Morag, Eddie, Mike and Stewart, all of whom kindly forgot that I ceased to be a University of Edinburgh student some time three years ago, and who not only let me continue to make use of equipment and facilities, but who offered a great deal of support (both technical and personal) along the way.

Over the course of the last 3 or 4 years, I have been subjected to strange, esoteric, and often bizarre ramblings from the Common Room Crowd, and have, in return, subjected them variously to my stress, crises, and general PhD anxieties, all of which they bore with amazing good humour! Thanks for lending ears Simon, Dave, Lynda, Anna, Dan and all the rest of ya. Thanks, too, to all the QMUC pre–and post–PhDs for commiserating at the right moments, and in particular to Jocelynne for fortifying me with Valvona & Crolla pastries and much encouragement. A very big, wiggly thank you to Anna, Mhairi and Lynne, for putting up with me, and just generally keeping me sane for the last 18 months.

I owe my family a great deal, probably more than I'm able to express here. Thank you, Mum and Dad, for being an unbelievable and unwavering source of encouragement and support. Thank you to Grama, for being such an inspiration, and thanks to Chris for always reminding me of the lighter side of things.

Rob, this is for you, because without you it never could have happened.

But maybe, my children, you're wondering whether
There's some common thread that ties all this together?
There is, but be patient a moment or two,
My story begins with a very strange cue.

Ignatius G. Mattingly 1985,
from **Lines for the Fiftieth Anniversary of the Founding of Haskins Laboratories**

Contents

List of Tables

List of Figures

CHAPTER 1

Introduction: Setting the stage

The communication of a linguistic message can be undertaken by means of a number of different media. Two of the most basic of these for the average communicator are speech and writing, with speech perception and reading as the means of understanding the message in each case. A fair amount of evidence has been collected that indicates that there is a relationship between certain aspects of these two methods of understanding. Recent studies have, for instance, found differences in perceptual performance between good and poor readers (e.g. Adlard & Hazan 1998, Mody, Studdert-Kennedy & Brady 1997, Nittrouer 1999, Werker & Tees 1987), and between alphabetic and logographic readers (e.g. de Gelder & Vroomen 1992). Further studies (e.g. McBride-Chang 1995*b*, Nittrouer 1996*b*) have also found specific correlations between aspects of perception and the development of awareness of phonemic segments—a skill which we will later see is highly related to alphabetic literacy.

At first glance it is perhaps unsurprising that these two processes should be found to be related. The *function* of both speech perception and reading is, after all, the same—specifically to understand a message that has been transmitted by means of some sort of encoding. Indeed the messages themselves are conventionally seen as having the same underlying format—a phonological representation of words and phrases—and it is often assumed that both speech perception and reading decode this representation in the form of a string of phonemes.

However, speech perception research has traditionally emphasised that the *processes* by which a phonological message is decoded in speech perception and in reading are very different. First, the *form* of the message itself is different for perception and reading. Liberman, Cooper, Shankweiler & Studdert-Kennedy

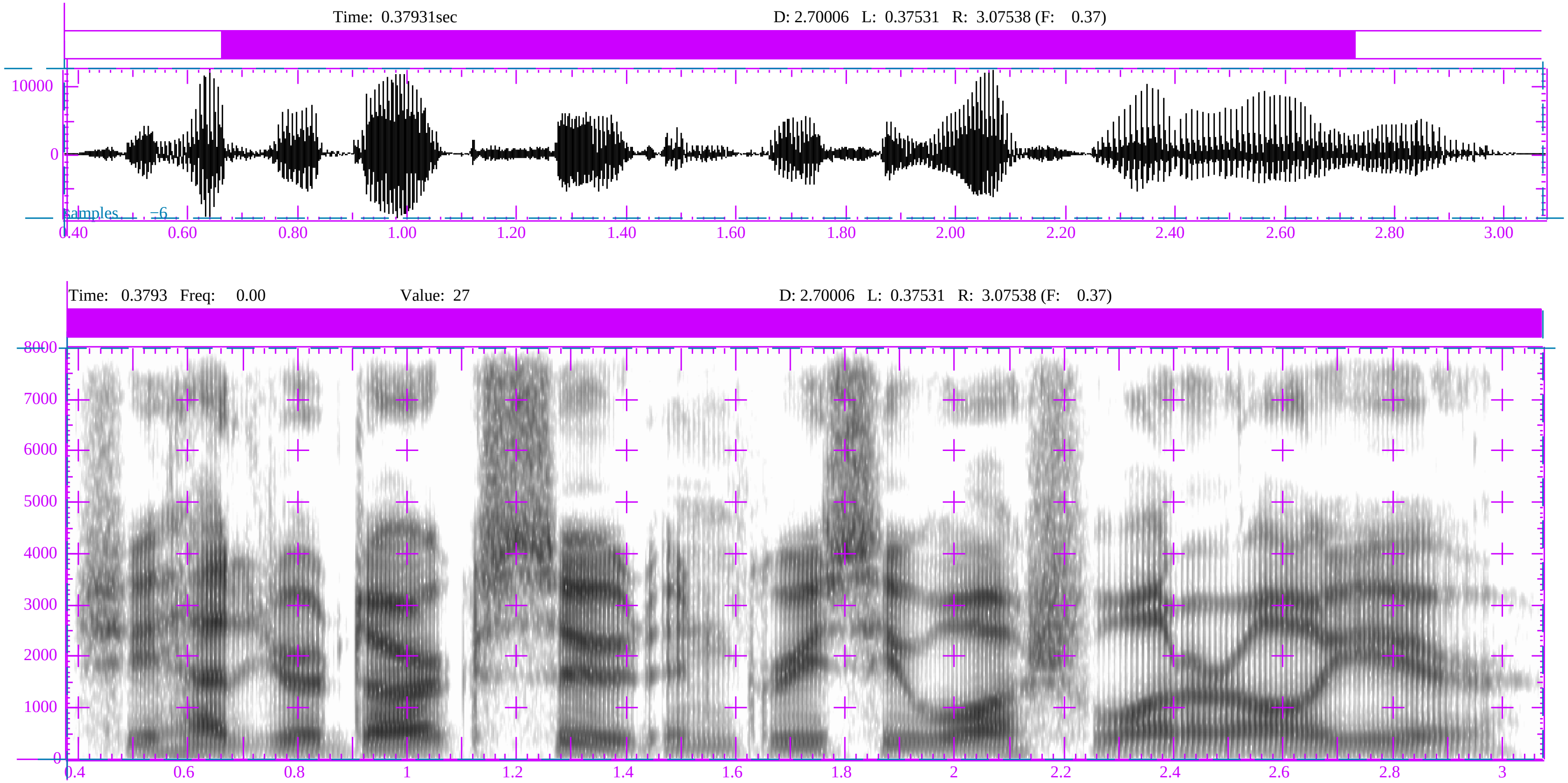

Figure 1.1: **Spectrogram of the utterance 'She had your dark suit in greasy wash water all year.'** Note, for example, that the section from approximately 2.25ms to the end of the utterance, which corresponds to 'water all year,' is virtually continuous, with little or no indication of possible word or segment boundaries.

(1967) describe this difference by saying that an alphabetic writing system transmits a phonological message by means of "a simple cipher" (p. 433), while a spoken message is transmitted by means of "a complex code" (p. 433). The terms 'cipher' and 'code' serve to underline the difference in the intricacy of the relationship between the message and the message medium in each case. In the case of writing, for those languages that make use of an alphabetic script, there is something approaching a one–to–one relationship between the minimal units of the underlying phonological message—phonemes—and the letters or letter combinations that represent them—graphemes. This may seem like an oversimplified description of the relationship if one examines a language like English, where a number of different graphemes may represent the same phoneme—/f/, for example, is represented by 'f' in 'fish', 'ph' in 'phone' and 'gh' in 'enough'—and one grapheme may represent a number of different phonemes—'o' is /ʊ/ in 'to', /ɒ/ in 'top' and /oʊ/ in 'no'. However, for many languages (e.g. German, Greek, Turkish) the relationship is much more transparent, and very much closer to one–to–one (see Oney & Goldman 1984, Porpodas 1989, Wimmer & Goswami 1994). In fact, in all languages that use an alphabet, the relationship between phonemes and graphemes is, at the very least, very much more regular than the relationship between the same phonemes and the aspects of the acoustic signal that represent them in speech.

The spectrogram in Figure 1.1 illustrates the first problem in understanding the relationship between the acoustic signal and the message that it is transmitting: specifically, that it is virtually impossible to neatly divide the signal into sections which correspond to individual phonemic segments. Additionally, the aspects of a signal that cue a particular segment are invariably multiple, some of these multiple cues overlap with aspects which cue other segments and may be shared by other segments, and all are affected by the context in which they are spoken. Thus in order to *read* a phonemic segment, a reader must be able to make the connection between a written grapheme and the phonemic segment it represents. In order to *perceive* the same segment in speech, on the other hand, a listener has to be able to reconcile this percept to the many, varied, and heterogeneous aspects of the speech signal.

This brings us to the second major difference between speech perception and reading: the *manner* in which the message is decoded. Listeners decoding a spoken message do not need to be, and in fact are not, aware of the complex process by which they are able to make a connection between a phonemic segment and the many and varying acoustic cues to that segment. Speech perception is a subconscious process: listeners simply understand the message. Reading, on the other hand, is a conscious process. In order to decode a written message, a reader must become consciously aware of the underlying phonological form of the message—specifically, they must become aware of phonemes—in order to be able to make a phoneme to grapheme connection.

Finally, as pointed out by A. Liberman (1996), there is a fundamental difference in *naturalness* between speech perception and word reading. All humans, given an environmental language, will learn to speak and to perceive speech—these subconscious processes are seen as essentially maturational. Writing and reading and their accompanying metalinguistic abilities (i.e. the ability to think consciously about the phonological form of a message), on the other hand, do need to be taught for the most part. Furthermore, the nature of these differences—the fact that speech perception is a more complex process than reading, and is subconscious and maturational, while reading is conscious and learned—are taken as support for the view that speech perception as a whole is the more fundamental of the two skills. Under this view, the development of alphabetic literacy, and all its component skills, builds on the phonological organisation laid down by the perceptual system (see e.g. Liberman 1996).

How does one reconcile these apparent differences in the *processes* by which messages are understood in speech perception and reading, with the studies (noted above) that have shown certain aspects of these same processes to be related? This reconciliation becomes particularly problematic when one begins to investigate the *nature* of the relationship between speech perception and reading processes, and in particular the possibility that the relationship could be causal. If the relationship was found to be causal, the direction of causality could either move from speech perception to literacy—that is, perception would have an effect on later literacy development—or from literacy to speech perception—that is, the development of literacy skills would have an impact on some aspect of speech perception. It is the possibility that causality might move in this second direction that is particularly problematic for conventional views of speech perception and literacy. Finding that literacy development impacts on perception would mean that we would have to entertain the idea that some aspect of speech perception could be affected by conscious, learned processes. We would thus have to consider that speech perception might not be the wholly subconscious, maturational and primary process that it has been assumed to be.

One particular study, by Nittrouer (1996*b*), does encourage us to consider this possibility. Nittrouer's study found a correlation between certain mechanisms which underlie perception and reading—specifically the use of acoustic cues in speech perception, and the development of conscious awareness of phonemes. The crucial aspect of this study is that it is inconclusive with respect to the possible causal direction between cue use in perception, and the development of phonemic awareness. At the very least, therefore, we must consider that both causal directions are possible.

The aim of this thesis is to determine which of the following two possibilities is more likely: i) that cue weighting has an impact on phonemic awareness, or ii) that phonemic awareness has an impact on cue weighting. It is also hoped that in exploring the relationship between these two processes, some new light will be shed on our understanding of the larger processes of which they are a part. This first chapter will therefore attempt to situate the studies in this thesis in the broader contexts of speech perception and metalinguistic awareness studies.

1 Speech perception

Some of the seminal work in speech perception was carried out under the auspices of speech synthesis research by Alvin Liberman and colleagues in the 1950s. These researchers experimented with using spectrograms to synthetically recreate the impression of speech. The system used in this work was a form of optical reader (eventually called a 'Pattern Playback', see Liberman 1996), which shone frequency modulated light through a spectrogram onto a phototube. The phototube then vibrated at the frequencies of the speech represented on the spectrogram. This work was first attempted with spectrograms derived from real speech. As can be seen in the spectrogram in Figure 1.1, a speech signal is made up of both fairly distinctive, gross–grained details—such as the high frequency noise at time 1.2ms—and less distinctive, fine–grained acoustic details. The Pattern Playback's vibrating phototube was able to reproduce all of these elements, creating sounds which very closely approximated real speech.

Importantly for speech perception research, however, Liberman and colleagues also found that they were able to achieve the same speech–like sounds with schematic copies of the spectrograms. In particular, it was found that a speech percept could be created with only a stylised representation of the more gross aspects of a spectrogram.

Having thus determined that not all of the fine–grained detail in the acoustic signal is necessary to create the percept of speech, these researchers then moved on to determine what parts of the signal *were* absolutely necessary for a listener to perceive a specific speech sound, and moreover, what parts of the signal were simply sufficient to signal a percept. What was found was that almost every distinguishable aspect of the acoustic signal is sufficient to signal or *cue* a speech percept. It was found, for instance, that a short burst of energy can be sufficient to cue a stop consonant, and that the frequency of this burst relative to the frequency of the formants of the following vowel allows the listener to differentiate between various stop places of articulation—e.g. between /p/, /t/, and /k/ (Liberman, Dellatre & Cooper 1952). A second possible cue to stop place of articulation was found to be the onset frequency and direction of movement of the post–consonantal vowel formants, in particular the second formant (Liberman, Dellatre, Cooper & Gerstman 1954). Differences in manner of articulation, such as that between the stop consonant /b/ and the glide /w/, were found to be cued by the duration of post–consonantal vowel formant transitions (Liberman, Dellatre, Gerstman & Cooper 1956), while differences in voicing were found to be cued by the timing of the onset of the first formant (later known as voice onset time or VOT, Liberman, Delattre & Cooper 1958). Thus speech perception came to be seen as a process of determining, and making use of, the relevant aspects of the acoustic signal to understand the phonological message.

1.1 What is an acoustic cue?

A. Liberman (1996) states that the term " 'cue' is a term of convention, useful for the purpose of referring to any piece of signal that has been found by experiment to have an effect on perception" (p. 22). Under this definition, therefore, we can see that the various configurations of bursts, transitions, and VOT which were found to signal different speech percepts in the experiments described above, can all be said to be acoustic cues.

Unfortunately, while this term is convenient and will be used in this way throughout this thesis, it is also slightly misleading. First, this definition suggests that a cue is a discrete, definable entity: that a burst is a cue to the presence of a stop, while a lack of burst signals a lack of stop, for instance. This, however, fails to capture the potential gradation of cues. The articulators (lips, tongue, teeth, jaw, vocal folds, etc.) that are used to create speech do not simply vary in binary configurations—i.e. mouth open or closed, vocal folds vibrating or not vibrating—but are free to vary in almost all configurations between these binary settings. The speech that results from the movement of these articulators thus varies in as many different configurations, leaving the speech perception system to reconcile a highly variable acoustic signal with a limited phonological lexicon. Liberman (1996) notes that there is a great deal of "data now available that indicate how exquisitely sensitive the listener is to *all* the acoustic consequences of phonetically significant gestures" (p. 22). What this data means, according to Liberman, is that "any definition of an acoustic cue is always to some extent arbitrary" (p. 22). It is possibly more clear, then, to think of a cue less as a discrete aspect of the acoustic signal, and more as an acoustic variable with a potential function—that is, a short burst of energy is a potential cue to the presence of a stop consonant.

More importantly, the use of the term cue to refer to discrete aspects of the signal wrongly suggests some sort of one–to–one relationship between an aspect of the speech signal and a single percept. As noted in the previous section, this is not actually the case. Repp, Liberman, Eccardt & Pesetsky (1978) point out that "In the articulation of an intervocalic stop consonant, for example, the characteristically rapid closing and opening of the vocal tract has acoustic consequences that include, among others, the following: various rising and falling transitions of the several formants; a period of significantly reduced sound intensity; and then a second, acoustically different set of formant transitions, plus (in the case of voiceless stops in iambic stress patterns) a transient burst of sound, a delayed onset of the first formant, and for the duration of that delay, band–limited noise in place of periodic sound in the higher formants" (p. 621).

In addition to this one–to–many relationship between a speech percept and the aspects of the signal that cue that percept, there is also a many–to–one relationship between percepts and cues. Because the same articulators are used to simultaneously convey other aspects of speech than simply the identity of phonemes, it is often the case that the multiple "acoustic consequences" of the movement of the articulators referred to by Repp et al. (1978) above might additionally have to serve as cues to suprasegmental, or other, aspects of speech. Fowler & Rosenblum (1991) note, for example, that there are multiple influences on the frequency of a speaker's pitch or F0, each of which cues a different percept. F0 is essentially controlled by the rate at which the vocal folds open and close. This will be locally increased or decreased by the speaker to convey lexical and phrasal pitch, but can also be increased and decreased by the intrinsic pitch differences between vowels, by the voicing (or lack thereof) of an obstruent, and simply by the deflation of the speaker's lungs as he breathes out. This means that as well as cuing differences in lexical or phrasal meaning, F0 cues vowel height, obstruent voicing, and the placement of elements in an utterance relative to the beginning of the utterance (i.e. relative to the point at which the speaker began to breathe out). In addition to this, because F0 is affected by the size and shape of the speaker's vocal tract, pitch range can also cue the sex, age and possibly the size of the speaker (e.g. Laver & Trudgill 1979).

This last characteristic is true not just for pitch, but for all acoustic cues: the size of the speaker and the nature of their vocal tract, as well as the language and dialect that they speak, affect the way in which all aspects of the speech stream are formed. This means that each acoustic cue will be different depending on the speaker. Moreover, as noted by Jusczyk (1997) "the production of any speech sound requires the coordination of many different components. As with any complex motor skill, it is virtually impossible to produce the same speech sound in the same way on two different occasions" (p. 47). Therefore even very small changes in a speaker's rate of speech, loudness, speech register, etc. will have an effect on the speech stream. As a result, cues do not only vary between speakers, but also between two productions of the same utterance by the same speaker. Thus, while a speech percept must be the same each time it is heard, the aspects of the signal that cue that percept will never be invariable in a straightforward way[1].

1.2 Coarticulation

One further aspect of the articulation of speech which has proved particularly problematic for speech perception research is a phenomenon called *coarticulation*. In effect, coarticulation refers to the articulation of aspects of more than one segment at the same time. It arises because the articulation of speech is both rapid and continuous—that is, the speaker does not slow down or stop the articulatory process at segment boundaries, or even between most words and phrases.

Coarticulation is problematic at both a practical and a theoretical level. From a practical point of view, coarticulation presents yet more variability in the speech

[1] Although note that some researchers do posit more complex invariants in the speech stream, as discussed in Section 2.1, this chapter.

stream that the perceptual system must cope with. For example, the configuration of the formants at the onset of the vowel /a/ will be different depending on whether the vowel follows /b/, /d/ or /g/ (see e.g. Delattre, Liberman & Cooper 1955). In fact, the acoustic cues that correspond to any and all segments are affected by the context in which these segments are spoken.

Theoretically, coarticulation poses a slightly different problem—namely how exactly one should characterise it. Jusczyk (1997) states that

> Coarticulation occurs because speech production involves moving our articulators...from one configuration to another in a very short time span. Because it takes time to move the articulators into the proper position for each sound, the articulatory apparatus is forced to find a compromise solution that involves starting the articulatory gestures for one segment prior to finishing the gestures relevant to a preceding segment. This causes the segments to overlap as speech is produced. For this reason a given slice of the speech wave includes information about the articulation of several different sounds in an utterance. (p. 5)

This definition, based on the concept that segments overlap in speech, is in the middle ground between definitions which see coarticulation as the 'assimilation' of certain characteristics of segments (e.g. Daniloff & Hammarberg 1973) and definitions which state that coarticulation is "gestural layering—a temporally staggered realization of gestures that sometimes do and sometimes do not share one or more articulators" (Fowler & Rosenblum 1991, p. 47, see also Öhman 1966).

From the point of view of acoustic cues, the results of coarticulation are usually defined as transitional cues—that is, aspects of the signal which are 'in transition' between one configuration and the next. However, the definitions of 'transition' are as varied as the characterisations of coarticulation described above. At one end of the scale is a model in which transitions are simply the dynamic sections of the signal that join together supposedly segment–intrinsic cues—the so–called 'steady–state' cues, such as fricative noise and vowel target formant frequencies. At the opposite end of the scale are theories which state that the effect of two segments on each other can be seen throughout both the segments, effectively subsuming any 'transitional' cues into the acoustic information for the segments themselves (e.g. Fowler & Rosenblum 1991, see also Section 2.1, this chapter).

It is fairly clear from the spectrogram in Figure 1.1 that the strict division of an acoustic signal into 'steady–state' cues and 'transitional' cues would be nearly as difficult as the division of the signal into individual phonemes—the signal is instead best described as almost constantly changing. Additionally, the division of acoustic cues into those which are intrinsic to a segment and those which only give information about the effect of coarticulation on that segment, ignores the numerous studies that have found that transitional information (in particular the onset frequency and movement of vowel transitions) is crucial to the perception of a number of phonemic segments (e.g. Delattre et al. 1955, Liberman et al. 1952).

As will be seen throughout this thesis, transitions can be seen as potential cues to phonemic segments in their own right, in much the same way as release bursts and VOT. However, as will also be seen, there may be some fundamental differences in the way that transitional cues and non–transitional cues are used by the perceptual system in certain situations.

1.3 How do listeners make use of acoustic cues?

Keeping in mind the complex relationship between acoustic cues and the percepts that they engender, the question that then arises is how exactly listeners make use of cues to arrive at a percept. Determining the answer to this question has been (and still is) the goal of a great deal of speech perception research. This research has uncovered a number of perceptual phenomena, such as categorical perception, cue trading relations, and cue weighting, among others, all of which may operate to enable the perceptual system to cope with the multiplicity and variability of acoustic cues.

Categorical perception

Categorical perception has to do with the way in which listeners organise acoustic cues into a finite number of perceptual categories. As noted above, speech articulators do not vary in binary configurations. The variation in the acoustic signal resulting from the movement of these articulators is therefore also not binary. Just as the tongue can touch the roof of the mouth in virtually any spot from the teeth to the soft palate, for example, it is possible for any one acoustic cue to vary in the same way from one extreme configuration to another: e.g. vowel onset formant transitions could potentially vary gradually from low rising to high falling, and VOT could potentially vary from long to short. However, in terms

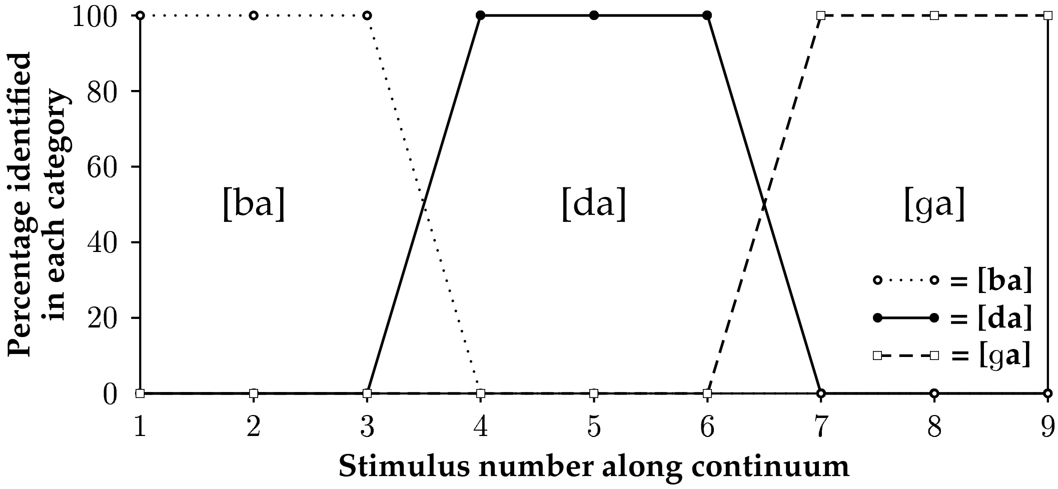

Figure 1.2: **Idealised graph of identification responses illustrating the phenomenon of categorical perception**. This graph represents responses to stimuli along a formant transition continuum from [ba] through [da] to [ɡa]. Each line on the graph represents the percentage of the stimuli at each point on the continuum that would be given a particular category label. Note that there is very consistent labelling within each category, and a very sharp change in labelling at each category boundary.

of speech understanding, it would be very inefficient for a perceptual system to automatically differentiate between all possible variants in formant configurations or VOT. Instead, what is needed from a perceptual system is an ability to determine which changes in the configuration of an acoustic cue signal a meaningful contrast—e.g. a change in formant transition configuration which signals the difference between [ba] and [da]—and the ability to 'ignore', or treat as irrelevant, a change which ought not to signal a meaningful contrast—e.g. a change in formant configuration which signals two different productions of [ba] spoken in two different segmental contexts. Research has shown that, for a large number of cues to certain phonetic contrasts, this is precisely how the perceptual system operates. If listeners are asked to identify speech sounds which are synthetically designed to vary along a gradual continuum, they do not give gradually less of one label and gradually more of another to these sounds. Instead, for a continuum of consonant–vowel (CV) syllables which vary, for example, in formant configurations from those appropriate for [ba] to those appropriate for [da] and then for [ɡa], there is a sharp change in labels from [ba] to [da], and [da] to [ɡa] at certain points on the continuum, and very consistent labelling of the syllables in between these changeover points (see Figure 1.2)[2]. These areas of consistent labelling are said to correspond to phoneme categories, with the changeover points corresponding to phoneme category boundaries.

In addition, listeners' discrimination abilities seem to correspond to their labelling tendencies. Asked to discriminate between stimuli from the same [ba]–[da]–[ɡa] continuum as described above, listeners perform at chance when attempting to tell the difference between two stimuli within the same category. Their ability to tell the difference between stimuli which fall on either side of a category boundary, on the other hand, is significantly better, even if the *absolute* acoustic differences between the stimuli from the same category and between those from different categories are identical. In other words, unless the difference between two stimuli is considered by the perceptual system to be meaningful, listeners treat the stimuli as equivalent, both in terms of phonetic labeling and in terms of discrimination (Liberman, Harris, Hoffman & Griffith 1957, Liberman, Harris, Eimas, Lisker & Bastian 1961).

This last point is interesting, given that there have been a number of suggestions put forward that labeling and discrimination tasks place different linguistic demands on a subject. Specifically, labelling tasks are often seen as requiring the subjects to make a linguistic, phonological evaluation of the stimuli, while discrimination tasks may not require this type of evaluation (e.g. Simon & Fourcin 1978). However, while discrimination tasks may not require a linguistic level of processing, they are often seen as allowing for a more stringent assessment of auditory sensitivity to small acoustic differences in speech (e.g. Sussman 1993). The implications of these possible differences on theories of perceptual development, and on the design of the current studies, will be discussed in Chapters 2 (Section 1.2) and 3 (Section 2.1).

It should be noted at this point that categorical perception is possibly one of the most widely studied perceptual phenomena (Rosen & Howell 1987). A wide range of listeners (including normal adults, normally developing children, clinical subjects, and non–human animals) have been tested on their tendency to categorically perceive numerous types of cues to different speech contrasts, as well as cues to non–speech contrasts. As a result of this extensive testing, it is often assumed that categorical perception is not simply a remarkable function of

[2]Although note that not all speech contrasts engender perception responses which are as strictly categorical as those shown in Figure 1.2: the perception of vowels, for example, has been shown to be more continuous than perception of stop consonants (see Rosen & Howell 1987).

a speech perception system, but rather is synonymous with speech perception as a whole. This in turn has led to categorical perception testing being used as an indicator of general perceptual ability for a number of studies, as opposed to simply an indicator of ability to class sounds into phonological categories.

Trading relations and perceptual cue equivalence

As noted above, there is a many–to–many relationship between acoustic cues and speech percepts. Having shown how listeners might cope with variation in single cues to a percept—i.e. they tend to perceive categorically—we can now ask how listeners cope with multiple cues to a single percept. The phenomenon of categorical perception has proved useful in allowing researchers to determine a possible answer to this question. Research has found that when multiple, rather than single cues are presented to a perceptual system, these cues interact and influence each other to form a percept.

One study of this phenomenon was carried out by Fitch, Halwes, Erickson & Liberman (1980), who examined the relative influence of two cues to the presence of post–fricative stop–consonants (as in the words 'slit' and 'split') on listeners' categorical perception. The two cues in question were i) duration of silence between the fricative noise and the onset of the liquid, and ii) presence or absence of vocalic onset transitions appropriate for bilabial stop closure. The authors were able to determine the relative effect of these two cues on the listeners' perceptions, by designing a categorical perception–type test in which the two aspects of the signal being manipulated did not always cue the same percept. A continuum of silence durations was created, ranging from a long silence duration, which clearly cued the presence of a stop, to a short silence, which did not cue stop presence. Then, two different vocalic onset transition conditions were created, one which was more likely to cue the presence of a stop, and one which was less likely to do so. These two transition conditions were combined with each of the silence durations, meaning that for each point on the silence continuum there were actually two different test stimuli. The listeners were presented with these stimuli, and asked to label them as 'slit' or 'split'.

What was found was that the two cues interacted in what has been called a *trading relation*. Overall, in both transition conditions, a longer amount of silence cued a stop consonant percept, and a shorter amount cued the percept of no consonant. However, for each different transition condition, the amount of silence needed to change the percept was different. When the formant transitions were appropriate to cue a stop consonant, only approximately 55ms of silence was required for a stop consonant to be perceived. When no formant transitions were present, approximately 80ms of silence was needed before a stop consonant was perceived. This means that when only one cue, for instance the silence duration, is available to signal a contrast, the effect of varying that cue (i.e. making the silence duration longer or shorter) is to change the percept. When a second cue to the contrast is available, however, it interacts with the first cue, and 'carries some of the load' of creating a percept, therefore less of the first cue is needed. This is the origin of the term *trading relation*—the two cues are able to 'trade' in the amount they are needed perceptually.

In the same study, Fitch et al. (1980) also showed the *perceptual equivalence* of cues to a contrast. The cues and contrast were the same as those described above: duration of silence and configuration of transitions as cues to the presence of a stop following [s]. Keeping in mind that both a long silence duration and the presence of transitions in the vocalic portion of the syllable cue the *presence* of a stop, the authors created stimuli in which these two cues either i) co–operated—that is, both cued a stop (a long silence plus transitions) or both cued the absence of a stop (a short silence with no transitions)—or ii) were in conflict—that is, one cued a stop and the other cued its absence (stimuli with long silence plus no transitions, or a short silence plus transitions). The results showed that those stimuli with both cues signalling a stop were easily discriminated from those with both cues signalling the absence of a stop. However, those stimuli with only one cue signalling the presence of a stop were much less easily discriminated from those stimuli with only the other cue signalling a stop. This lack of discriminability is the same result as is found in conventional categorical perception testing: as noted above, listeners are unable to discriminate between two configurations of the same cue when the two configurations signal the same percept (that is, when the two stimuli are from the same category). This led Fitch et al. (1980) to conclude that, in the same way that the perceptual system is able to treat two slightly different configurations of the *same* cue as perceptually equivalent, the perceptual system seems able to treat two *different* cues to the same contrast as perceptually equivalent. This means that despite the obvious acoustic differences between the silence duration cue and the formant transition cue (as noted by Fitch et al. (1980) "one of them is silence, the other is sound" p. 349), the presence of either one is sufficient to arrive at the same percept.

Acoustic cue weighting

Numerous other studies have been carried out which have shown trading relations between, and spectral equivalence of, cues to numerous other phonemic segments (e.g. Best, Morrongiello & Robson 1981, Repp et al. 1978, Stevens & Klatt 1974). However, a number of other studies have also found that two cues to the same percept are not necessarily always equivalent in terms of the *relative role* that they play when both are present to cue a percept. That is, in arriving at a percept listeners may not make equal use of all of the cues available in all situations.

Dorman, Studdert-Kennedy & Raphael (1977) conducted a study to determine the role of stop bursts and formant transitions in the perception of place of articulation of voiced stops. By cross–splicing natural syllables recorded by two different speakers, these authors created CVC stimuli with 5 different combinations of bursts, aspiration and vowel onset transitions: i) stimuli with all three cues, ii) stimuli with no aspiration or transitions, iii) stimuli with no transitions, iv) stimuli with no burst or aspiration, and v) stimuli with no burst. The results of this study showed that the degree to which the listeners made use of, or weighted, the burst and transition information—indicated by the degree to which they were affected by the absence of each cue—differed depending on the place of articulation of the consonant, the quality of the following vowel, and the speaker. Similar results were found by Whalen (1991), who showed that the perceptual weight given to cues to fricative place of articulation depends on the identity of both the fricative and the following vowel.

Walley & Carrell (1983) also examined the relative importance of different cues to listeners' percepts of stop place of articulation. This study made use of stimuli in which burst spectrum cues either agreed with or conflicted with (i.e. didn't cue the same percept as) the vowel formant transition cues. The results of this study showed that when the two cues to the percept of a particular place of articulation were in conflict, adults' identification of the stimuli generally corresponded to the percept cued by the formant transitions. Further studies (e.g. Ohde & Haley 1997, Wardrip-Fruin 1982, Wardrip-Fruin 1985) have gone on to show similar differences in the status or weighting of other cues to different percepts.

It should be noted that this apparent division of cues into those that are weighted more heavily by listeners, and those that are weighted less heavily, should not be confused with those theories that divide the acoustic signal into 'primary' and ‘secondary’ percepts (e.g. Stevens & Blumstein 1981). The main premise of these theories (which will be discussed in more detail in Section 2.1, this chapter) is that invariant properties, which it is claimed can be found at specific points in the acoustic signal, are the main cues to speech percepts, and all other, context–dependent aspects of the signal are only secondary cues. It is clear that this idea of cue use does not correspond well with the results of Walley & Carrell’s (1983) study described above, which showed that adults give more perceptual weight to the context–dependent transitional cues. Nor does it correspond well with the studies of Dorman et al. (1977) and Whalen (1991), who showed relative cue use to be dependent on a number of factors, including the contrast being perceived, and the nature of the cues themselves. All of these studies instead indicate that the perceptual system is free to make use of whichever acoustic cues are appropriate for any given percept.

Further evidence of the flexibility of cue weighting in speech perception comes from a number of studies which have induced listeners to change the cue that they weight more heavily. Some of these studies used some form of signal manipulation—either background noise (Wardrip-Fruin 1982, Wardrip-Fruin 1985) or reverberation (Watson 1997)—to mask certain cues in the signal. These studies found that listeners will weight cues differently when the stimuli are masked than they do when perceiving in silence—that is, the listeners will switch the weight given to certain cues if this is necessary for accurate perception (although see Chapter 2, Section 1.3 for a discussion of the same type of study with children).

More surprising are studies which found that listeners’ typical cue weighting behaviour can be shifted without resorting to masking parts of the signal. Gordon, Eberhardt & Rueckl (1993) found that when listeners were able to give their full attention to a phoneme identification task, they gave most weight to VOT when labeling a /ba/–/pa/ contrast, and gave most weight to formant patterns when labeling an /i/ to /ɪ/ contrast. However, when the same listeners were distracted from the perception task by attempting to complete an arithmetic calculation at the same time, the relative importance of the cues to the percepts changed: the listeners used F0 onset frequency more heavily to identify the /ba/–/pa/ contrast, and used vowel duration more heavily to identify the /i/–/ɪ/ contrast. A second study, carried out by Christensen & Humes (1997), made use of non–speech correlates of acoustic cues found in speech: specifically, frication noise, sloping frequency transition and a silent gap. Listeners were trained to label stimuli with various configurations of these three cues as 'circle', 'triangle' and 'square.' While an identification experiment found that untrained subjects tended to give the most weight to the frequency transition, these authors found that they were able to train the same subjects to classify these stimuli according to a different cue, specifically the duration of the silent gap.

Clearly, therefore, relative weighting of acoustic cues is dictated both by the demands of the perceptual system to make use of the most informative cue in any given context, and by demands external to the perceptual system itself.

Development of acoustic cue weighting

The apparent ability of the perceptual system to shift the weighting or importance of cues as required has also been found to be important developmentally. A number of studies have shown that young children do not make the same use of acoustic cues as do older children and adults in perception (e.g. Greenlee 1980, Krause 1982, Morrongiello, Robson, Best & Clifton 1984, Nittrouer & Studdert-Kennedy 1987, Nittrouer 1992, Ohde & Haley 1997). Morrongiello, Best and colleagues (Best et al. 1981, Morrongiello et al. 1984), for example, found that children weight transitions more than do adults in their labeling of a 'say–stay' contrast with varying gap duration and vowel onset transitions. Ohde & Haley (1997) also found that very young children (3–4 years) make more use of formant transitions than do older children and adults—in this case in the identification of stop consonants (all of these studies will be discussed further in Chapter 2, Section 1.3).

One particular set of studies by Nittrouer and colleagues (Nittrouer & Studdert-Kennedy 1987, Nittrouer 1992, Nittrouer 1996*b*, Nittrouer & Crowther 1998) indicates, according to the authors, a possible explanation for these differences in cue weighting between adults and children. Nittrouer & Studdert-Kennedy (1987) examined adults' and children's weighting of two cues to fricative place of articulation in the fricative–vowel syllables 'sue', 'shoe', 'sea', and 'she': i) frequency of fricative noise, and ii) vowel onset transition configuration. The study found that adults tended to be influenced only a relatively small amount by the transitional cues, basing their labelling decision more on whether the fricative noise was closer in frequency to the /s/ or /ʃ/ end of the continuum. The children, however, were much more influenced by the transitional cues, basing their labelling decision more heavily on whether the transition was appropriate for having followed /s/ or /ʃ/. Additionally there appeared to be a developmental effect: the youngest children (aged 3 years) showed the strongest effect of the transitions, the 5–year–olds showed a weaker transitional effect, and the 7–year–olds were very similar to the adults in their cue weighting. Further studies have gone on to replicate these developmental cue weighting results: for the same fricative contrast (/s/ vs. /ʃ/) in different vowel contexts, for different consonant contrasts (presence or absence of a stop consonant, as in 'say' vs. 'stay'), for various combinations of synthetic and natural speech, and for different dialects of English (Nittrouer, Manning & Meyer 1993, Nittrouer 1996*b*, Nittrouer & Miller 1997*a*, Nittrouer & Miller 1997*b*, Watson 1997).

However, one particular experiment carried out by Nittrouer (1992) showed that in a certain context young children give *less* weight than do adults to transitional cues. It had been shown in studies previous to Nittrouer's (1992) study that adults' identification of a CV syllable as either /da/ or /ga/ is influenced by the transitional effects of a preceding VC syllable: specifically, whether the syllable is either /al/ or /ar/ (see Mann 1980). Nittrouer, however, found that children are much less influenced by transitions in identifying this contrast.

At first glance this incongruous result appears to indicate that children may simply process transitional information differently depending on the *context of the transition*—i.e. more weight is given to transitions between fricatives and vowels than transitions between a stop consonant and a glide. Nittrouer, however, reconciles these two results by suggesting instead that children process transitional information differently depending on the *position of the transition* in relation to the overall syllabic structure of the stimuli—i.e. within–syllable transitions are given more weight by children than across–syllable transitions (Nittrouer 1992).

Nittrouer's hypothesis is founded on evidence from studies of phonological development in both perception and production that has suggested that young children's speech processing is focussed on units roughly the size of a syllable (e.g. Studdert-Kennedy 1987). Under this account, children do not need to give much weight to across–syllable transitions, as these occur outside the boundary of their perceptual units. Children will, however, give more weight to those transitions that occur within a syllable "because these properties help them recognize syllable structure in the speech stream" (Nittrouer & Miller 1997*b*, p. 2254). The same phonological development studies also suggest that there is some movement away from syllable organisation, towards organisation around something more akin to a phonemic segment, as the child ages. Therefore the reason for adults' different perceptual behaviour with regard to transitions, according to Nittrouer, is that they have "ceased to favor properties that specify movement and instead [have] come to emphasize those properties...that are most informative about place and shape of vocal–tract constrictions" (Nittrouer & Crowther 1998, p. 810), i.e. those cues which more closely relate to sub–syllabic segments. This model of the development of perceptual strategies—in which child listeners weight syllable–intrinsic cues more heavily and adults weight segment–intrinsic cues more heavily—has been termed the Developmental Weighting Shift by Nittrouer and colleagues (e.g. Nittrouer et al. 1993).

2 Changing units in phonological organisation

Nittrouer's theory that perceptual development moves from being syllable–based in childhood to phoneme–based in adulthood raises a number of questions. First, is there any other evidence that such a shift might take place in perceptual development? Second, what is the evidence that such shifts in units take place in any other area of phonological organisation? And finally, is there any evidence to suggest that changes in units at all of these different levels might be at all connected, and thus might reflect a more general reorganisation of phonological units? This section will attempt to give a very general overview of what is already known (or hypothesised) regarding each of these issues.

2.1 How do adults perceive? Theories:

We begin our investigation at what is assumed to be the end point of development—perceptual behaviour in the average adult listener. As will be seen, although there is a great deal of disagreement as the the exact *manner* in which adults perceive speech, there is actually a reasonable consensus as regards the *units* with which they perceive it.

In general, perceptual theories can be divided into two groups: the articulatory or motor group (e.g. Fowler 1986*a*, Fowler & Rosenblum 1991, Liberman et al. 1967, Liberman & Mattingly 1985), and the acoustic group (e.g. Blumstein & Stevens 1980, Kingston & Diehl 1995, Stevens 1980, Stevens & Blumstein 1981). For the strict motor theorist, "the objects of speech perception are the intended phonetic gestures of the speaker" (Liberman & Mattingly 1985, p. 2). A gesture, in this theory, is related to articulatory movements: a bilabial nasal [m], for example, consists at a fundamental level of a labial stop gesture and a velum lowering gesture. Gestures themselves generally involve the movement of two or more articulators, "thus 'lip rounding', for example, is a collaboration of lower lip, upper lip and jaw" (Liberman & Mattingly 1985, p. 22). However, because the movement of these articulators may be perturbed by coarticulation, rate of speech, speaker characteristics etc. (as described above), what is perceived cannot be the *actual*, highly variable movement or gesture. Instead what is perceived is the *intended* gesture. According to the motor theory, listeners perceive these intended gestures by means of a specific speech module, which compares the input signal with potential descriptions of that signal. The module is constrained in the number of descriptions that it produces by the fact that it derives them "by an analogue of the production process" (Liberman & Mattingly 1985, p. 26)—that is, the module "guess[es] how the [incoming] signal might have been produced" with reference to information about the physical characteristics of the vocal tract, (Fowler & Rosenblum 1991, p. 39).

In terms of the *units* of perceptual organisation, early versions of the motor theory (Liberman et al. 1967) specifically spelled out that the theory was intended to account for the part of the perceptual system that "lies between the acoustic stream and a level of perception corresponding roughly to the phoneme." (p. 431). In a revised version of the theory (Liberman & Mattingly 1985), the term 'phoneme' was replaced by 'phonetic category' or alternatively 'phonetic unit.' Phonetic category is used presumably to account for the fact that in the revision of the motor theory, gestures are taken to be more directly related to groups of features than to phonemes. It also takes into account contemporary changes in phonological theory, in particular the introduction of non–linear phonology. The term phonetic unit seems to have been chosen to avoid having to specify the size of unit (other than a gesture) into which the speech stream might be parsed by the listener. However, it is reasonably clear that the authors are at some level maintaining the concept of perception as a process of determining the sequence of phonemes in an utterance. This is reflected in the fact that the gestures that they claim are perceived by the listener are said to pattern in groups to produce phonetic segments. A specific example of this can be seen in the suggestion that there may be parts of an utterance which will contain information about only one 'phonetic unit': Liberman & Mattingly (1985) go on to state that such a part could be "the middle of the frication in a slowly articulated fricative–vowel syllable, and in vowels that are sustained for artificially long times," (p. 13) both of which (the fricative and the vowel) are single phonemes.

A slightly modified version of the motor theory, the direct–realist theory (Fowler 1986*a*, Fowler 1986*b*, Fowler & Rosenblum 1991), also makes the claim that listeners perceive in terms phonetic gestures. In this theory, the claim is made that invariants do actually exist in articulation. Pardo & Fowler (1997, p. 1142) note that "coordinative relations" among articulators mean that gestures are formed "flexibly and equifinally." By this they mean that the amount that each articulator is used in a gesture will vary depending on the coarticulatory demands being made on the articulator (flexibility), and that despite this variability, the "coarse grained gestural goal" (Fowler 1994, p. 608) will be reached (equifinality). The example given by Pardo & Fowler (1997, p. 1142) is that of bilabial closure: "the jaw will contribute less and the lips correspondingly more, to closure during /ba/ than during /bi/" due to the influence of the vowel, but in either case the gesture 'bilabial closure' is invariantly achieved. Additionally, unlike motor theorists, direct realists propose that perception of these gestures can be achieved without recourse to a special speech module. Their claim is that listeners directly parse the underlying phonological structure of an utterance from the acoustic signal. This is done by detecting, in the acoustic signal, "the acoustic signatures of gestures as a means of identifying the gestures themselves, which constitute the speaker's phonological message" (Pardo & Fowler 1997, p. 1150).

The direct realist theory appears, like the motor theory, to assume that segments are fundamental at some level of perception. Pardo & Fowler (1997, p. 1141) state that "phonetic segments are specified not only by their spectral characteristics, but also by their temporal properties", which they claim means that "overlapping phonetic gestures can be perceived as just that—physical events that occur over time and overlap, rather than merely influence, one another." These authors then go on to explain that the "perceptual parsing" which takes place under this theory is a process of separating these temporally overlapping gestures from one another to derive a sequence of "gesturally parsed segments" (p. 1143).

Acoustic theorists, in contrast to the motor theorists, posit a much more direct relationship between the acoustic signal and the phonetic percept. Stevens & Blumstein (1981), for example, propose that there are invariant properties in speech—in this case in the acoustic signal—which correspond to phonetic categories (i.e. bilabial stop vs. alveolar stop vs. velar stop) or distinctive features (i.e. [+/− voice]). These invariant properties are clearly not the individual, context–dependent acoustic cues (e.g. burst frequency, duration of silence, formant transitions etc.) discussed earlier in this chapter, although elements of these cues may be present in the invariant properties. Instead the invariance in the speech stream is claimed to be found in the overall shape of the spectrum at particular points or time frames in the acoustic signal, specifically at acoustic boundaries or discontinuities (i.e. the boundaries between segments in an utterance). In the original specification of this theory (e.g. Blumstein & Stevens 1980, Stevens 1980, Stevens & Blumstein 1981) one single spectral sample, or spectral template, was proposed per feature: for stop consonants, for example, the sample consisted of the first 20 ms at the release of the consonant burst. In later versions of this theory (e.g. Stevens 1985) the spectral samples were increased to two, one from either side of an acoustic boundary, which were intended to be compared with each other. Under either version, however, these invariant patterns or properties of the acoustic signal are considered to be the primary cues to phonetic representations or features, and are perceived by innately endowed 'property detecting mechanisms.' All other aspects of the signal—i.e. all context–dependent cues—are seen as secondary cues, used only when phonetic features are represented ambiguously by the primary cues.

H. Sussman and colleagues also propose that invariant specifiers can be found in the acoustic signal: in this case in the form of locus equation coefficients (e.g. Sussman, Fruchter & Cable 1995). Locus equations are "linear regressions of the onset of F2 transitions on their offsets in the vowel nucleus" (Sussman & Shore 1996, p. 936). These researchers claim that the context–conditioned variability of vowel onset formant transitions "gives way, when displayed as locus equation plots, to a lawful acoustic representation of the entire stop category" (Sussman & Shore 1996, p. 936). Originally designed as 'phonetic descriptors' of place of articulation, the theory has been reinterpreted from a perceptual point of view: locus equation coefficients have been suggested (by e.g. Sussman et al. 1995) to be correlates of listeners' perceptions of place (however, see also Fowler 1994).

Kingston, Diehl and colleagues (e.g. Kingston & Diehl 1995), also acoustic theorists or auditorists, propose a theory which directly contradicts that proposed by motor or gestural theorists. Kingston and colleagues claim that the reason that multiple acoustic cues or properties 'cohere' into one percept is not necessarily because acoustic cues have a common articulatory source (i.e. as proposed by the gesturalists), but because they have similar auditory effects. These researchers go on to propose that speakers actively choose to "covary articulations precisely because their acoustic consequences are auditorily similar enough to be integrated into more comprehensive perceptual properties, intermediate between the acoustic properties and distinctive feature values" (Kingston & Diehl 1995, p. 7). Some intermediate properties which have been proposed are *C:V duration ratio*, and a *low frequency property*, which both cue intervocalic voicing. These intermediate perceptual properties (and others like them) are presumed to coincide with the general sensitivities of the auditory system: it is these properties which, under this theory, are the objects of perception.

However, while the mechanisms underlying acoustic theories may differ from those underlying motor theories, in terms of the units of perception acoustic theories tend in general to be similar to the motor theory. Much as the gesturalists, auditorists assume that the units of perception are phonetic features. Stevens and Blumstein's spectral template theory, for example, specifies that perception involves "the analysis of the speech signal into discrete phonetic features" (p. 2), as is clear from the fact that each of the invariant properties that the theory proposes is meant to be a direct correlate of a phonetic feature. Inasmuch as the locus equations posited by H. Sussman and colleagues are intended to be invariant correlates of particular sets of features (e.g. place of articulation features), this acoustically–based theory can also be seen as specifying some sort of reference to phonetic features in perception. Finally, Kingston, Diehl and colleagues' auditory theory also specifies features—for these theorists, "acoustic properties get mapped onto a specific distinctive feature value when a speech sound is identified by listeners" (Kingston & Diehl 1995, p. 7).

2.2 How do adults perceive? Some evidence:

The conclusion that can be drawn from all of the theories described above is that the goal of mature speech perception is the derivation of a sequence of phonemes, whether directly or in terms of features. However, one can also ask if the terms phoneme and segment are used because it is conventional to do so, or because there is evidence that this might indeed be the way in which listeners behave perceptually.

There is a reasonable amount of evidence from various sources that the description of adult speech perception in terms of phonemic units may be correct. Shattuck–Huffnagel (1983, 1987), and MacNeilage & Davis (1990), for example, have shown that productive errors in adults' speech often involve the exchange, confusion, or incorrect serial ordering of phonemic segments. Other researchers have found further evidence of some level of segmental organisation in studies of backward talkers (Cowan, Leavitt, Massaro & Kent 1982) and aphasic speakers (Blumstein 1981).

It would appear from these studies that there might be some empirical grounds for the assumption that perception in normal adults is organised around a phonemic segment. However, we will reconsider this evidence in Chapter 6 when we discuss the implications of alphabetic literacy on perceptual processes.

2.3 What do children do in perception?

We turn now to an examination of what is known about perceptual behaviour in infants and children. A great deal of research into infant speech perception has found that infants' speech *discrimination* abilities are equal, if not superior, to adults'. Experimental methods, such as the *high amplitude sucking* technique, which measures infants' rate of non–nutritive sucking (on a pacifier/dummy) in response to familiar and new stimuli, and the *head turn* paradigm, in which infants respond to new stimuli by orienting themselves towards the sound, have allowed researchers to explore infants' perceptual abilities. Eimas and colleagues, for instance, (Eimas, Siqueland, Jusczyk & Vigorito 1971, Eimas 1974, Eimas 1975, Eimas & Miller 1980, Miller & Eimas 1981), have shown that infants from 1 to 4 months can categorically discriminate speech sounds which differ along a VOT continuum ([ba] to [pa]), along a place of articulation continuum ([bæ] to [dæ]), and along a manner continuum ([ba] to [wa]), and [ra] to [la]). Further studies have gone on to show that not only are infants able to discriminate virtually all sounds which are categorically perceived by adult speakers of their own environmental language, but they are also, *unlike* adults, able to discriminate sound contrasts which are *not* distinctive in their own environmental language. Trehub (1976), for example, showed that English–learning infants could discriminate two non–English contrasts: a Polish or French nasal/non–nasal contrast along a [pa] to [pã] continuum, and the Czech contrast between [řa] and [za]. Streeter (1976) showed that Kikuyu–learning infants were able to discriminate a non–native (English) [ba] to [pa] contrast, and Werker and colleagues (Werker & Tees 1984, Werker & Lalonde 1988) found further evidence that English–learning infants are able to discriminate non–native consonant contrasts—in this case along two Hindi retroflex–dental continua: [ʈa] to [ta] and [ɖa] to [da], and a Nthlakapmx[3] glottalised velar to uvular continuum: [k'i]–[q'i]—and are able to

[3]A native North American (specifically British Columbian) language belonging to the Salish family. Also known as Thompson.

discriminate non–native vowel contrasts—German [ʊ] to [Y] and [u] to [y] (Polka & Werker 1994).

However, despite the fact that these studies show that infants possess what might be considered highly sophisticated or 'adult–like' discrimination abilities, it should not immediately be concluded that the studies also indicate that infants make these discriminations *in the same way* as adults. As noted at the beginning of this chapter, there is a difference between the *function* of a speech perception system and the actual *mechanism* of that system. Jusczyk (1997, pp. 70–71) notes that "In most of the investigations, the stimuli were designed to contrast with respect to a single phonetic segment. Therefore, there is an inclination to view the infant who successfully discriminates "bug" from "dug" as detecting a difference between the initial phonetic segments [b] and [d]. Although this is an accurate enough description of how the experimenter views the contrast, it may not be a valid description of what infants are doing." In fact, as Jusczyk goes on to say, there is no need for the infants to analyse the stimuli into segments in order to know that they are different: it is possible for them to perceive speech in terms of much more global units, like words or syllables.

There is some evidence from perception studies that infants and young children do indeed perceive in terms of units which are more like syllables or monosyllabic words than phonemes. Studies by Jusczyk & Derrah (1987) and Bertoncini, Bjeljac-Babic, Jusczyk, Kennedy & Mehler (1988) have shown that infants are not sensitive to similarities at the level of the phonemic segment. In these studies, infants were familiarised with a series of syllables that had a common initial consonant, e.g. [bi], [ba], [bo], [bɚ]. The authors then tested the infants' responses to one of two new syllables: one that shared the same initial consonant: [bu], and one that did not: [du]. The prediction was that if the infants perceived the syllables in terms of a sequence of individual segments, they would detect the syllable with the new initial segment, but would be less likely to detect the syllable with the same initial segment. The results showed that the infants treated *both* the new syllables in the same way, suggesting to the authors that the individual syllables were not being perceived in terms of sequences of segments, but rather in terms of a whole CV unit.

Additional studies by Bertoncini, Mehler and colleagues (Bertoncini & Mehler 1981, Bertoncini, Floccia, Nazzi & Mehler 1995, Bjeljac-Babic, Bertoncini & Mehler 1993) have shown that infants are better able to discriminate three phoneme units which conform to a syllabic pattern than those that do not: two–month–olds can discriminate [tæp] from [pæt] more easily than they can discriminate [tsp] from [pst]. They have also found that infants detect differences in the number of syllables between stimuli, but do not detect differences in the number of phonetic segments, or morae[4].

Evidence from other perception studies suggests that some flexibility in the exact definition of children's unit of organisation may be required. A study by Walley, Smith & Jusczyk (1986), for example, examined 5–year–old and and 7–year–old children's judgments of similarity between pairs of 2–3 syllable CVCV stimuli. These stimuli were designed so that the pairs shared an initial consonant, an initial CV unit, or an initial CVC unit. The older children in the study had little difficulty in judging that two stimuli which shared only one phoneme were similar. The younger children, on the other hand, had most difficulty with the pairs of stimuli which only shared an initial consonant, supporting the hypothesis that phonemes are not the unit of perception for infants and young children. However, the size of unit that the youngest children were most successful at classifying as similar was not the CV unit as in the studies above, but the CVC unit. Additionally, in the study carried out by Bertoncini et al. (1988, noted above) it was found that when neonates were tested on the same CV syllables as the slightly older infants, they did not treat all new syllables as different from the familiar syllables. Instead, neonates appear only to notice new syllables if they contain a different vowel from the original syllables, prompting Mehler, Dupoux, Nazzi & Dehaene-Lambertz (1996) to propose that it may be the vocalic nucleus of the syllable which is important to the perceptual system developmentally.

These studies suggest that we should not jump to the conclusion that the syllable is the *exact* unit of perceptual organisation for infants and young children. However, they do make it clear, along with the results of the other studies discussed in this section, that it is unlikely that the initial perceptual unit is a phoneme.

2.4 What do children do in production?

Having shown that it is indeed possible that a change in the unit of perception might occur from childhood to adulthood, we now turn to other areas of language development to determine whether such shifts might occur in any other

[4]A mora is a Japanese phonological unit. Cutler (1996) defines a mora, briefly, as "a CV structure, or a single vowel, or a syllabic coda (usually a nasal consonant); thus *Honda*, for example, has three moras: Ho–n–da" (p. 93). The infants in Bertoncini et al.'s (1995) study were French–learning rather than Japanese–learning.

phonologically organised system. There is a great deal of evidence from studies of phonological *output* development to suggest that children's earliest phonological representations may by organised around something like a one– or two–syllable word. Studdert-Kennedy (1987), in a review of these studies, divides the evidence up into three main points. The first two of these points come from a study by Ferguson & Farwell (1975). These authors observed a 15–month–old who produced phonetic forms correctly in certain contexts but not in others—[n] was produced correctly in 'no' but [m] was produced instead of [n] in the target word 'night' and [b] was produced instead of [m] in the target word 'moo.' Studdert-Kennedy (1987) notes that in this example it appears that "the child does not contrast [b], [m], [n], as in the adult language, but the three words with their insecurely grasped onsets" (p. 76).

The second piece of evidence that Studdert–Kennedy notes from Ferguson & Farwell's (1975) study is again an observation of a 15–month–old's attempts to produce an adult target—in this case the word 'pen.' In the course of half an hour, the child produced [mã^{ə}, ^{v}ʌ̃, dɛ^{dṇ}, hɪn, ^{m}bõ, p^{h}ɪn, t^{h}ṇt^{h}ṇt^{h}ṇ, ba^{h}, ɖau^{N}, buã]. Studdert–Kennedy notes that most of this child's productions actually have many of the same gestures as the adult model—the child simply has not timed these gestures correctly relative to each other[5]. Again, according to Studdert-Kennedy (1987), the child is not attempting to reproduce a sequence of phonetic segments, but rather a "holistic pattern of gestures" (p. 78), indicating the perceptual importance of the unit over which these gestures operate—i.e. the monosyllabic word.

The third piece of evidence presented by Studdert-Kennedy (1987, see also Nittrouer, Studdert-Kennedy & McGowan 1989) is the extensive use of consonant harmony or consonant deletion in child utterances. Children appear to have difficulty with, or avoid, switching place or manner of articulation, thus: "one child may attempt *fish* with [fɪˈ], another with [ɪʃ]; faced with *duck*, one child may try [gʌk], another [dʌt]" (Nittrouer et al. 1989, p. 120). Menn (1983) describes child productions of this sort as a single unit which has been "assembled before it is spoken" (p. 16).

In terms of a change in units of organisation in production, Vihman (1996) notes that evidence of this can be seen in studies which follow children's production

[5]Ferguson & Farwell (1975) actually label the aspects which are common to both the adult model and the child's attempts as 'features', stating that the child "seems to be trying to sort out the features of nasality, bilabial closure, alveolar closure and voicelessness" (p. 423).

behaviour over time: "such studies 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 of liquids)" (p. 216).

Additional evidence that children and adults make use of different sized units in production also comes from studies by Nittrouer and colleagues of coarticulation in speech (Nittrouer et al. 1989, Nittrouer 1993, Nittrouer 1995, Nittrouer, Studdert-Kennedy & Neely 1996). Nittrouer (1993) suggests that if children's unit of production is the syllable, then they may articulate syllabic units as "largely undifferentiated entities" (p. 960), beginning and ending all gestures in the syllable at roughly the same time. If this is the case, then, according to Nittrouer, individual gestures in children's speech will exert their influence over a greater proportion of the utterance than will the same gestures in adults' speech—that is, children's speech will be more coarticulatory. The studies carried out by Nittrouer and colleagues did indeed find that young children show a greater degree of within–syllable coarticulation than do older children and adults, leading the authors to conclude that "young children organise their speech over a wider temporal domain than adults do" (Nittrouer et al. 1989, p. 130).

2.5 What do children do metalinguistically?

There is one further area of linguistic development that displays shifts in units of phonological organisation similar to those suggested for perception: the development of metalinguistic awareness, which begins with awareness of large units, and involves an emergence of awareness of gradually smaller units.

Metalinguistic awareness can be defined as the ability to consciously think about and manipulate (i.e. count, delete, segment, correct) variously sized segments of language. As described by A. Fowler (1991), the process of "gaining access to these segments in order to count, label or manipulate them...is akin to becoming aware of the many movements that go into walking for the purpose of learning ballet" (p. 99). I. Liberman and colleagues (Liberman, Shankweiler, Fischer & Carter 1974) divide those units of speech which can be thought about consciously into 'meaningful' units, such as sentences, words, and morphemes, and 'meaningless' units—i.e. units that have no semantic meaning on their own—such as syllables, onset–rime units[6], phonemes and features. Within each of these two groups, awareness of units of different size develops at different times.

Of particular interest to the current line of questioning is the development of awareness of meaningless units—also known as phonological awareness. It has been found that young children are able to count the number of syllables in a word and correctly judge rhyme and alliteration before they can count the number of phonemes in a word or judge whether words end with the same segment (e.g. Bradley & Bryant 1983, Liberman et al. 1974). Thus, phonological awareness begins with awareness of syllables and onset–rime units, and moves to the awareness of phonemes, in much the same way as has been suggested for perception. It should be noted at this point that the development of awareness of phonemic segments has been found to be highly linked to the development of alphabetic literacy. This means that the change in size of unit at a metalinguistic level may be more complicated than has been posited for the comparable change at other levels. The details of and criteria for the development of phonemic awareness will be discussed in more depth in Chapter 2 (Section 2.2).

At this point, we have amassed evidence not just for the possibility of a change in size of organisational unit in perception, but also for similar changes in production and in metalinguistic abilities. The question to be addressed now is why such a change should occur.

2.6 *Why start with syllables?*

Various reasons have been proposed for why children's early phonology might be organised around syllable– or monosyllabic word–sized units. Studdert-Kennedy (1987) states that "a child's entry into language is mediated by meaning; and meaning cannot be conveyed by isolated features or phonemes." Instead, the earliest unit of meaningful contrast for the child is "the word (or formulaic phrase)" (p. 67, see also Studdert-Kennedy 1991).

Menn (1983) gives a description of a child's approach to word production which suggests an additional possible explanation for the apparent cohesion of syllables for children. Menn notes that the child in question seemed to have

[6] Throughout this thesis the unit which corresponds to the part of a word which follows a consonant/consonant cluster onset will be referred to as the *rime*. The term *rhyme* will be reserved for the action of producing or recognising words with the same rime.

"learned an articulatory program of opening and closing her mouth that allow[ed] her to specify two things: the vowel and one point of oral closure" (p. 5). This description corresponds well with theories that propose a motor–organisational explanation for syllable–based phonologies in child language (e.g. Davis & MacNeilage 1995, MacNeilage & Davis 1990, MacNeilage & Davis 1993, MacNeilage 1994). MacNeilage (1994), for example, states that a syllable is effectively an alternation of opening and closing of the jaw: "a cycle of elevation and depression of the mandible, the two phases of which are associated with vowels and consonants, respectively" (p. 185). Davis & MacNeilage (1995) suggest that these syllabic cycles of opening and closing are the 'frames' into which 'content' or segmental sequences are organised. MacNeilage (1994, p. 186) goes on to point out that "the first truly speech–like vocalization of infants, their babbling, is highly syllabic, though relatively undifferentiated at the segmental level" (i.e. all 'frames', little 'content'), emphasising the early motoric importance of syllables over segments for children.

Research by Vihman and colleagues (e.g. Vihman 1992, Vihman 1993, Vihman, Macken, Miller, Simmons & Miller 1985, Vihman, Velleman & McCune 1994, Vihman & Velleman 1989), has shown that children's first words tend to capitalise on these early babbling patterns: there is continuity between a child's range of "vocal motor schemes" (Vihman 1993) and the same child's early words, both in terms of the repertoire of sound patterns, and the overall word shape (Vihman et al. 1985). Vihman and colleagues suggest that a child's earliest (recognisable) attempts at adult targets are selected for production by the child based on the similarity of the targets to the child's existing vocal motor schemes. There is also a suggestion that the child may perceive adult speech via an articulatory filter built around its own motor patterns which specifically enhances those adult patterns that match its babbles (Vihman et al. 1994). These adult targets are then altered to further increase the fit to the child's pre–existing motor schemes (e.g. Vihman & Velleman 1989). These studies all underline the probability that the initial organisation of children's vocalisations around a word or syllable motor cycle continues into early speech.

2.7 *Why change from syllables?*

Having shown that it is quite feasible that children's initial unit of perception could be something like a syllable or word, the question that remains to be asked is why this unit would need to change. There are two hypotheses, not necessarily mutually exclusive, which have been proposed to explain changes in representational unit (whether immediate perceptual representation or more long–term phonological representation). The first comes from studies of phonological output development, while the second is more closely related to the development of metaphonological skills.

The first hypothesis is the most commonly proposed explanation for a change in organisational unit for both perception and production—specifically, that this change takes place under pressure from a child's growing lexicon (Lindblom, MacNeilage & Studdert-Kennedy 1983, Lindblom 1989, Studdert-Kennedy 1987). Describing this in terms of articulatory or motor–type theory, Studdert-Kennedy (1987) states that

> As long as the child has only a few words, it needs only one or two articulatory routines. Initially it exploits these routines by adding to its repertoire only words composed of gestural patterns similar to those it has already 'solved' and by avoiding words with markedly different patterns. Once the initial routines have been consolidated, new routines begin to emerge under pressure from the child's accumulating vocabulary. New routines emerge either to handle a new class of adult words, not previously attempted, or to break up and redistribute the increasing cohort of words covered by an old articulatory routine. (p. 79)

This first, lexically–based hypothesis is the explanation that Nittrouer and colleagues tend to support. It is fairly clear that the Developmental Weighting Shift is intended to describe the *maturation* of the perceptual system under the influence of a child's growing experience with language and the resultant expansion of the child's lexicon. Nittrouer & Miller (1997*b*) state that

> This model [the DWS] specifically suggests that children initially show a preferential weighting of dynamic, acoustic properties (i.e. those that change spectrally over time) because these properties help them recognise syllabic structure in the speech stream. *With linguistic experience* [my emphasis], children gradually decrease their attention to these dynamic properties, and increase their attention to the static properties (i.e. those that do not involve spectral change over time) that are particularly informative about phonetic structure in their native language. (p. 2254)

However, Nittrouer also briefly posits an alternative and less conventional hypothesis regarding changes in units of perception. Nittrouer (1992, 1996*b*) proposes that the shift from words or syllables to phonemic segments that she has observed in perception may not just reflect a change in sensitivity to phonological structure at a *subconscious* level. She suggests that this change may also be related to a change in access at a *conscious* level. It has been noted above that the development of phonological awareness involves a change in conscious access from awareness of syllables to awareness of phonemes. This shift appears to parallel the perceptual shift documented by Nittrouer and colleagues, suggesting a possible relationship between the two. It is at this point that we return to the connection between perception and literacy touched on at the beginning of this chapter.

The idea that phonological organisation can be influenced by developments in metalinguistic processing, though unconventional, *has* been suggested by other researchers. A. Fowler (1991) for instance, states that "there are many reasons to think that metalinguistic factors may play a role in developing a phonemic representation" (p. 110), giving as an example the importance that language play has been hypothesised to have on the further refinements in word representations in toddlers. A. Fowler notes that evidence of phonological processing in deaf children also indicates the possibility that the development of phonemic awareness may impact on phonological representations. Hanson (1991), for example, states that "deaf children's lack of complete access to the auditory aspects of English provides them with a considerably different language experience than that of hearing children. Whatever knowledge they may acquire about English phonology will be largely influenced by visual experiences such as lipreading and reading and by gestural experiences such as speaking" (p. 154). All of these visual and gestural experiences can be classified as metalinguistic exercises for a deaf child, as the process of learning them will require a conscious awareness of language on the part of the deaf learner. This suggests in a very circumstantial way that the development of metalinguistic skills has some sort of impact on phonological processing.

More specific empirical evidence of the potential effect of metalinguistic development on phonological processing comes from studies of deaf good and poor readers. Hanson (1991) describes studies in which children from these two groups were tested on their ability to recall lists of printed letters (i.e. 'A, Q, B, D' etc.) which either did or did not rhyme. Studies (previous to Hanson's) of word recall by hearing subjects have found an effect of rhyme: hearing subjects are more likely to confuse items in rhyming lists than in non–rhyming lists (e.g. Conrad 1971), reflecting their sensitivity to the phonological similarity of the rhyming words. Deaf children, on the other hand, should be unaffected by rhyme. However, the studies described by Hanson (1991) show that deaf good readers have more difficulty remembering lists of rhyming letters than lists of non–rhyming letters, while no such effect was seen for deaf poor readers. This suggests that those deaf children who have been successfully taught to read have been able to acquire a system of phonology of some sort which interferes with their short–term memory in the same way as it would hearing subjects.

Morais & Kolinsky (1995) also propose that orthographic knowledge and phonemic awareness may have an effect on phonological representations at some level. These authors provide evidence of misperceptions from dichotic listening tests which suggest that literate, semiliterate and illiterate subjects have different perceptual strategies with regard to the segmental structure of speech. Dichotic listening tests are tests of perceptual processing, in which the listener is presented with two different stimuli, one stimulus in each ear, and asked to report what they have heard. One of the possible results is a 'fusion' error, in which parts of the two stimuli are combined perceptually by the listener. What was found by Morais and Kolinsky was that the proportion of errors which were 'global errors'—that is, errors in which a whole syllable or more was completely misrecognised, e.g. 'dono' for 'cano'—was highest in the illiterates, and grew less with literacy. Contrastively, the proportion of errors which were 'segmental'—that is, errors in which only one segment was misrecognised, e.g. 'pano' for 'cano'—was lowest in the illiterates, and grew more with an increase in literacy. The authors conclude that it is the literate subjects' conscious awareness of phonemes which influences their particular type of phonological error.

All of these examples taken together suggest that this second, metalinguistically–driven hypothesis is at least a reasonably plausible explanation for a change in perceptual units.

A study by Nittrouer (1996*b*) of the relationship between the Developmental Weighting Shift and explicit phonemic awareness skills did indeed find a relationship between the two. The results of this study showed that those children who gave more weight to fricative cues in distinguishing /s/ from /ʃ/ had better phonemic awareness, while those children who gave more weight to transitional cues had poor phonemic awareness. This in itself opens the door to the possibility that shifts in perceptual strategy could be influenced by the development of phonemic awareness.

Unfortunately, however, the results of Nittrouer's (1996*b*) study could be explained by reference to either a lexical or a metaphonological explanation. As noted by Nittrouer (1996*b*), it is equally likely that "learning how best to weight acoustic properties may be a requisite for recognizing phonetic structure" (p. 1067) or that "discovering syllable–internal structure [i.e. developing phonemic awareness] may actually create pressure to develop the most effective processing strategies for providing access to that structure" (pp. 1067–1068). This leaves us to determine which of these two possible hypotheses is correct.

2.8 Possible implications for perceptual development

At this point we should consider the possible implications of these alternative outcomes on our understanding of perceptual development as a whole.

If it is possible that some aspect of speech perception might be influenced by the development of literacy skills, and metaphonemic awareness, then we have to consider that perceptual development is not an entirely maturational process. This in turn means that we would have to reconsider the way in which 'mature' speech perception is characterised.

Additionally, if a process such as phonemic awareness, which is so closely related to alphabetic reading skills, can have a developmental impact on speech perception, we would have to drastically redefine our understanding of the relationship between speech perception and literacy, and indeed all higher cognitive processes.

Furthermore, if the change in perceptual unit from word or syllable to phonemic segment is the result of the development of a non–essential cognitive process, then it should also follow that the change in perceptual unit is a non–essential change. As a result, we would then have to entertain the possibility that perception of phonemic segments might not be the goal of perceptual development. This brings into question the status of the phoneme in perception itself.

Finally, it should be pointed out that the two hypotheses regarding the cause of a change in perceptual unit are not necessarily mutually exclusive. A. Fowler

(1991, p. 111), for instance, suggests that “Although it may well be that metalinguistic experience in general, and orthographic experience in particular, may aid us in refining our phonological representations, these findings need not commit us to the view that phonemes are arbitrary or epiphenomenal in nature.” Instead, she suggests that a theory which aims to describe the emergence of the phoneme must simply take into account the fact that “phoneme–level representations, implicit as well as explicit may not come for free but rather must emerge over time, in the course of lexical expansion, language play, and, potentially, orthographic experience” (p. 111). If this is the case, then we must consider that speech perception is not a unitary construct, but rather is multifaceted, with each facet under the potential influence of different developmental demands.

CHAPTER 2

Theoretical background and goals

The starting point of the current study, both theoretically and methodologically, is a study carried out by Nittrouer (1996*b*), which showed a correlation between a change in representational unit at the conscious level, and the Developmental Weighting Shift model of speech perception. This study (which will be discussed in more detail below) is one of a number of studies carried out by Nittrouer which aimed to determine the course of development of subconscious perceptual behaviour. The aim of Nittrouer's *first* perceptual studies, however, was not to establish or support a model of speech perception. Instead, these studies were designed to evaluate an aspect of the argument between acoustic and gestural theorists: that of the role of 'coarticulatory' cues in speech perception. As noted in Chapter 1 (Section 1.2), the acoustic variability in speech segments caused by their articulation in the context of other segments causes problems for perceptual theories. The question that drove Nittrouer & Studdert-Kennedy's (1987) study was:

> Is coarticulation necessary and intrinsic to production, and must a listener therefore draw on the contextually variable information that it carries to recover the phonetic message? Or...are the acoustic consequences of coarticulation merely noise that a listener filters out? (p. 319)

Nittrouer & Studdert-Kennedy believed that they could determine the answer to this question by examining the *development* of use of coarticulatory, or transitional cues. If young children are not sensitive to transitional cues, then these cues must be something which a developing perceptual system has to learn to cope with—indicating that they play a secondary role in perception, as suggested by acoustic theorists (e.g. Stevens & Blumstein 1981). Alternatively, if children *are* sensitive to transitional cues, then it is more likely that such cues play an important role in perception from the onset of development—indicating, as suggested by gestural theorists, that "the listener uses the systematically varying transitions as information about the coarticulation of an invariant gesture with various vowels, and so perceives this gesture" (Liberman & Mattingly 1985, p. 6). The results of Nittrouer & Studdert-Kennedy's (1987) study showed that children *are* sensitive to transitions, which the authors state "runs counter to the claim of Stevens and Blumstein (1978) that sensitivity to coarticulation in adult speech perception is a secondary effect, learned by association with a primary invariant" (Nittrouer & Studdert-Kennedy 1987, p. 329). This result was therefore taken as support for a gestural description of speech perception.

Additionally, and importantly for Nittrouer's future model of perceptual development, the results of this study and of those that followed also went beyond addressing the above issue. These studies showed that children are not simply sensitive to transitional information: in certain contexts they are *more* sensitive to this information than adults. It is this finding upon which the Developmental Weighting Shift (DWS) model is based.

The DWS model states 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 1996*b*, p. 1060). We will examine the two parts of this model individually. The first part of the model is a developmental shift in weight given to acoustic cues. In the initial definitions of the model (e.g. Nittrouer & Studdert-Kennedy 1987) this was simply a shift from heavier weighting of dynamic transitional cues by children, to heavier weighting of more static aspects of the signal by adults. Later, in response to further studies of the phenomenon, Nittrouer and colleagues elaborated on the model: the shift was said to move from dynamic transitional cues to those cues which are most 'informative' about the segmental structure of the speech stream—i.e. if transitional cues happen to be the most 'informative' in a certain situation, then listeners will continue to weight them more heavily through development (Nittrouer & Miller 1997*b*).

The second aspect of the model is the fact that the shifts in attention described above are related to increases in sensitivity to phonetic structure. Thus children attend to transitional cues more heavily than adults because children initially process all speech into syllable sized units, while adults process speech into segment–sized units. Nittrouer explains this by relating acoustic transitional cues to the articulation of a CV syllable:

> vocalic formant transitions in CV syllables result from the vocal–tract changes associated with moving from a consonant to a vowel configuration. Using the term 'constriction' to refer both to the close constriction of the tongue tip and/or blade used in the production of /s/ and /ʃ/, as well as to the looser constriction of the tongue body during vowel production, it can be seen that acoustic properties of these transitions are determined by the constriction locations of both consonant and vowel.

Children therefore attend to syllable–internal transitions because they reflect the whole CV syllable structure. Adults, on the other hand, attend to cues like frequency of fricative noise spectrum, and vowel target formant values, because these more closely reflect individual (possibly segment sized) components of the syllable.

Again, earlier definitions of the model were more simplistic than later versions: initially this increase in sensitivity was seen as operating at a subconscious, processing level only. In Nittrouer's (1996*b*) study, however, the idea of "increases in sensitivity to phonetic structure" was expanded to include increases in *conscious* sensitivity to phonetic structure[1]. In this study, Nittrouer explored the possibility that the shifts in cue weighting which had been observed in pre–school and early–school–age children might be related to the development of phonemic awareness—the conscious awareness of, and ability to manipulate, phonemic segments. This skill is related to the development of literacy skills, and therefore develops in children at roughly the age at which cue weighting shifts had been found to occur—i.e. early school age. The results of Nittrouer's (1996*b*) study did indeed find strong correlations between ability on phonemic awareness tasks, and the degree to which dynamic vs. static cues were weighted: children with poor phonemic awareness weighted cues similarly to the children in Nittrouer's previous studies, while children with good phonemic awareness weighted the cues similarly to the adults in earlier studies.

[1]A possible relationship between shifts in acoustic cue weighting and the development of phonemic awareness was posited briefly by Nittrouer & Studdert-Kennedy (1987), and Nittrouer (1992), however was not expanded on until Nittrouer (1996*b*).

Because the development of phonemic awareness involves an expansion at the conscious level from the ability to focus on larger units like syllables and onset–rime units, to the ability to focus also on phonemes, the results of the 1996 study have been taken by Nittrouer as strong support for her DWS model, in particular her hypothesis that perception develops from syllable–based organisation to phoneme–based organisation. However, this study also raises issues regarding the perceptual phenomenon which the model is attempting to describe. Specifically, this study leaves open the question of possible developmental influences within the relationship. From the results of this study it is, as Nittrouer (1996*b*) points out, equally possible to conclude that phonemic awareness could affect speech perception as it is to conclude that speech perception affects phonemic awareness development. This being the case, one can ask whether shifts in cue weighting are, as Nittrouer claims, strictly developmental, or whether they are affected by the development of conscious awareness of phonemes.

The central goal of this thesis will therefore be to attempt to answer this question. **This thesis will be an investigation of the relationship between shifts in acoustic cue weighting in perception, and the development of phonemic awareness skills. In particular, this study will focus on the possible causal nature of the relationship, and will attempt to determine the extent to which one of these processes might affect the development of the other.**

Before this investigation can be carried out, however, it is important first to establish what is already known about the development of the two processes in question. This chapter will describe those studies which have explored the development of acoustic cue weighting, and phonemic awareness, both in isolation and in relation to each other.

1 The development of acoustic cue weighting

1.1 The Developmental Weighting Shift model

The Developmental Weighting Shift model is based primarily on the results of two perceptual studies: Nittrouer & Studdert-Kennedy (1987) and Experiment 3 of Nittrouer (1992). Nittrouer & Studdert-Kennedy (1987) examined perception of fricative–vowel stimuli, in particular 'sue', 'shoe', 'sea' and 'she', by 3–, 4–, 5–, and 7–year–old children and adults. This study found that the younger children (3–5 years) made significantly more use than the older children and adults of within–syllable transitional information in their perception of these syllables. In fact, the younger the children were, the more use they made of this information: 3– to 4–year–olds made significantly more use of transitional cues than did the 5–year–olds. Contrastively, the 7–year–olds and the adults made significantly more use of information provided by the frequency of the fricative noise than did the younger children.

As noted both above, and in Chapter 1 (Section 1.3), the results of Nittrouer & Studdert-Kennedy's (1987) study have been interpreted by the authors as meaning that children are more influenced in their perceptual decisions by dynamic aspects of the speech stream (such as formant transitions), while adults are more influenced by relatively more static aspects of the speech stream (such as frequency of fricative noise). This in turn has been taken as support for the hypothesis that children organise their perception very globally, possibly in terms of syllable–sized units, while adults' perception is much more analytical, being organised around something more like a phonemic segment. However, the results of this one study do not actually provide any evidence that there is anything about the shift in cue weighting which has to do with syllable preference in children, or phonemic segment preference in adults, other than the fact that the transitions that the children weighted more heavily were within–syllable transitions. Fortunately, a study carried out by Nittrouer in (1992) provides what can be seen as critical support for this aspect of the model. The study examined perception of two syllable VCCV stimuli: 'arda', 'alda', 'arga', 'alga', again by children (age 5 and 7 years) and adults. The results of this study found that young children made *less* use of across–syllable transitional information than did older children or adults. This is taken as evidence that children are not so much attentive to transitional cues across the board, but rather that they attend more heavily only to certain transitions: those that occur within syllable boundaries.

How was Nittrouer able to uncover these apparent developmental trends in acoustic cue use? The answer lies in the experimental paradigm used for these two studies. This method, which was also used for all of Nittrouer's other speech perception studies, is based on Fitch et al.'s (1980) cue trading relations study (discussed in detail in Chapter 1, Section 1.3)[2]. The most straightforward way to explain this methodology, as well as the premise behind it and the results that

[2] Nittrouer has also conducted a number of studies of perception of non–speech sounds, which have not always made use of the same methodology.

might be expected from it, is to give a detailed description of its use in Nittrouer's first study of acoustic cue weighting.

Nittrouer & Studdert–Kennedy 1987

The main goal of the method employed by Nittrouer is to be able to compare relative use, or *weighting*, of two different cues to the same contrast. For Nittrouer & Studdert-Kennedy's (1987) study these two cues were i) the frequency of fricative noise and ii) the configuration of vowel onset formant transitions, in the fricative–vowel syllables /si/, /ʃi/, /su/ and /ʃu/ ('sea', 'she', 'sue', 'shoe'). Both of these aspects of the signal can be used in the identification of /s/ and /ʃ/ word–initially. As can be seen in the stylised spectrograms of the two fricative–vowel syllables /su/ and /ʃu/ in Figure 2.1, the fricative noise cue and the formant transition cue are different for each syllable: the frequency of the fricative noise is relatively high for the /s/ and relatively lower for the /ʃ/, while F3 falls from a higher onset point, and F2 falls from a lower onset point, following /s/ than following /ʃ/.

These differences mean that either cue could potentially be used by the perceptual system to signal the place of articulation of the fricative. In order to determine listeners' relative use of these (or any other) cues to a contrast, stimuli are designed in which the two cues do not always agree as to the percept they should engender. First, a continuum is created in which one of the cues varies gradually from a configuration which cues one percept, to a configuration which cues the other percept. In the case of Nittrouer & Studdert-Kennedy's (1987) study, it was the fricative noise that was varied along the continuum, from a fricative frequency which is a strong cue to /ʃ/, to a frequency which is a strong cue to

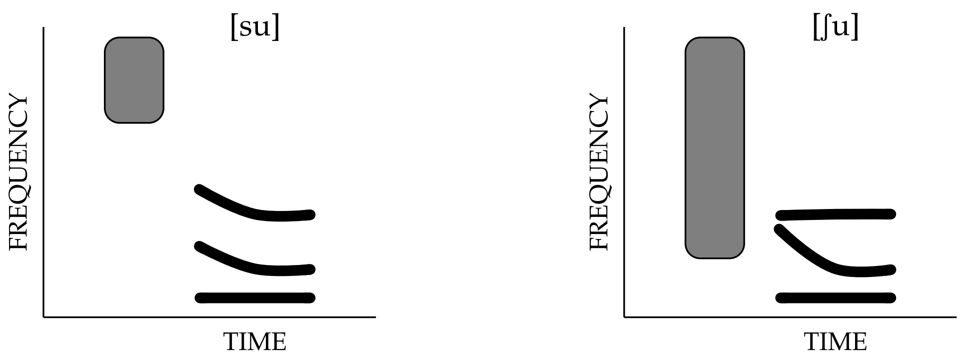

Figure 2.1: **Stylised spectrograms of /su/ (left) and /ʃu/ (right).** Note that both the frequency of the fricative noise and the onset configurations of the vowel formants differ between these two CV–syllables.

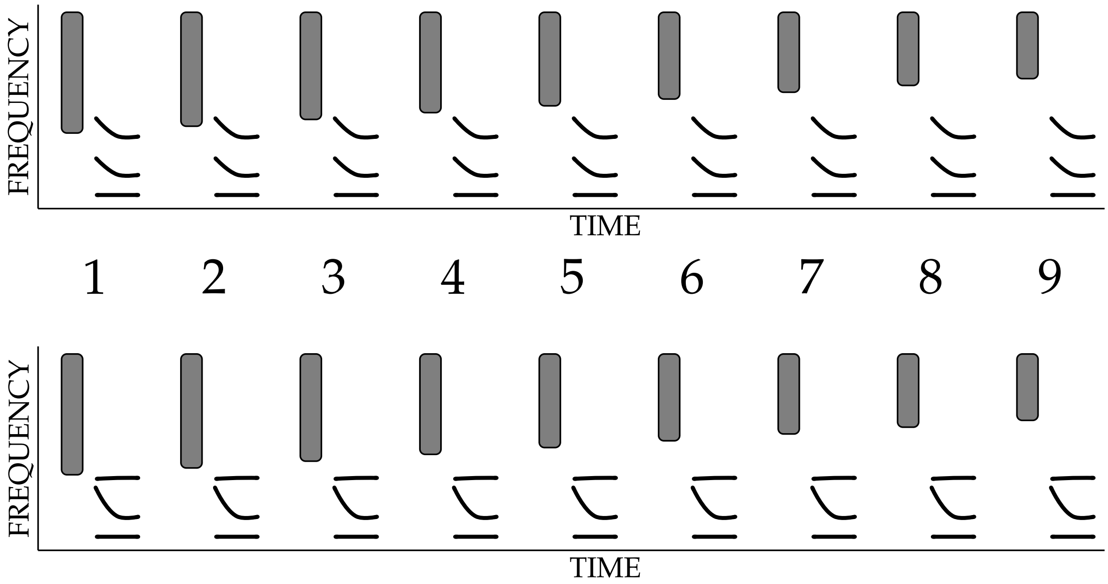

Figure 2.2: **Stylised spectrograms of Nittrouer–style /ʃu/–/su/ continua**. Numbers refer to stimulus number on the fricative noise continuum: 1 is the most /ʃ/–like fricative noise, while 9 is the most /s/–like fricative noise. The uppermost continuum has vowel onset formant frequencies which are appropriate for having followed /s/, while the lower continuum has vowel onset formant transitions which are appropriate for having followed /ʃ/.

/s/. The other cue—in this case the formant transitions—is set in one of two configurations, each strongly cuing only one of the two percepts: that is, either i) transitions appropriate for having followed /s/, or ii) transitions appropriate for having followed /ʃ/. To complete the stimuli, each of these binary varying cues (the transitions) is combined with each of the more continuously varying cues along the continuum (the fricative noises)[3]. This means that for each point on the fricative continuum there will be two different stimuli. Put another way, this means that (as illustrated in Figure 2.2) the stimuli used in Nittrouer & Studdert-Kennedy's (1987) study were effectively two /ʃ/–vowel to /s/–vowel continua, with identical fricative noises, and identical vowel targets, but different vowel onset formant transitions.

The reasoning behind the use of this type of stimuli is as follows. On its own a continuum of a single speech cue—e.g. the fricative noise—should engender a reasonably classical categorical perception response: a rapid change at one point

[3]Note that, as pointed out in Chapter 1, Section 1.3, neither the speech articulators, nor the signal that they create vary in strictly binary configurations—however, this configuration is useful for the purposes of these studies.

on the continuum from one phoneme label to the other, and fairly consistent labeling on either side of this change–over point. However, as explained in Chapter 1, Section 1.3, the addition of a second cue to the stimuli affects the labeling of the contrast. The second cue will interact with the first cue and, for those stimuli where the first cue is ambiguous, will either reinforce or contradict the percept which would be engendered by the first cue alone. In the case of the fricative–vowel syllables in Figure 2.2, without the vowel transition cue the listener might place the /s/–/ʃ/ category boundary between fricative noises 4 and 5. However, with the addition of formant transitions which cue /ʃ/, the boundary might be placed closer to the /s/ end of the continuum: between fricative noises 5 and 6 for example. The reason for this is that the addition of the /ʃ/–transition reinforces the /ʃ/ percept, meaning that a slightly less /ʃ/–like fricative noise is now sufficient to engender an /ʃ/ percept. The same is true of the addition of a set of formant transitions which cue an /s/: more of the stimuli should be perceived as /s/, so the phoneme category boundary should be shifted towards the /ʃ/ end of the continuum. This shift is illustrated in Figure 2.3.

This *trading relationship* between cues is what allows Nittrouer to determine the extent to which each cue is used by a listener: her testing materials are designed to take advantage of this perceptual phenomenon, as will be seen below. Because the stimuli in this type of test form two continua, each listener will have two sets

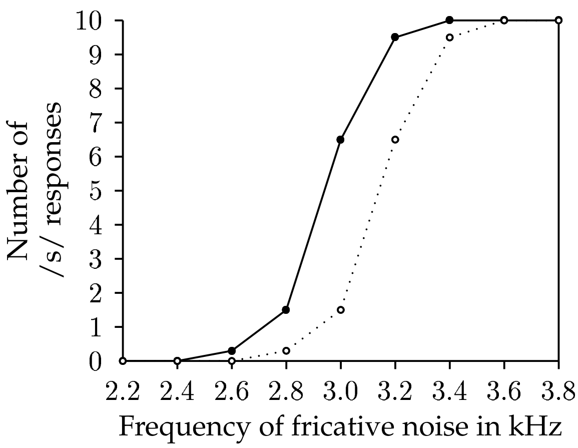

Figure 2.3: **Example responses to Nittrouer–style /ʃ/–/s/ continua.** The x-axis shows the continua of fricative noises, ranging in frequency from 2.2kHz (the most /ʃ/–like) to 3.8kHz (the most /s/–like). The solid line represents a listener's /s/ responses to stimuli with /s/–transitions; the dotted line represents the same listener's /s/ responses to stimuli with /ʃ/–transitions.

of responses, and thus two categorical–perception–type response curves, one per continuum (as illustrated in Figure 2.3 for an /ʃ/–vowel to /s/–vowel continuum). If the cue which is varied along both the continua—in this case the fricative noise—is used more heavily by the listener than the cue which is only varied between the continua—in this case the transitions—then such a listener should label both the continua relatively similarly (i.e. place their category boundaries in roughly the same place for each). In the case of Nittrouer & Studdert-Kennedy's (1987) stimuli, this is because the *fricative* parts of the fricative–vowel syllables in both continua are the same. If, alternatively, the listener makes more use of the cues which vary between the continua than those which vary along the continua, such a listener should label the two continua very differently. Again in the case of Nittrouer & Studdert-Kennedy's (1987) stimuli, this is because the *formant transitions* for each continuum are completely different.

Figure 2.4 illustrates two extreme, and highly hypothetical, sets of responses from listeners making sole use of only one of the two cues present: Graph (A)

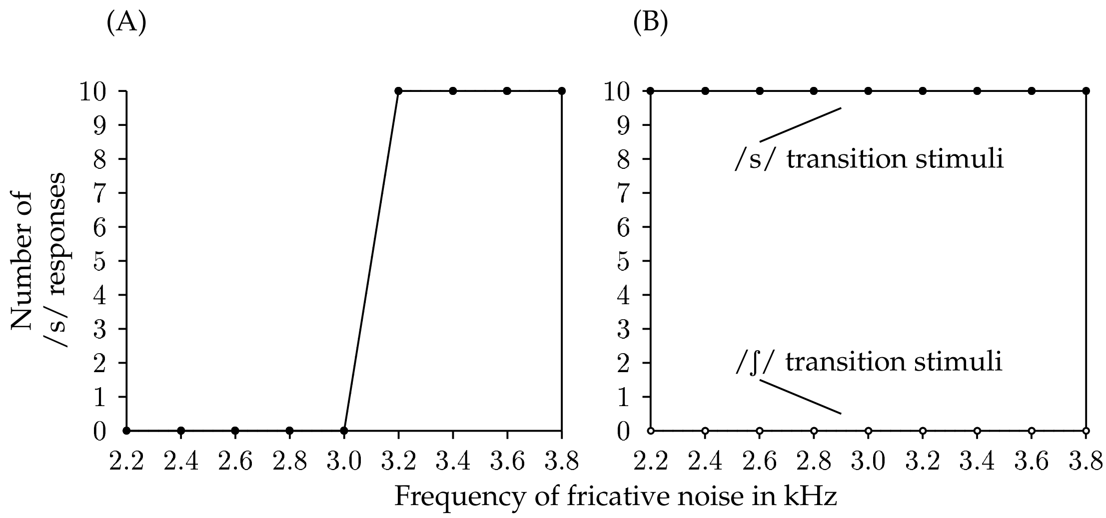

Figure 2.4: **Hypothetical extreme responses to Nittrouer–style /ʃ/–/s/ continua**. The x-axis shows the continua of fricative noises, ranging in frequency from 2.2kHz (the most /ʃ/–like) to 3.8kHz (the most /s/–like). On both graphs, the line with filled dots represents a listener's /s/ responses to stimuli with /s/–transitions; the line with open dots represents the same listener's /s/ responses to stimuli with /ʃ/–transitions (note that because the two response curves are identical in Graph (A), only the line with the filled dots is visible in this graph). Graph (A) represents responses from a listener who makes sole use of fricative noise cues, while Graph (B) represents responses from a listener who makes sole use of transitional cues.

is a set of responses for a listener who only makes use of fricative noises: both their sets of responses are the same, regardless of the transition following the fricative noise; Graph (B) is a set of responses for a listener who only makes use of transitional cues: the two continua are not perceived as continua, instead all the stimuli with /s/–transitions are perceived as /s/–vowel and all the stimuli with /ʃ/–transitions are perceived as /ʃ/–vowel.

It should be noted, however, that it is extremely unlikely that a listener would make *sole* use of one acoustic cue when others were present, unless such a cue was extremely unambiguous. A more realistic set of responses are those actually obtained by Nittrouer & Studdert-Kennedy (1987) shown in Figure 2.5. These graphs show the percent /s/ responses given to two /ʃu/–/su/ continua by the oldest and youngest listeners in the study: adults and 3–year–old children.

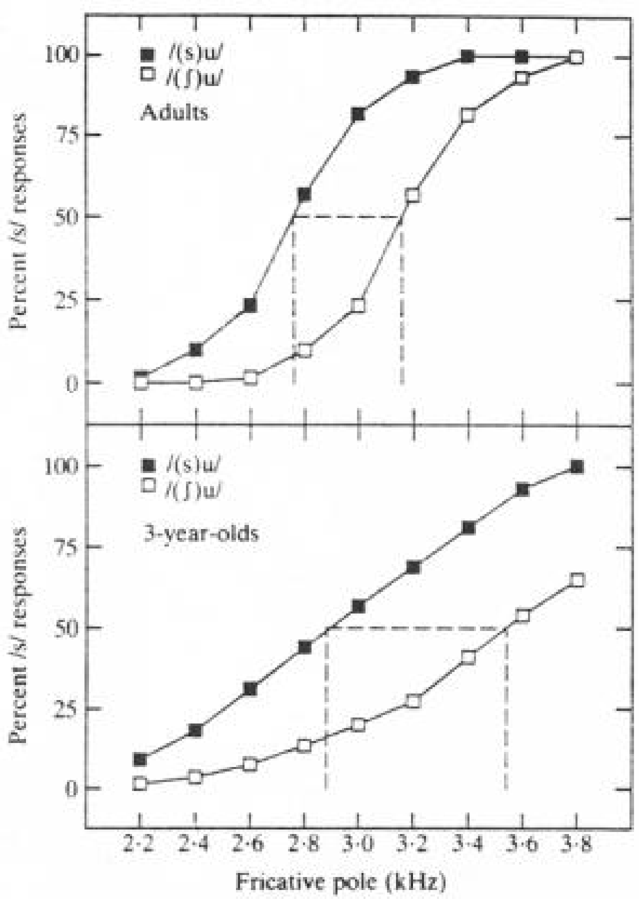

Figure 2.5: **Perceptual response curves from Nittrouer & Studdert–Kennedy 1987**. Note that the IPA symbol in parentheses indicates the transitional context of the continuum. Figure from Nittrouer (1992, p. 352) ©Academic Press Ltd. Reprinted by permission.

As can be seen, neither of the age groups show responses like the completely categorical or completely a–categorical responses illustrated above. Instead both groups of listeners display response curves which are roughly half–way between the two extreme types of response, suggesting that both groups are affected by both the fricative noise and the vowel formant transitions.

It is important to note, however, that the response curves of the two groups, while not as extremely different as the hypothetical responses in Figure 2.4, are not actually the same. This suggests that the children and the adults are not using the fricative noises and the formant transitions to the same extent. First, the adults' responses are relatively more categorical than the children's, to the extent that the extremely unambiguous fricative cues at the endpoints of the continua engender clear /s/ or /ʃ/ responses regardless of the transitional cue that follows. Additionally, the adults' two category boundaries are relatively less separated than the children's. Contrastively, the children's response curves are relatively less categorical than the adults' (note that the endpoints of the continua only engender clear /s/ or /ʃ/ responses when the transition that follows cues the same percept as the fricative), and their category boundaries are relatively more separated. According to Nittrouer & Studdert-Kennedy (1987) this combination of results indicates two things: i) the fact that the adults' response curves were more categorical than the children's indicates that the adults were using the fricative noise more heavily than the children, and ii) the fact that the children's response curves were more separated than the adults' indicates that the children were using the transitional cues more heavily than the adults. The validity of this conclusion will be discussed further in Section 1.2.

In order to be able to make statistical analyses of these qualitative observations, Nittrouer & Studdert-Kennedy (1987) also made quantitative measures of the relative influence of fricative noise and vowel formant transitions on listeners' judgments. Three measures were calculated from the listeners' response curves using a normalising equation (described in more detail in Chapter 3). The first of these is a measure of the *slope* of the response curve, which is calculated twice for each listener or set of listeners: one measure for each response curve. The slope of a response curve can be taken as a reflection of the degree of 'categorical–ness' of the responses: i.e. the rate at which the listener changed from one category label to the other. Examining the response curves of the adults and children from Nittrouer & Studdert-Kennedy (1987) displayed in Figure 2.5, it can be seen that the adults have relatively steeper slopes, meaning that it took fewer changes in fricative noise frequency for them to change from predominantly /ʃ/ responses to predominantly /s/ responses. The children, on the other hand, have relatively shallower slopes, meaning that it took a larger number of changes in fricative noise frequency for them to change from /ʃ/ to /s/ responses. This age difference in slope values was seen for all of the listeners in the study: response curves became steeper with increasing age. The two youngest groups (3– and 4–year–olds) had the shallowest slopes, the 5–year–olds had intermediate slopes, and the 7–year–olds and adults had the steepest slopes. These differences were significant, both between the youngest children (3– to 5–year olds) and the oldest groups (7–year–olds and adults), as well as between the very youngest groups (3– to 4–year–olds) and the 5–year–olds.

The second measure taken from the response curves is the *mean* of the responses. This is seen as corresponding to the 50% point on the response curve: that is, the point at which 50% of the responses are /ʃ/ and 50% of the responses are /s/. Described in another way, this is the point at which the listener places the category boundary, although it should be noted that because it is calculated by normalising the response curve values, this point may not always coincide exactly with one of the points on the continuum. In Figure 2.5 these points are marked by vertical dotted lines. Again, as for the slope, there will be two values per listener or listener group for *mean*—each corresponding to one response curve.

The third and final value is not calculated directly from the listeners' response curves, but from the means of the two response curves. This third value is the *separation* of the response curves, and is calculated by taking the difference between the means of each listener's two response curves. Separation is seen as corresponding to the degree of 'transition effect', that is, the degree to which the category boundary placement is affected by the presence of the two different sets of formant transitions—one would expect, for instance, to see "more 's' responses to tokens with /s/ vocalic transitions" (Nittrouer & Studdert-Kennedy 1987, p. 326). Only one separation value is calculated per listener or listener group. In Figure 2.5, the separation for the adults and the children is marked by a horizontal dotted line (between the two vertical lines marking the means): this line is longer for the children's responses than for the adults' responses. The calculation of this measure allowed Nittrouer & Studdert-Kennedy (1987) to show a transition effect for all age groups. Additionally, these authors were also able to show that the effect decreased with increasing age: a significant difference in transition effect was again found between the youngest and the oldest groups, as well as between the very youngest children (3– to 4–year–olds) and the 5–year–olds.

Nittrouer 1992: Experiment 3

Nittrouer employed essentially the same methodology as described above in her 1992 study of the weighting of across–syllable transitions. However, the relationship between the two cues to the contrast examined in the 1992 study is slightly different to the relationship between fricative noise frequency and formant frequency in the /s/ vs. /ʃ/ contrast. The contrast examined in the 1992 study is essentially a /d/–/g/ contrast, but in this case the /d/ and /g/ were embedded in the following two–syllable non–words: ‘arda’, ‘arga’, ‘alda’ and ‘alga’, as opposed to CV syllables. As noted in Chapter 1 (Section 1.3), a study by Mann (1980) has shown that adults are affected in their perception of the CV syllables /da/ and /ga/ by the transitional influences of preceding VC syllables /ar/ and /al/. A more detailed examination of the acoustic characteristics of these VCCV words should make the reason for this clear. In isolation, one of the main differences between the syllables /da/ and /ga/ is the configuration of F3 at the onset of the vowel: as shown in Figure 2.6, F3 falls after the release of /d/ but rises following the release of /g/. However, when these two syllables are preceded by /ar/ or /al/ the configuration of F3 is influenced by the place of articulation of the liquid, in much the same way that the configuration of the vowel onset formants following the fricatives (in the study discussed above) is influenced by the place of articulation of the fricative. Both the stop consonants /d/ and /g/ and the liquids /l/ and /r/ have contrasting places of articulation: /d/ and /l/ are alveolar sounds and are therefore formed relatively farther forward in the mouth, while /g/ and /r/ are velar sounds and are therefore formed relatively

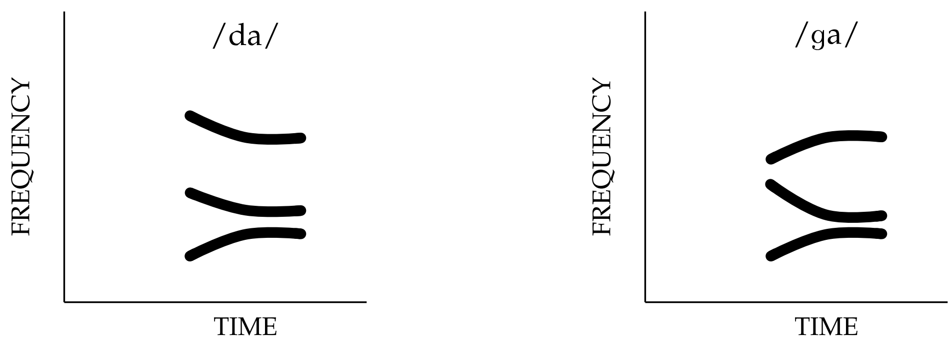

Figure 2.6: **Stylised spectrogram of /da/ (left) and /ga/ (right).** Note that F3 falls following /d/, but rises following /g/.

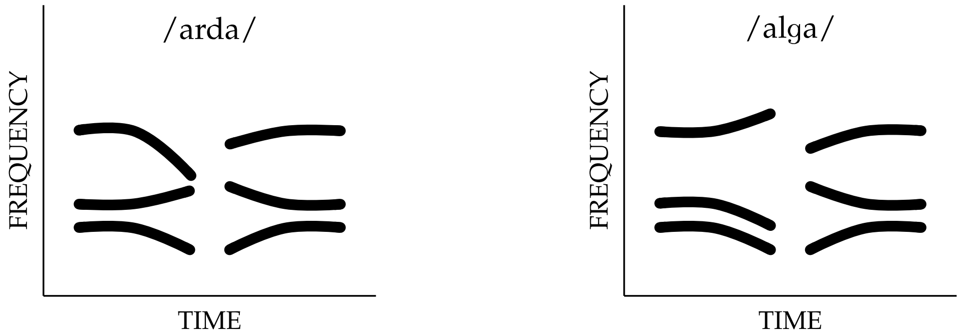

Figure 2.7: **Stylised spectrogram of /arda/ (left) and /alɡa/ (right).** Note the similar onset configurations of F3 in the second syllables as a result of coarticulation.

farther back in the mouth. The effect of articulating these sounds immediately after each other is to shift the place of articulation of the constrictions: /d/ is formed farther back after /r/ than after /l/, while /ɡ/ is formed farther front after /l/ than after /r/.

These assimilatory effects create a situation in which there is the potential for perceptual confusion: the vowel onset transitions in /da/ preceded by /ar/ and those in /ɡa/ preceded by /al/ are very similar in their configurations, as displayed in Figure 2.7, and are therefore ambiguous cues to the /d/–/ɡ/ contrast (see e.g. Lotto & Kluender 1998). However, Mann (1980) found that when these ambiguous CV syllables followed /al/, more of them were identified by adult listeners as /ɡa/, while the same ambiguous CV syllables following /ar/ were identified more often as /da/. Mann interprets this as meaning that these listeners were able to compensate for the lack of acoustic cues to differentiate between the two CV syllables by using knowledge of the way in which a preceding syllable would influence the articulation of the stop: if the syllable followed /ar/ then it was perceived as a /da/ which had been formed further back in the mouth than it would have been in isolation; if it followed /al/ then it was a /ɡa/ which had been formed further forward.

Nittrouer (1992) replicated this study with children, in an effort to determine whether this across–syllable transitional information is used in perception by younger listeners. For this study, the syllables /da/ and /ɡa/ were synthesised on a continuum in which the parameter which varied was the onset transition of F3. To the beginning of each of these syllables were added four other syllables—two /ar/ and two /al/. These four initial syllables were spliced from natural tokens of the pseudowords /alda/, /arda/, /alɡa/, /arɡa/. Each therefore retained transitional information about the CV syllable that they preceded.

According to Nittrouer (1992) children's responses to these stimuli were much *less* affected than adults' by the transitions between the two syllables. Nittrouer interprets this finding as support for the syllable–to–phoneme aspect of the DWS model. This study shows, she claims, that adults make use of whatever cues will give them the most information about individual segments—in this case the information regarding the across–syllable influence of liquids on stops. Children, says Nittrouer, use the information provided by across–syllable transitions less than do adults, because these do not give them information about their perceptual units, the syllables. The issues regarding this conclusion will be discussed in more detail in the following section.

1.2 Issues relating to the DWS model

The studies described above provide the core perceptual support for Nittrouer's Developmental Weighting Shift model. However, the development of acoustic cue weighting is slightly more elaborate than simply a movement from heavier weighting of syllable–internal dynamic cues to heavier weighting of segment–internal static cues, as a closer examination of Nittrouer & Studdert-Kennedy (1987), Nittrouer (1992) and the studies that followed them will show. There are a number of issues regarding Nittrouer's studies, and the conclusions she has drawn from them. Some of these issues have been raised by Nittrouer and colleagues themselves, others have been raised by other studies, however each has required a re–evaluation and/or elaboration of the original definition of the Developmental Weighting Shift model.

Limited experimental evidence

Perhaps one of the most problematic issues regarding the Developmental Weighting Shift model is the possibility that the results on which it is based might be due, not to a general developmental property of the perceptual system, but rather to certain acoustic characteristics of the contexts which have thus far been examined. In the two studies discussed to this point, Nittrouer and colleagues have shown a developmental weighting shift to occur for an /ʃ/–/s/ contrast, and a different shift to occur for a /d/–/g/ contrast, when preceded by /r/ and /l/. The question is, do these same shifts occur in any other phonetic context? Unfortunately, it is difficult to answer this question because shifts in cue weighting have been shown to occur for only a very limited range of phonetic contrasts.

The issues arising from this problem can be broken down into three slightly different problems. The first revolves around the fact that even within the limited number of contrasts that *have* been tested, a difference in transitional effect has been seen for different vowels. Specifically, the size of the transitional effect—that is, the amount that the response curves were separated— was greater for transitions from fricatives into /u/ than for transitions from fricatives into /i/. The effect of these different vowel contexts on the transition effect was evident for both child and adult listeners. However, while the effect was proportionally the same for both groups (the transitional effect was about 1/7 as large in the /i/ context as in the /u/ context, Nittrouer 1992), the *absolute* extent of the vowel effect was much greater for the children, due to the fact that their transitional effect was much larger than the adults' in the first place.

Nittrouer & Studdert-Kennedy (1987) hypothesise from this result that "the perceptual weight given to transitions is proportional to their extent" (p. 326). By *extent* the authors mean the duration of the transition and the amount of change in frequency from the onset to the offset of formant movement, which will differ depending on the physical distance to be traveled by the articulators from one constriction to the next. Both /s/ and /ʃ/ are constrictions which have a place of articulation near the front of the mouth. If these two fricatives are followed by an /u/, which is a back vowel in American English, then the distance to be travelled is relatively large—thus /u/ following a fricative has relatively extensive onset transitions. If the fricatives are followed by /i/, which is a front articulated vowel, then the distance to be traveled is relatively short—thus /i/ following fricatives has comparatively less extensive onset transitions.

Nittrouer (1992), however, notes that because only two vowel contexts were examined in the 1987 study, one which engendered a strong transitional effect, and the other which engendered a weaker transitional effect, it is possible that responses to either of these could have been caused by some artifact of the synthesis process. If this were the case, then a possible reason for the difference in response pattern between the two vowel contexts might simply be that one of the two types of response pattern does not reflect normal speech perception behaviour. Therefore, in order to determine whether it is indeed the case that the strength of the transitional effect is determined by the extent of the transitions, Nittrouer (1992) conducted a further study of the /s/–/ʃ/ contrast, using the same /u/ context and an additional one: /a/, which is articulated in a relatively more back position than /i/ (although generally not as far back as /u/).

Nittrouer found in this later study that the /a/ context engendered a much more similar set of responses to the /u/ context than had the /i/ context in the 1987 study. In response to both the /a/ context and the /u/ context stimuli, children showed much shallower and more separated response curves than adults. The response curves for /u/ were still slightly larger than those seen for /a/, however the difference was not as large as that seen between /u/ and /i/. Nittrouer (1992) takes this result as further support both for her Developmental Weighting Shift model and for her hypothesis that transitional effects depend on the extent of the transitions.

The second problem has to do with a lack of consistency in phonetic contexts between Nittrouer & Studdert-Kennedy's (1987) study, and Experiment 3 in Nittrouer's (1992) study. The problem is that the comparison which Nittrouer makes between weighting of within–syllable transitions and across–syllable transitions is *also* a comparison between weighting of transitions in two completely different phonetic contexts. That is, not only were the studies comparing the use of transitions between segments, and transitions between syllables, but they were also comparing the use of transitions between a fricative and a vowel (the /su/–/ʃu/ contrasts, for example), and transitions between a liquid and a stop (the /arda/–/alɡa/ contrasts). The fact that acoustically comparable contexts were not chosen to test within– and across–syllable transition use, raises the possibility that the difference in cue weighting seen between these two studies has to do more with the phonetic contexts of the transitions than with syllable structure.

The last problem is simply to do with lack of extensive proof for proposals regarding the cues that are attended to by adults and children. First, the DWS model specifies that infants and young children always attend more to within–syllable transitional cues. This suggests that children are unable to make efficient use of non–transitional information, which in turn suggests that they should have difficulty perceiving contrasts in which the transitional cues are minimal or perhaps even non–existent. There is some evidence to suggest that this might be the case (see Section 1.3 below for a discussion), however it has also been found that children will attend less heavily to transitions (and more heavily to other cues) when those transitions are less extensive: Nittrouer and colleagues' studies above showed that children attend more to transitions between fricatives and /u/, than between fricatives and /i/. Unfortunately no contrasts have been tested by Nittrouer and colleagues in which transitional cues are naturally very poor cues to the contrast, therefore no conclusions can be drawn about the extent to which children are 'programmed' to attend to transitions.

The model is slightly more complicated as regards the cues which adults are meant to weight most heavily. As noted in the introduction to this chapter, Nittrouer and colleagues originally hypothesised that adults would make most use of static cues, like fricative noise, to contrast with the dynamic cues used most by children. A study by Nittrouer & Miller in 1997, however, led the authors to amend this hypothesis. This study examined perception of a /sa/–/ʃa/ contrast and a /su/–/ʃu/ contrast, in which some of the transitional information was neutralised. The aim of the study was to further investigate the importance of transitional cues to listeners, by effectively removing the information that they carried. The neutralisation of the transitional cues did indeed affect the listeners' percepts, but to a different extent for each vowel context. For non–neutralised stimuli, adults had previously shown a very small transitional effect (similar, but much reduced in comparison to that shown by the children). For the stimuli with the neutralised transitions, however, this effect was even more reduced—but only for those stimuli with an /u/ vowel: i.e. only for those stimuli with extensive transitions. The authors claim again that this is due to the fact that fricative–vowel transitions are more extensive at the onset of /u/ than at the onset of /a/ (as noted above). The more extensive a transition, the greater the effect it should have on perception, and consequently the more detrimental the neutralisation of that transition should be. The authors hypothesise that

> It should be the case for phonetic environments in which formant transitions are not particularly informative about segment identity (and, reciprocally, in which another property is especially informative) that the weight children assign to those transitions should decrease with language experience. However, for phonetic environments in which formant transitions are informative (and, reciprocally, in which other properties are not) the weight children assign to those transitions should continue to be substantial. (Nittrouer & Miller 1997*b*, p. 2265)

There are, however, problems with this hypothesis—in particular with the definition of 'informativeness.' Nittrouer & Miller (1997*b*) appear to assume that the indicator of the informativeness of a cue is the degree to which it is used by adults to make their labeling decisions. The authors state that Harris (1958)

> demonstrated that the [fricative] noise is the primary property on which adults make this fricative decision [a /s/ vs. /ʃ/ decision] (and so is presumably informative concerning fricative identity). Conversely, Harris showed that the noise is not weighted heavily by adults when the fricative is /f/ or /th/ (and so is presumably not informative). (Nittrouer & Miller 1997*b*, p. 2265)

Unfortunately, however, this results in rather a circular argument: adults attend more to certain cues than others because these cues are more informative; the way to determine of a cue is informative in perception is to ascertain if it is attended to by adults.

The solution to all three of these problems is one which has yet to be undertaken—specifically, an in depth investigation of cue weighting shifts in a more extensive range of contrasts, in terms of the variety both of phonetic context, and of syllabic structure. In order to address the question of whether such shifts occur across the board in perception, it would be necessary to examine cue weighting for contrasts for which the 'steady–state' and/or the transitional cues were much more or much less informative than they are for an /s/–/ʃ/ contrast. Possible candidates for this type of test include fricative contrasts for which the fricative noise is comparatively weaker in amplitude than the noises in /s/ and /ʃ/ (e.g. /f/, /θ/), or stop consonant contrasts for which the majority of information comes from the transition configuration rather than the 'steady–state' burst noise (e.g. /b/, /d/). This type of study is outside the scope of this thesis, however, these concerns will be borne in mind in the design of the stimuli to be used in this study.

Auditory processing

The question asked in the previous section was whether the results of Nittrouer and colleagues' studies might be due, not to a developmental perceptual phenomenon, but rather to characteristics of the contrasts examined. The question asked in this section is whether the results might indeed be due to a developmental perceptual phenomenon—but *not* the phenomenon that Nittrouer proposes. A number of researchers (e.g. Elliott, Hammer, Scholl & Wasowicz 1989, Sussman 1993, Sussman & Carney 1989) have proposed that "immature sensory processing" (Sussman 1993) in children might be the source of the differences between adults' and children's perceptual behaviour. In other words, the phenomenon observed is not simply related to the manner in which children learn to process speech, but rather is related to children's developing auditory systems, and their ability to process all sounds.

J. Sussman (1993) suggests that the reason for both Nittrouer & Studdert-Kennedy's (1987) results, and their explanation of these results, is the fact that Nittrouer & Studdert-Kennedy only collected "phonetic identification data" (p. 392)—that is, data on identification of stimuli by means of phonemic labels. As noted in Chapter 1, Section 1.3, there is some question as to whether speech labeling and speech discrimination access the same perceptual processes. J. Sussman claims that phonemic labeling tasks do not allow for any assessment of subjects' sensitivity to fine–grained acoustic differences between speech stimuli. The type of test that *does* allow for this type of assessment is, according to Sussman and various other researchers (e.g. Elliott et al. 1989), a discrimination test. In this type of test the listener is asked to differentiate between two sounds which differ to varying degrees along a particular parameter. The object of the test is to determine the smallest degree of difference that a listener can successfully discriminate (this degree is known as a *just noticeable difference* or j.n.d.). The results of such a test give a measure of the listeners' auditory sensitivity to the varying parameter. In a criticism of other studies which make use of labeling rather than discrimination tasks (specifically, Tallal & Peircy 1973, Tallal & Peircy 1975), J. Sussman (1993) suggests that the performance of subjects on such a task "may reflect their ability to label the tokens differentially, rather than their ability to differentiate the lower level acoustic information in the stimuli" (p. 1287).

In response to criticism from the auditory sensitivity camp, Nittrouer and colleagues have carried out a number of tests of children's and adults' discrimination abilities. The first study, carried out by Nittrouer (1996*a*), compared labeling of fricative–vowel stimuli (of the same design as those used in Nittrouer 1992) with discrimination of stimuli which varied along a fricative noise spectrum, and discrimination of stimuli which varied in F2 transition configuration. The results of this study did show a difference between labeling and discrimination behaviour: children had shallower and more separated labeling response slopes than the adults (much as had been found in Nittrouer's previous studies), but required greater differences in fricative frequency and transition frequency in order to discriminate the stimuli than did the adults. The discrimination results on their own would seem to indicate that children are indeed less sensitive to changes in both these parameters. A later study by Nittrouer & Crowther (1998) found similar results. This second study examined discrimination of non–speech stimuli: both dynamic spectral (glide or transition) stimuli and static spectral (steady–state) stimuli. Again it was found that children required larger differences than adults to discriminate both types of stimuli, and that both groups required larger differences to be able to discriminate dynamic spectral stimuli than static spectral stimuli.

These results do support some sort of theory that children's auditory sensitivity must develop to a certain level in order to be able to detect fine–grained differences in stimuli. However, as noted by Nittrouer & Crowther (1998), the results do not in any way explain the shifts in acoustic cue weighting seen in Nittrouer and colleagues' previous studies. In order to be able to explain the differences in *labeling* behaviour by reference to differences in auditory sensitivity, discrimination studies would have to show that children were less sensitive to acoustic differences in fricative–like stimuli, and more sensitive to differences in transition–like stimuli. That is, the children's apparent preference for transitions in cue weighting tasks would have to be borne out by a greater sensitivity to this cue in discrimination tasks. As the discrimination results show the opposite behaviour, Nittrouer & Crowther (1998) conclude that while adults and children may show differences in auditory sensitivity, these differences cannot be the source of adults' and children's differences in cue weighting behaviour.

At this point it is not clear why two such different sets of results should be found for two tests which are assumed to both tap speech perception abilities. The possibility that these conflicting results may be the result of more than one perceptual phenomenon will be discussed further in Chapter 6.

Synthetic vs. natural stimuli

As noted above, Nittrouer & Studdert-Kennedy (1987) made use of 'hybrid' stimuli, which were a combination of synthetic fricative noises, and naturally produced vowel portions (transition–plus–vowel target). However, as noted by Nittrouer & Miller (1997*b*, p. 2254), "it is possible that children show a greater weighting of vocalic formant transitions than adults, while demonstrating a lesser weighting of the fricative noise, because children fail to process synthetic speech components as they do natural speech." In other words, there is a possibility that the children attended to the transitions in the 1987 study not because of their 'transitional' nature, but because of their naturalness. Nittrouer

& Miller (1997*b*) attempted to address this issue by replicating the original perceptual weighting study using wholly synthetic stimuli. In general these stimuli engendered a slightly more categorical response than the hybrid stimuli from both adults and children. Despite this, though, the general developmental trend of less categorical, more separated curves in the younger listeners was nonetheless replicated.

However, there remains a possibility that the results seen in this later study could again be due to the nature of the stimuli. The synthetic vowel portions used by Nittrouer & Miller (1997*b*) were ostensibly based on a combination of parameters provided by Whalen (1981), and values gained from acoustic analysis of natural speech vowel portions similar to those used in the original study. However, while the transition onset and offset frequencies were taken from frequencies in the natural tokens, the values for all the frequencies *between* these two points were linearly interpolated. This means that although the beginning and end values of the formant transitions were realistically modeled, the overall shape of the transitions was highly stylistic. Hazan & Rosen (1991) have suggested that in highly simplistic stimuli of this sort, the stylised aspects of the stimuli may be more salient than the equivalent, more complex cues present in natural speech, and thus may be potentially more likely to draw some listeners' attention.

Slope of response curves

As noted above, the perceptual response curves obtained for these studies are analysed in terms of two values. The first is the separation of the two response curves, which is taken as an indicator of the extent to which the listener's responses have been influenced by the transitional cues which differ between the two speech continua. The second measure is the slope of the response curves, which is taken as an indicator of the degree to which the listener's responses have been influenced by the fricative noises. Nittrouer and colleagues consider these two values to be the result of the same phenomenon: the degree to which the listener attends to either syllable– or phoneme–specific information. It is possible, however, that the slope, or categorical–ness of the response curves may not only be a result of the listener's degree of attention to the fricative noises. The first possibility is that the children's shallower response curves are due to difficulty in maintaining attention to the perception task. A sharply categorical response curve is the result of a listener giving extremely consistent responses (e.g. 8–10 out of 10) to stimuli on either side of a category boundary. This consistency is due partly to the categorical perception phenomenon, but also naturally requires that the listener maintain steady attention to the task. The response curve of a listener who is not paying attention to the task, and is consequently giving inconsistent responses, will be much shallower, and in extreme circumstances may even sit on or around the 50% point (as the listener has a 50% chance of randomly choosing one or the other of the two available labels). If the children who had participated in Nittrouer's early studies had been easily distracted from the perception task, it is quite possible that their responses would be less consistent, and their response curves therefore more shallow than the adults'. Nittrouer (1992) therefore decided to test children's transitional cue weighting in a context which would engender age–related differences in perceptual behaviour in terms of boundary placement, but not in terms of slope of response curves.

The context chosen to test the source of the children's shallower response curves was one which had been used by Morrongiello et al. (1984) to examine weighting of cues to the presence of a post–fricative stop consonant, as in 'say' vs. 'stay'. The two cues to this contrast are the duration of silence (or 'gap duration') following the fricative noise, and the configuration of the vowel onset formant transitions. As in the /ʃ/–/s/ studies described above, the stimuli used by Nittrouer (1992) were designed so that one cue varied along a continuum, while the other cue varied in one of only two configurations. For this study, the length of the post–fricative gap was varied from a duration which strongly cued the absence of the stop, to one which strongly cued the presence of a stop. Each of these gap durations was combined with the two vowel formant transition configurations: one which was appropriate for having followed /s/, the other which was appropriate for having followed /st/.

It should be noted at this point that these transitional cues are slightly different from those used in the 1987 study, in that they are not each unambiguous cues to different places of articulation. Both /s/ and /st/ are formed with a relatively closed vocal tract, with both closures occurring at the alveolar ridge. The formant transitions which follow these noises will reflect this: both the transition configurations will indicate that they have followed some sort of alveolar closure. However, because of the slight difference in manner of articulation between /s/ and /st/ (/st/ involves a complete closure of the vocal tract, while /s/ involves an incomplete closure) the degree to which these two cues indicate a complete closure (i.e. a stop) will differ. Morrongiello et al. (1984) found that adults classified /eɪ/ vowel portions with transitions which were appropriate for having followed /st/ as 'day' with 100% accuracy[4], while they classified /eɪ/ vowel portions with transitions which were appropriate for having followed /s/ as 'day' only approximately 50% of the time, and as 'ay' the other 50% of the time. It has also been found (as has been found for other contrasts) that the strength of these transitional cues in signalling the presence of a stop interacts with the gap duration cue: Best et al. (1981) found that for adult listeners,

> If unequivocal spectral [i.e. transitional] information about the occurrence of a medial /t/ is provided, listeners hear 'stay' when the duration of a silent gap between /s/ and the vocalic syllable minimally specifies a stop closure. However, when spectral information provides only equivocal information about an alveolar stop, listeners need stronger evidence for stop closure from another acoustic cue (e.g., longer closure gap) in order to perceive 'stay'. (p. 205)

What Nittrouer found in her 1992 study of this contrast was that, for the stimuli with ambiguous transitional cues, the children's response curves were significantly less separated than the adults'. The interpretation of this is that the children needed a much smaller amount of silence than did the adults to perceive "stay" from these ambiguous transitional cues (which is essentially the same conclusion drawn by Morrongiello et al. 1984). Nittrouer proposes that the reason for these results is once again the fact that young children give more perceptual weight to transitions than do adults. For adult listeners, the ambiguous transitional cue, as found by Morrongiello et al. (1984), was labeled as 'ay' or 'day' essentially at chance when presented in isolation, and required a larger amount of silence to be perceived as 'stay' when presented in the context of frication+silence. For the children in Nittrouer's (1992) study, however, this ambiguous cue required less support from the gap duration than was needed by the adults, indicating (according to Nittrouer) that the ambiguous transitional cue was worth more towards a stay response for children—i.e. they gave it more weight perceptually (see also Nittrouer, Crowther & Miller 1998, for a replication of these results with the addition of a burst cue).

[4] In English one of the main differences between voiced and voiceless stops is in degree of aspiration—/d/ is unaspirated; /t/ is aspirated. However, unvoiced stops following /s/ are unaspirated, thus 'stay' with the fricative noise removed will sound more like 'day' than 'tay' (Ladefoged 1993).

Importantly, however, Nittrouer (1992) also found that although young children's response curves for stimuli with *ambiguous* transitional cues were shallower than adults', children's response curves for stimuli with *unambiguous* transitional cues were similar in slope to the adults'. Nittrouer states that if children's shallower responses to this type of stimuli were caused by lack of attentiveness to the task, they should have displayed equally shallow responses for both sets of stimuli. Nittrouer therefore concludes that she is justified in considering the slope and the separation of the response curves as resulting from the same perceptual phenomenon.

However, another possible explanation for the children's shallower response curves compared to the adults is that this gradual change from a shallow slope to a steep slope is a characteristic, not of a change in fricative cue use, but of the development of labelling consistency. It may simply be that children's ability to categorise speech sounds improves as they mature. Simon & Fourcin (1978) carried out a study of English and French children's perception of voicing contrasts based on a combination of durational cues (long vs. short VOT) and spectral cues (for the short VOT: rising or flat F1 transition, for the long VOT: no F1 transition). What these authors found was that although English– and French–learning children differ in the rate at which they learned to successfully make use of VOT, both groups appear to go through three basic developmental stages of stimulus labelling in terms of this cue. Simon & Fourcin (1978) showed that i) very young children label unambiguous stimuli (i.e. continuum endpoints, which are stimuli with extreme VOT values) reasonably clearly, but give quasi–random responses to ambiguous stimuli (i.e. those stimuli between the endpoints on the continuum), ii) slightly older children label stimuli 'progressively', giving gradually more of one label and less of the other as they move along the continuum, and iii) older children label stimuli in a basically categorical manner.

Support for this explanation comes from a study by Hazan & Barrett (1999). These authors examined the perceptual weighting of both transitional and non–transitional cues to 5 CV contrasts: /g/–/k/, /d/–/g/, /s/–/z/, /s/–/ʃ/, /tʃ/–/ʃ/. The child subjects in this study were older than any of the child subjects examined in Nittrouer's perceptual studies, ranging from 6 through 12 years of age. The results of the study showed that response curves continued to get progressively steeper from 6 through 12 years, and that even the oldest children did not display response curves which were as steep as those of the adults. As noted by Hazan & Barrett (1999), this supports the claim that "sharpness of categorisation, and hence labelling consistency of phonemic contrasts continues to increase until adulthood" (p. 2496). Contrastively, these studies did not find any conclusive evidence that children from 6 through 12 years make any greater use of transitional cues than do adults. The authors point out that this does not contradict Nittrouer's Developmental Weighting Shift model, as the shift from heavier use of transitional cues to heavier use of fricative cues is presumed to have taken place by that point in development. However, this result does raise the question as to whether the slope of a listener's response curve and the separation between the listener's two response curves are really the result of the same perceptual phenomenon, as proposed by Nittrouer. If these two measures were as strongly linked as has been suggested, one would not expect to find one of them continuing to develop after the other has ceased to develop.

1.3 Other studies which have examined the development of transitional cue use

There are a number of other studies which lend support to Nittrouer's theory that children have an initial perceptual preference for dynamic spectral cues. Some of these have directly evaluated infants' and children's use of transitional cues in comparison with other cues. Nittrouer et al. (1998) also note that indirect evidence for the view presented in the DWS model can be found in studies which show that children have difficulty perceiving contrasts when the information provided is non–transitional only.

As noted above, Morrongiello et al. (1984) found the same age–related differences in the weighting of two cues to the presence of a stop consonant (as in 'say' vs. 'stay') as did Nittrouer (1992). That is, it was found that children needed much less silence than did adults to compensate for a transitional cue which only weakly cued the presence of a stop consonant. These authors drew the same conclusions as did Nittrouer: that transitional cues, even when weak, are more important to children than to adults.

A set of studies by Ohde and colleagues (e.g. Ohde 1994, Ohde, Haley & McMahon 1996, Ohde & Haley 1997), for instance, have also shown a developmental shift in the use of transitional cues, in this case cues to place of articulation of stop consonants in CV syllables, and to vowel quality in certain segmental contexts. The three cues varied in these studies were formant transitions, noise bursts, and voicing duration. These studies made use of a slightly different methodology to that used by Nittrouer and colleagues, in that none of these three cues were varied along a continuum. Instead, for each cue there were two potential conditions, one 'informative', the other 'uninformative'. For the formant transitions, these two conditions were 'moving' or 'straight' (the straight condition was assumed to be less informative than the moving condition); for the noise burst the conditions were 'present' or 'absent'; and for the voicing duration the conditions were '46ms' of voicing or '10ms' of voicing. The results of these studies showed that 3– to 4–year–old children had difficulty identifying consonants when the formant transition cues were attenuated ('straight' formant condition), while formant motion (or lack thereof) had little effect on the responses of either older children (5 to 11 years) or adults. This effect was particularly pronounced for perception of the velar consonant [g]. Interestingly, for the identification of vowel quality, younger children were again more influenced by formant transitions than older children and adults, but this time only for vowels in the context of a velar consonant.

Parnell & Amerman (1978) also examined adults' and children's use of two types of cue to stop place of articulation: burst–plus–aspiration, and vowel formant transitions. This study made use of natural speech stimuli from which the burst and aspiration, the formant transitions, and the vowel target formants were variously excised. The older children (11 years) and the adults were able to identify the place of articulation of the stops from the stimuli which contained only burst and aspiration information. The younger children (4 years), on the other hand, performed at chance on these stimuli. When transitions were added to the stimuli with only burst and aspiration information, however, these younger children performed well above chance. This would again seem to suggest that the transitional information was more important to the younger children than the older children or adults.

A study by Lacerda (1992) has extended the study of cue weighting to infant perception abilities. Making use of a high amplitude sucking technique (see Chapter 1, Section 2.3), this study examined infants' (age 16 to 230 days) and adults' ability to discriminate place of articulation distinctions based on transitional cue differences only. The stimuli used in this study were CV syllables which consisted only of formant transitions and steady states, with the transitions in various configurations (straight, low and rising, high and falling) and in both vowel initial (CV) and vowel final (VC) position. Lacerda reports that while the adults were least able to discriminate those stimuli with a maximum transition rate, the infants' discrimination performance improved as transitions became more dynamic, indicating the importance of these cues to young listeners.

Watson (1997) replicated Nittrouer & Studdert-Kennedy's (1987) /s/–/ʃ/ study with groups of normal adults, normally developing children, and children with expressive phonological disorders. This study found that older phonologically disordered children had more global cue weighting strategies than their normally developing peers—that is, they weighted syllable–internal transitions more heavily than did the normal children. Additionally, when reverberation was used to mask the transitional cues, it was found that while the adults and the older normal children were able to increase the weight they gave to the fricative frequency cue, the younger normal children, and the phonologically disordered children were not able to shift from heavier weighting of the (now much less salient) transitional cue. The results of this study again support the hypothesis that young children's cue weighting strategies are indeed more global than older children's and adults'. The study also suggests that the shift in cue weighting strategies observed by Nittrouer and colleagues might have something to do with the development of an ability to become more flexible in cue use.

Finally, three studies of acoustic cue use in the perception of voicing in final stops (Greenlee 1980, Krause 1982, Wardrip-Fruin & Peach 1984) seem to indicate that younger children have more difficulty identifying voicing when no transitional information is provided. Greenlee (1980) used a deletion method to create stimuli in which the only cue to the voicing of the final consonant was the duration of the preceding vowel. The youngest children tested (3 years) were unable to identify the stimuli on the basis of only vowel duration, however their identification improved for stimuli which had not had transitional and other cues to voicing deleted. The older children (6 years) had less difficulty identifying the stimuli as voiced or voiceless based on just the vowel duration, but this group also benefitted from the addition of other cues to voicing. Wardrip-Fruin & Peach (1984), who used a similar deletion technique, also found that 6–year–old children had much more difficulty making a voiced/voiceless judgment when they were not provided with transitional information. Additionally these authors found that when this group of children *were* provided with transitional cues, they weighted them more heavily than did adults. Krause (1982) investigated perception of the same voiced/voiceless contrast, but in this case using synthetic stimuli which varied along a continuum from a short to a long vowel duration. Krause found that the youngest children that she tested (3 years) required a longer vowel duration to perceive voicing than either the older children (6 years) or the adults, indicating that they weighted changes in this cue much less than older listeners. Interestingly, Krause also notes that a certain number of children were influenced in their labeling decisions by one of the non–test characteristics of the stimuli, namely the presence or absence of an F1 transition. There was a tendency in these children to label those stimuli with an F1 transition as voiced, and those without as unvoiced, regardless of the duration of the vowel, once again indicating the potential importance of transitional cues to younger listeners.

It would appear, then, that despite the many questions raised regarding the Developmental Weighting Shift, one premise on which it is based—that children have a perceptual preference for within–syllable transitions—has experimental support, at least for a limited range of contrasts.

2 The development of phonemic awareness

As noted in Chapter 1 (Section 2.5), phonemic awareness is one of a number of skills which fall under the umbrella of phonological awareness skills—i.e. those metalinguistic skills which involve conscious awareness of 'meaningless' units of speech. These 'meaningless' units include larger units such as syllables and onset–rime units, and smaller units such as phonetic features, as well as phonemes. However, as will become clearer in this section, awareness of phonemes is not simply another in a list of metaphonological skills that a child acquires over the course of development. Studies by I. Liberman and colleagues in the 1970s showed that the development of phonemic awareness is strongly linked to the acquisition of alphabetic literacy skills. These studies found that children at different stages of reading acquisition performed differently on tests of phonological awareness. Specifically, children who could not yet read were only aware of the phonological structure of words to the level of syllables, while children who were reading were aware of phonological structure to the level of phonemes (e.g. Liberman et al. 1974). Thus the development of phonemic awareness involves a shift from syllables to phonemes which (as mentioned above) appears to parallel the shift that Nittrouer and colleagues have suggested occurs at a perceptual level. Before we examine the connection between these two shifts, however, we will first examine the development of phonological awareness as a whole, and establish how this development relates to literacy acquisition.

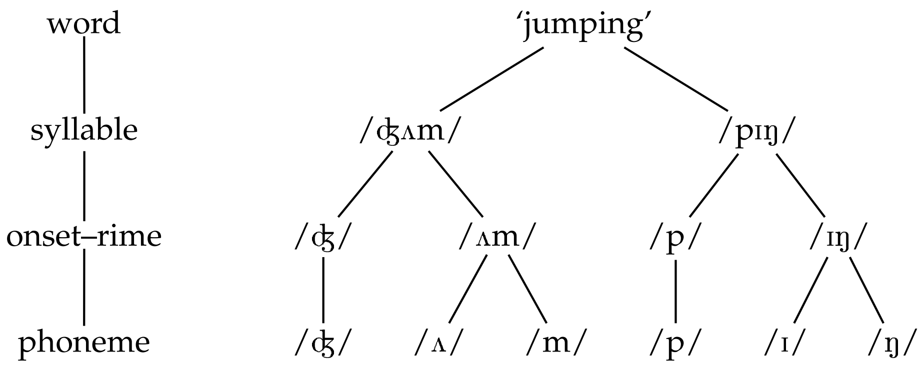

Figure 2.8: **Hierarchical model of the phonological structure of 'jumping'.**

2.1 What is phonological awareness?

Phonological awareness skills encompass a wide variety of metalinguistic abilities. Specifically, phonological awareness is defined as an ability to consciously think about and manipulate speech sounds which are smaller than a morpheme (this is often more likely to mean conceptually smaller rather than physically smaller). Therefore, a child who is able to recognise that 'cat' and 'hat' rhyme, or is able to produce 'frog' as a rhyme for 'log' is said to possess phonological awareness, as is the child who can say that 'pat' is the odd one out in the list 'mud, miss, pat, moon,' and the child that is able to say that the sounds in the word 'sheep' are /ʃ–i–p/. However, a number of studies have found that these different metalinguistic analyses are not equivalent in terms of the ease with which children can carry them out. This has led researchers to suggest that phonological awareness is a heterogeneous set of skills which develop in stages, rather than one homogeneous skill which develops gradually (Bertelson & de Gelder 1991, Goswami & Bryant 1990, Morais 1991, Treiman & Zukowski 1991, Treiman & Zukowski 1996). Treiman & Zukowski (1991), for example, claim that phonological awareness can be broken up into levels. One way of seeing these levels is in terms of 'linguistic' level. By this Treiman & Zukowski mean level of linguistic unit in a hierarchical sense: words, which occur at the top level, can be broken down into syllables, and syllables can be broken down into onset and rime units[5], and finally all of the above can be broken down into phonemes (see also Gussenhoven & Jacobs 1998), as illustrated in the tree diagram in Figure 2.8.

[5]In theories of syllable structure, rimes may also be broken down into a nucleus (also known as a peak) and a coda (e.g. Gussenhoven & Jacobs 1998), however these units have not been proposed as valid metaphonological units.

Therefore, a child who is able to access syllables, for example, will be able to say that the word 'elephant' has 3 syllables. A child that is able to access onset/rime units will be able to segment the word 'phone' into /f/ (the onset) and /on/ (the rime), and will be able to to produce and recognise words that rhyme with 'phone', or that have the same onset. Finally, a child that has good phoneme awareness will be able to name all of the component sounds in a word: 'cat', for example is /k–æ–t/, while 'sneeze' is /s–n–i–z/.

Treiman & Zukowski (1991) note that "children's performance depends on. . . the linguistic level that the task taps. For example, tasks that require children to segment speech at the level of words seem to be easier than tasks that require children to segment speech at the level of phonemes" (p. 67). Here the authors cite a study by Fox & Routh (1975) in which 3– to 7–year–old children were asked to say 'just a little bit' of a sentence, a word, or a syllable: i.e. just a little bit of the sentence 'Peter jumps' is the word 'Peter,' just a little bit of the word 'Peter' is the syllable 'Pete,' and just a little bit of the syllable 'Pete' is the phoneme 'P.' Fox & Routh (1975) found that the children's success at sentence and word level tasks was good, however the phoneme task was found to be more difficult for the younger children. Treiman and Zukowski go on to suggest in a later paper (Treiman & Zukowski 1996) that the reason for the difference in ease with which different linguistic units can be analysed

> is that the ability to segment speech into higher–level phonological units develops earlier than the ability to subdivide these units into their lower–level constituents. According to this hypothesis, children first gain the ability to segment speech into words. They next become able to divide words into syllables, then syllables into intrasyllabic units, and finally intrasyllabic units into phonemes (Treiman 1992). (Treiman & Zukowski 1996, p. 194)

While not all researchers agree that conscious analysis of words into intrasyllabic units (i.e. onsets and rimes) constitutes a separate level of metaphonological analysis (see e.g. Carlisle 1991), most will agree that there is a fundamental difference between analysing utterances into larger units like syllables and onset–rime units, and analysing utterances into smaller units like phonemic segments. There is a great deal of empirical evidence to support this divide, for the most part in the form of studies of the relationship between phonological awareness and literacy.

It is not at all surprising that phonological awareness is implicated in the acquisition of literacy. As noted in early studies by I. Liberman (Liberman 1973), and reiterated by numerous other researchers since, speech understanding does not require the listener to have a conscious concept of units smaller than a word, or perhaps a morpheme. Reading in an alphabetic orthography, on the other hand, requires a much more explicit understanding of speech sounds. This is not the same as knowing what sounds each letter is supposed to correspond to: if this were the case then "The child who is told that /bə/, /æ/, /tə/ are the sounds of B,A,T, respectively, would read 'bat' as the nonsense word 'buhatuh' " (Morais 1991, refering to Liberman 1973). Instead what is needed is for the child to become consciously aware of the phonological structure of speech, and to relate this structure to an orthographic representation of speech.

The seminal study of the relationship between specific levels of phonological awareness and literacy skills was carried out by Liberman et al. (1974). I. Liberman and colleagues found that while young, pre–reading children were able to successfully count the number of syllables in a word (indicated by their ability to tap out the syllables in the word), they were unable to count the number of phonemes. Older, reading children, on the other hand, were able to count both the number of syllables *and* the number of phonemes. Thus, the shift from syllable awareness to phoneme awareness appears to imply literacy in some way. The question that then arose from these findings, and which continues to be quite contentious, is what exactly the nature and direction of causality might be in the relationship between shifts in phonological awareness and literacy.

Clearly, if pre–reading children are able to count the number of syllables in a word, this is not a skill which is likely to be caused by literacy acquisition, although it might be one which would itself aid in learning to read. A number of researchers have proposed such a relationship: i.e. between the levels of phonological awareness which clearly precede literacy acquisition, and later literacy acquisition itself. Studies (both longitudinal and training) by Bryant and colleagues (Bradley & Bryant 1983, Bryant, MacLean, Bradley & Crossland 1990, Bryant 1998, Goswami & Bryant 1990, Kirtley, Bryant, MacLean & Bradley 1989) have shown that awareness of onset–rime units, which arises before reading, is particularly predictive of later reading ability. These researchers suggest that the importance of onset–rime awareness to reading becomes clear when one examines the complex phoneme–grapheme relationship in English, referred to in the introduction to Chapter 1:

> English is a capricious orthography in general, but it is much less predictable at the level of the single letter than of groups of letters. Thus a word like "light" cannot be easily read letter–by–letter, because the individual letters represent sounds which do not add up to the word "light." But it is quite possible that a child could come to read this word by learning that there is a group of written words which end in the letters "–ight," and which always end in the same rhyming sound. (Goswami & Bryant 1990, p. 27)[6]

However, while the spontaneous emergence of syllable, and perhaps onset–rime, awareness clearly precedes reading, it is not clear from Liberman et al.'s (1974) study what causes the emergence of the ability to analyse words at the level of phonemes. The two groups of children in the study differed *both* in literacy level (pre– and beginning–readers), and in age (pre–school and kindergarten). From the results of this study, therefore, it is impossible to conclude whether phonemic awareness is a skill which develops maturationally, or one which develops as a result of literacy acquisition.

Differences in phonemic awareness ability have also been found between children of the same age who are either good or poor readers: poor or dyslexic readers have much more difficulty with phonemic awareness tasks than do good readers (e.g. Bradley & Bryant 1983, Fox & Routh 1975, Treiman & Baron 1981). However, while these studies underline the relationship between phonemic awareness and literacy, they do not address the question of causality in the relationship: an inability to become phonemically aware could be either the source of poor readers' literacy problems or the result of them.

A clearer picture of the relationship between phonemic awareness and literacy emerges from studies of groups of subjects who have different levels of literacy *not* because of age differences or clinical disorders, but because of different educational experiences. The groups that meet these requirements that have been studied most extensively are illiterate vs. literate adults, and alphabetic

[6]For an argument *against* the importance of onset–rime awareness in reading acquisition, see also research by Snowling and colleagues, e.g. Muter, Hulme, Snowling & Taylor (1997), Hulme, Muter & Snowling (1998).

vs. non–alphabetic literates. The first pair of subject groups—illiterate and literate adults—has been studied in great depth by Morais, Bertelson and colleagues (Bertelson, de Gelder, Tfouni & Morais 1989, Morais, Cary, Alegria & Bertelson 1979, Morais, Bertelson, Cary & Alegria 1986). The studies of these researchers have shown that while literate adults (whether literate from childhood, or 'ex–illiterate'—i.e. only literate from adulthood) were able to perform phonemic awareness tasks at the same level as reading children, illiterate adults performed at a comparable level to pre–reading children (i.e. poorly)[7]. These studies would seem to suggest that alphabetic literacy acquisition, and the understanding of phoneme–to–grapheme correspondences that it entails, is the catalyst for the development of phonemic awareness, and that without this catalyst, phonemic awareness does not develop. A very specific example of this comes from a study by Morais (1991) of two illiterate poets. These two subjects were found to have highly developed awareness of all levels of phonological structure *except* phonemes. Both poets were able to correctly judge rhyme, distinguish rhyme (e.g. 'povas–movas') from assonance (e.g. 'chomba–zonta') (Morais 1991, p. 11), and find the odd–word–out of a list based on differences in onset. However, even when one of the subjects was explicitly taught to analyse CV and CVC syllables into phonemes, he was unable to segment new CVC syllables.

Morais & Kolinsky (1995) note that a number of objections have been raised against the results of the above studies, on the grounds that the illiterates' poor scores could be due to less well developed general cognitive abilities as a result of lack of schooling. Morais goes on to note that the *same* differences in phonemic awareness ability are seen between illiterate adults and ex–illiterate adults, both of whom score at the same level on tests of general cognitive ability. However, there are studies which do go some way further to addressing this specific problem. These studies examine subjects with different literacy backgrounds, but no difference in overall educational background: alphabetic vs. non–alphabetic literates. The best referenced of these studies is an examination of two groups of Chinese subjects, carried out by Read, Zhang, Nie & Ding (1986). The standard Chinese orthography (for both Mandarin and Cantonese) is a logography—i.e. each symbol is an ideogram. There is, however, an alphabetic script, called Hanyu Pinyin, which is also taught in China. The two groups in Read et al.'s (1986) study consisted of one group who could read only logograms, and a second group who could read both logograms and the alphabetic Pinyin. Consistent

[7]Note that these illiterate adults are adults who have never been exposed to literacy training, as opposed to adults who have failed to acquire literacy skills.

with the results of all of the above studies, the logogram–only readers performed very poorly on phonemic awareness tasks, while the logogram–plus–alphabet readers performed very well on the same tasks.

Bertelson & de Gelder (1991) report on a pilot study which replicated the above study with Chinese–Dutch bilingual subjects. All of the subjects in this later study could read Chinese logograms, but only half had learned to read in Dutch (none had learned the Chinese alphabetic Pinyin). The subjects were tested in Dutch (using both real and non–real words) on their ability to judge rhyme, and to segment and delete consonants. Again, both groups scored well on rhyme judgment, but the logogram–only readers scored much worse than the logogram–plus–alphabet readers for consonant segmentation and deletion.

The results of all of these studies provide evidence that, at the very least, the development of phonemic awareness skills is very closely related to alphabetic literacy acquisition.

2.3 Issues

Having provided what appears to be reasonably conclusive evidence for the way in which phonemic awareness develops, it should be pointed out that there are a large number of unresolved issues regarding the emergence of this skill. These issues, which relate to the development of phonemic awareness, its relationship with literacy, and the way in which it is tested, will be discussed in this section.

Absolute link between alphabetic literacy and phonemic awareness

The results of the above studies of illiterate and logogram–only literate adults have led Morais, Bertelson and colleagues to propose that the link between phonemic awareness development and alphabetic literacy is exclusive: that is, unless a child learns to read in an alphabetic orthography, they will not become phonemically aware (e.g. Morais & Kolinsky 1995). There are, however, some questions which have been raised about the validity of this claim. Mann (1986), for example, carried out a study which compared the phonemic awareness skills of Japanese school children to those of American first grade children. Mann believes the results of this study show evidence of phonemic awareness development without alphabetic literacy acquisition. The Japanese children in this study were first through sixth grade students, learning to read in both a kanji logography (based on the Chinese logography) and a kana syllabary. The Japanese syllabary transcribes speech in terms of moras (see Chapter 1, Section 2.3 for a definition of a mora). The older children in the study (those in the sixth grade) had also begun to learn to read using an alphabet called Romaji, which is based on the Roman alphabet. The Japanese first grade children were found to perform poorly on phoneme counting and deletion tasks compared to American first grade children. However, the Japanese fourth grade children were found to perform reasonably well compared to the American first graders, despite the fact that at this point they had not begun to learn to read in an alphabet. Finally, both the Japanese fourth grade children and the American first grade children were outperformed by the Japanese sixth grade children, who (like the American children) had begun to learn to read in an alphabet.

The performance of the first grade Japanese children in comparison to their American counterparts is what might be expected from the results of the studies discussed to this point. On the other hand, the performance of the Japanese fourth grade children suggests that access to an alphabetic orthography may not be absolutely necessary for phonemic awareness development. However, Bertelson & de Gelder (1991) propose that Mann's study of Japanese readers is not a true test of the constraints for the development of phonemic awareness. These authors suggest that there may be some aspect of the kana syllabary which gives Japanese children access to phonemic structure:

> The kanas are actually not pure syllabaries. First, some characters represent single segments: the five vowels /a/, /ε/, /i/, /o/ and /u/, and the consonant /n/. Also, the fact that different kana stand, for instance, for /pa/, /ta/, /ka/, /ma/, and /a/ may indirectly draw attention to the consonants. The probability of such discovery is increased, because kanas are usually presented to the pupils in matrix arrangement with columns corresponding to the initial consonant of the represented mora and the rows corresponding to the vowel. (Bertelson & de Gelder 1991, p. 402)

Mann (1986, 1991), however, notes that although some of the subjects in the study did report relying on a strategy which made use of the kana syllabary matrix chart, these subjects did not perform any differently than the rest of the subjects. Mann instead suggests that the difference seen between the Japanese and American pupils, in particular the better performance of the Japanese sixth grade pupils compared to the American first grade pupils, indicates that "the age of the child has an impact on the degree of phoneme awareness, and also on the ability to profit from instruction in the alphabetic code" (Mann 1991, p. 58). Mann then goes on to propose a sort of critical period for phonemic awareness development, which she suggests could "reconcile the presence of awareness among children who lack knowledge of an alphabet and its absence among illiterate adults" (Mann 1991, p. 62).

Mann, however, does not suggest that the development of phonemic awareness is wholly dependent on age or maturation—just that this may itself place constraints on the development of phonemic awareness. Instead she proposes that phonemic awareness should develop in the context, not just of alphabetic literacy training, but of any "experience in manipulating the internal structure of words" (Mann 1991, p. 62). An example of such an experience, suggests Mann, can be found in the acquisition of 'secret' or play languages spoken by many children (and adults) in literate, illiterate, and non–alphabetic cultures. The premise of most of these languages is the manipulation (i.e. reversal, addition, deletion, etc.) of the phonological structure of words, at various different levels, including the phonemic level. Examples of these languages are 'pig latin', an English–based play language which manipulates words at the level of onset–rime—thus 'please' becomes /lizpɛɪ/—and the 'la–mi' language of Cantonese (e.g. Mann 1991) which manipulates words at the level of the phoneme—'ha:ng' becomes /la:ng hɪng/ (the initial consonant, the vowel, and the final consonant separate, the initial consonant and the vowel reverse, /l/ is inserted before the vowel, /ɪ/ is inserted after the consonant, and the final consonant is added back to the ends of both new syllables). Mann (1991) proposes that experience with language play such as found in these games can serve to encourage phonemic awareness development in both alphabetically literate and illiterate children without requiring specific alphabetic reading instruction.

It should also be noted that there is some evidence from clinical studies that phonemic awareness can be directly trained in pre–reading children. Studies have shown that the type of metaphonological training that is sometimes used as part of the remediation process for phonologically disordered children, encourages the development of phonemic awareness before the onset of literacy training (e.g. Howell, Hill, Dean & Waters 1993, Innes 1995).

Mann (1991) does however go on to note that her suggestion—i.e. that *any* experience with accessing the 'internal structure' of words can lead to phonemic awareness—does not explain the original invention of secret languages in non–alphabetic cultures. Mann goes on to note that "There is also the problem that certain children demonstrate surprising levels of phoneme awareness that their teachers and parents are at a loss to explain" (Mann 1991, p. 62). Other researchers explain these phenomena by claiming that some level of phonemic awareness *may* develop without any particular training, either formal alphabetic literacy training or otherwise. Lundberg (1991), for example, notes that in a certain number of studies (e.g. Lundberg, Olofsson & Wall 1980, Lundberg, Frost & Petersen 1988) a very small number of pre–reading children (ranging from 8 out of 387 pre–readers, to 9 out of 51 pre–readers) were able to correctly perform half or more of the phonemic awareness tasks. Lundberg believes that "the fact that such children exist apparently indicates that it is possible, at least in principle, to develop phonemic awareness without the support of formal reading instruction at school" (Lundberg 1991, p. 50).

It does seem, from the results of these studies, that phonemic awareness may in fact develop in response to various sorts of metaphonemic training, and, in rare cases, may even develop spontaneously to some extent. Alphabetic literacy training may not after all be the sole source of phonemic awareness ability. It should be noted at this point, however, that the opposite does not appear to be true—that is, while it has been shown that phonemic awareness may develop without alphabetic literacy, it has not been shown that it is possible to become literate in an alphabetic language without, as a result, developing phonemic awareness. Therefore, while there is some argument as to the exclusive nature of alphabetic literacy's causal relationship with phonemic awareness, it can at least be said that the relationship is absolute.

How aware is aware?

Lundberg's findings, that in some rare cases children may develop phonemic awareness spontaneously without reading instruction, raises another issue: what constitutes awareness? How successful does a subject need to be at a given phonemic awareness task to be considered phonemically aware? And, does the same subject then need to show equally high levels of awareness in *other* phonemic awareness tasks in order to be considered phonemically aware? Lundberg (1991) clearly believes that a score of 50% on any test constitutes phonemic awareness, and does not appear to believe that it is necessary for subjects to achieve this score on all tests of phonemic awareness: while 9 of 51 pre–readers that she reports on performed to this level on a phoneme segmentation task, only

3 of the same 51 performed to the same level on a phoneme synthesis task. All of these subjects were considered by Lundberg to have some phonemic awareness ability. Morais & Kolinsky (1995), on the other hand, suggest that "A non–negligible score in one particular "phonemic awareness" task does not necessarily indicate the presence of phonemic awareness. Rather it may reflect the fact that the subjects have found some strategy that is appropriate to deal with the particular task they have to perform" (p. 318). As evidence of this, the authors describe the performance of an illiterate adult on various phonological awareness tasks. This subject achieved a 50% success level on a consonant deletion task, performed at chance on a consonant oddity task (i.e. 'miss, moon, *p*ig'), and was unable to correctly perform *any* of a syllable deletion task. Based on the first two results, Lundberg might state that this subject was phonemically aware, however Morais & Kolinsky (1995) claim that the results of the syllable deletion task, which is usually quite easy for pre–and illiterate subjects, show that this subject is not actually precocious at phonemic awareness tasks, rather that she has devised a strategy which allows her to perform one specific type of task. This leaves the question open, then, as to what level of success at phonemic awareness tasks is required for an individual to be considered phonemically aware. This question will be addressed further in Chapter 6.

What constitutes a phonemic awareness test?

In addition to the issue regarding the level of success required at a phonemic awareness task, there is also the issue of the actual design of the phonemic awareness tests themselves. The question is, what type of task taps phonemic awareness and can be said not to tap anything else?

The first problem in this area is the potential that exists for confounds between units of the same *actual* size, but different *linguistic* size or level. In the hierarchical model of phonological structure described in Section 2.1, a word is made up of a certain number of syllables, each of which can be broken up into onset–rimes, and finally all of the above can be broken up into phonemes. However, in many instances, elements of a word may be valid units at more than one linguistic level. As noted by Mann (1991), "deleting the initial consonant from a word like *cat* can be regarded either as "phoneme" or "onset" deletion; the first phoneme of a word like *open* is simultaneously a phoneme, a rime, and a syllable" (p. 56). This is illustrated in Figure 2.9.

Keeping in mind the possible differences in reading and other experience that have been found to be necessary for different levels of phonological awareness, it becomes clear that such confounds pose potential problems for phonemic awareness testing.

A second possible problem area in phonemic awareness (and in fact in all metalinguistic awareness) testing has to do with the level of *explicit* vs. *implicit* awareness that is being tested. Morais (1991) gives a good example of the difference between these two levels in his report on the study of two illiterate poets. As noted above, both were able to produce rhymes and alliterative pairs, and to make correct judgments on pairs of words based on rhyme vs. assonance. However, they were unable, even with training, to delete the onset of a word, and produce just the rime. Morais suggests that the two poets are able to produce and judge rhymes

> not because they are able to make the onset–rime distinction, but because they are more sensitive to phonological similarities that arise from common onset or from common rime. Alliteration and rhyming abilities cannot be equated with the ability to analyse syllables in terms of onset and rime. (Morais 1991, pp. 11–12)

The poets seem to have an *implicit* awareness, or sensitivity as Morais calls it, to the phonological structure of words which allows them to play with language in the way that is necessary for them to create poetry. They do not, however, seem to have the *explicit* awareness of the units of speech that they are playing with which would allow them to manipulate these units individually. If this divide between implicit and explicit awareness does exist, then there is a fundamental

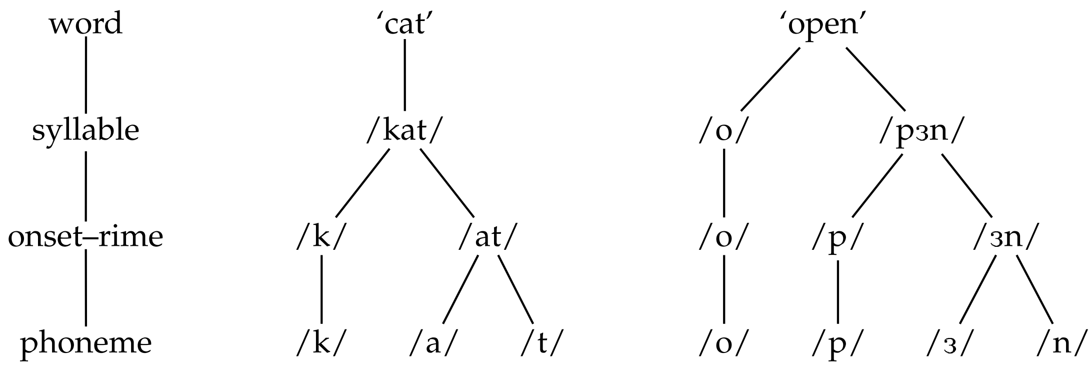

Figure 2.9: **Phonological structure of 'cat' and 'open'.**

difference between phonemic (or any metalinguistic) awareness tasks which require some sort of passive judgment to be made (e.g. judging which word in a list starts with a different sound from the others), and those tasks which require a deliberate manipulation of elements of speech (e.g. deletion of a phoneme from a word).

Finally, there are a number of non–metalinguistic processes which have been found to be implicated in phonemic awareness. Treiman & Zukowski (1991), as well as noting that phonological awareness tasks differ depending on the linguistic level that is being tapped, note that tasks may also differ depending on the level of cognitive development necessary to complete the task: "children have more difficulty manipulating the phonemes in a word, as in saying 'sun' backwards, than in recognising that 'sun' contains 's', 'u', and 'n'." (Treiman & Zukowski 1991, p. 67). In other words, the more cognitive processes that are called into play (e.g. in the first 'sun' example these include an ability to understand what 'backwards' means, and the ability to apply this concept to a string of elements), the more difficult the task will be. In a study which directly tested this relationship, McBride-Chang (1995*b*) found strong correlations between the general cognitive ability of subjects and their success on a phonemic awareness task.

McBride-Chang (1995*b*) also found that subjects' ability on tasks which tested their memory correlated with their phonemic awareness. This is not entirely surprising: as McBride–Chang notes, in order for a subject to be able to complete a phonemic awareness task, they must be able to remember the stimulus, the nature of the operation to be performed, and (potentially) all of the individual phonemes in the stimulus. Studies by other researchers have also found that dyslexics and poor readers, who as noted above generally have poor phonemic awareness, often also have poor short–term memory skills, and in particular perform very differently to good readers when asked to recall lists of rhyming words (Conrad 1971, Liberman, Shankweiler, Liberman, Fowler & Fischer 1978).

Most importantly for this current study, the third process that McBride–Chang and colleagues have found to be implicated in the process of phonemic awareness is speech perception (McBride-Chang 1995*b*, McBride-Chang 1996, McBride-Chang, Chang & Wagner 1997, Manis, McBride-Chang, Seidenberg, Keating, Doi, Munson & Petersen 1997). Again this finding is not entirely surprising—at the very least the successful completion of a phonemic awareness task requires successful perception of the stimuli. However, as will become clear in the next section, where the potential correlations between perception and awareness will be discussed in detail, the relationship is much more complicated than this.

It appears, therefore, that the answer to the question posed at the beginning of this section—i.e. what type of task taps phonemic awareness—has three parts. First, because of the possibility of confounds between phonemes and onsets, we can only say for certain that a task is tapping *phonemic* awareness if there is no possibility that the unit being manipulated is also an onset. Second, it has been shown that tasks that ask for an implicit judgment and those that ask for an explicit manipulation of phonemes may require different types of phoneme awareness. Finally, it appears that cognitive, memory and perceptual demands are also implicated in phonemic awareness tasks.

3 The relationship between acoustic cue weighting and phonemic awareness

As noted in the introduction to this chapter, the central goal of this thesis is to investigate the *nature* of the relationship between perceptual weighting of acoustic cues, and phonemic awareness, with particular emphasis on the possible causal direction of the relationship. The first step in this process was to establish what is known about the way that each process develops individually—this was the goal of the last two sections of this chapter. The next step in the process is to determine what is known about the way in which these two processes correlate.

Clearly the *primary* evidence for the existence of a relationship between the two processes comes from Nittrouer's (1996*b*) study—in fact, this appears to be the only study which has explicitly looked at the relationship between these two very specific aspects of speech perception and metaphonological awareness. However, additional support for the existence of a relationship can be found in studies from two slightly different, but related areas of research. The first of these is the study of the relationship between speech perception and literacy skills. Keeping in mind the reasonably well documented relationship between reading acquisition and the development of phonemic awareness (as discussed in the previous section), any studies which show a relationship between perception and literacy should also indicate a potential relationship between perception and phonemic awareness.

The second area of research that provides additional support for a possible relationship between cue weighting and phonemic awareness encompasses those studies of speech perception which have also looked at some aspect of phonological awareness. Again, because phonemic awareness is *part* of the spread of phonological awareness skills, studies that show a relationship between speech perception and phonological awareness also lend support to Nittrouer's (1996*b*) findings.

Some of the studies in both areas do investigate acoustic cue weighting, while others specifically investigate phonemic awareness. However, very few look at the two together, and none (besides Nittrouer's (1996*b*) study) actually look for a correlation between the two. Additionally, most of these studies suffer from the same problem as Nittrouer's (1996*b*) study—specifically that they are predominantly cross–sectional rather than longitudinal. However, having said this, all of the studies provide some degree of support for the actual existence of a relationship between perception and awareness, and many also allow for a level of speculation regarding the causal direction of the relationship: this should aid in the formulation of possible hypotheses for this thesis.

3.1 The relationship between speech perception and literacy

For the most part, studies which have examined the possibility of a relationship between speech perception and literacy have made use of contrasting groups of good and poor, or dyslexic, readers. The main reason for this seems to be that the goal of these studies was not to simply establish the existence of a relationship between perception and literacy. Rather the goal was to determine whether perceptual problems could be the source of poor readers' literacy deficits. Unfortunately this goal often gives the conclusions drawn from the studies a rather one–sided slant: perceptual development is almost always seen as a building block for later reading ability, rather than reading ability as a possible cause of changes in perceptual behaviour, despite the fact that the predominantly cross–sectional studies preclude conclusions being drawn either way. However, the results of the studies themselves do suggest that a relationship of some sort exists between perception and literacy.

A variety of methodologies have been used in these studies to test perception. The first is potentially the most 'life–like'—that is, the most like the perceptual situations encountered by the average listener—specifically, word recognition or repetition. In this type of test, the subject listens to a list, or lists, of words and repeats what they think they've heard back to the examiner. The drawback of this type of test is that it is more difficult to control for variation in all aspects of the stimuli, and therefore difficult to determine the exact aspect of the stimuli which might be problematic for listeners. However, the tests are often made more difficult (and thus more likely to show up more subtle differences between listeners) by masking the words with white noise, or by using non–or pseudo–words (which prevents the listener from using context to help them decipher misperceived sounds).

This type of test was used in a study by Brady, Shankweiler & Mann (1983) which attempted to determine the possible source of dyslexics' poor short term memory difficulties. As noted in Chapter 1, Section 2.7, it has been shown that listeners are affected by rhyme when remembering lists of words—that is, they are more likely to make recall errors for a list of rhyming words than for a list of non–rhyming words (Conrad 1971). It has also been shown that there is a difference between good and poor readers in the amount they are affected by rhyme: poor readers are much less affected than good readers by whether words in the list rhyme or not (Liberman et al. 1978, Shankweiler, Liberman, Mark, Fowler & Fischer 1979). Brady et al. (1983) suggest that this difference is due to poor readers' "failure to fully exploit phonetic coding" (p. 346, see also Liberman et al. 1978)—in other words, poor readers are less sensitive to the phonological similarities between rhyming words, and thus are less likely to confuse them. Brady et al.'s (1983) study was designed to test whether this apparent problem can be traced to perceptual difficulties. Good and poor readers were played words, both with and without noise masking, and were asked to repeat them. In the unmasked condition there was no difference in perceptual performance between the good and the poor readers. However, in the masked condition the poor readers had significantly more difficulty correctly perceiving the words than did the good readers. Brady et al. (1983) conclude that poor readers have less effective perceptual skills than good readers, but that the difference is so slight that it can only be seen when listening conditions are made more demanding.

A similar study was undertaken by Snowling, Goulandris, Bowlby & Howell (1986). These authors expanded on Brady et al.'s (1983) study by including non–words, as well as high and low frequency real words, and by assessing perceptual ability not just in good and poor readers of the same chronological age, but also in normally developing children of the same reading age as the dyslexics (i.e.

younger children who are reading at the same level as the dyslexics). This study again made use of noise masking. The results showed that while dyslexic readers were equally good as the age–matched good readers on high–frequency real word repetition, they were much worse on low–frequency and non–word repetition. When compared with the reading–age–matched children, the dyslexics were much worse at non–word repetition, but were equal in ability for both high and low frequency real word repetition. Interestingly, contrary to the findings of Brady et al. (1983), the dyslexics were not affected by the noise masking any differently from the normally developing readers. Snowling et al. (1986) conclude from this that not all aspects of speech perception are impaired in poor readers: the authors suggest that, because the noise masking affected all readers similarly, dyslexic readers' perceptual difficulties are not related to perception "at input" (p. 504). Instead, the authors claim that the fact that the dyslexics had most difficulty with the non–word repetition indicates that the aspect of perception which is impaired is some post–input aspect which deals with immediate analysis of the phonological structure of new words[8].

A number of studies have used a variation of this type of test, in which listeners are asked to discriminate between pairs of words, instead of simply recognising them (e.g. Adlard & Hazan 1998, Masterson, Hazan & Wijayatilake 1995, Mody et al. 1997). The pairs of words may differ in a number of different ways: they may be minimal pairs (differing in only one feature, e.g. 'date–gate'), consonants appearing in one word may be omitted from the other (e.g. 'pay–play'), and consonants may be changed across word pairs (e.g. 'spill–still') (all examples from Adlard & Hazan 1998). All of these studies report certain discrimination deficits for dyslexics compared to normal readers, although in all studies the deficit was found to be restricted in some way. Masterson et al. (1995) found the two adults dyslexics that they tested to have perceptual discrimination difficulties only for specific contrasts (predominantly fricatives). Adlard & Hazan (1998), on the other hand, found that only 30% of the dyslexic children that they tested showed perceptual difficulties, and that while the error rate for this subgroup was significant, it was also quite small. The study by Mody et al. (1997)

[8]This multi–layer view of speech perception put forward by Snowling is one of a number of different hypotheses which account for the mechanisms of speech perception. This particular view is based on a model of word recognition in which there are two routes to verbal word repetition—one which requires phonological analysis, and one which only requires access to the phonological representations stored in the mental lexicon (see Snowling et al. 1986, p. 491). Under this hypothesis, new words cannot be recognised by reference to a lexicon, and must therefore be analysed phonologically before repetition.

tested poor readers' ability to discriminate pairs of the following (computer–generated) syllables: /ba/, /da/, /sa/, /ʃa/. The poor readers were found to be worse than good readers on discriminating between /ba/ and /da/, but not on discriminating /ba/ from /sa/ or /da/ from /ʃa/. The authors suggest that poor or dyslexic readers may have specific perceptual problems with phonological contrasts which are phonetically similar (i.e. which differ in only one feature).

As noted above, while word repetition tasks, and to a lesser extent word or syllable discrimination tasks, are fairly natural tests of general perceptual ability, they are not very flexible tests of more specific aspects of the perceptual process. For instance, although word repetition test materials can be controlled to a reasonable extent in terms of the number and placement of different types of segments—i.e fricatives vs. stops vs. nasals etc.—it is difficult to use these tests to assess listeners' ability to cope with the minute variations in different aspects of the speech stream that they would have to cope with when listening to (for example) multiple different speakers—e.g. variations in fricative frequency, or of VOT. Word discrimination tasks can test a listener's ability to cope with more global variations (i.e. the change from /s/ to /ʃ/) but a more sensitive test of the effect of smaller variations in stimuli on the perceptual system is a categorical perception test. In fact, because the perceptual phenomenon of categorical perception is well tested and reasonably well understood, this type of testing has been used as a measure of perception in a number of studies of perception and literacy.

One well known study of categorical perception and literacy was carried out by Werker & Tees (1987). These authors believed that there might be a link between the categorical perception phenomenon and the access to phonological structure needed for the phoneme–to–grapheme conversion process: "This categorical perception capability imposes an initial phonetic categorisation on spoken language and is thought to provide the basis from which phonological categories are constructed" (Werker & Tees 1987, p. 49). Presumably any developmental problems in perception would affect the organisation of these categories, and thus impinge on the development of phoneme awareness. Werker & Tees (1987) examined average–reading and reading–disabled children's perception of a /ba/–/da/ contrast, which varied along a continuum in terms of the onset frequency and configuration of F2. The subjects were presented with these stimuli in four different speech perception tasks. The first of these was a 2–item, forced choice labeling task: the subjects were told that they would hear either /ba/ or

/da/ and were asked to say which they had heard. The second and third tasks were both discrimination tests: an AX task, in which the subjects were told they would hear pairs of stimuli and were asked to label the pairs as same or different, and and ABX task in which the subjects heard three stimuli and were asked to state whether the third was the same as the first or the second. The final task was a change/no change task, in which subjects were played repetitions of one or two stimuli, which, at irregular intervals would either change or remain the same. The results of the study found that the disabled readers' perception was less categorical than the average readers' for all but the ABX task (for which both groups performed poorly, suggesting that it is a fairly demanding task). These differences were small, but both significant and consistent (Werker & Tees 1987).

Other studies have also found differences on categorical perception tests between good and poor readers, although much as the results of Werker & Tees, the differences are often small, or apparent for only one subset of the subjects. Godfrey, Syrdal-Lasky & Knox (1981), for example, tested labeling and discrimination of /ba/–/da/, which varied in terms of F2 and F3 onset transitions, and of /da/–/ɡa/, which varied in terms of F3 onset transitions. These authors found that poor readers labelled both contrasts more variably and less categorically than good readers, and were significantly less accurate than good readers at discriminating between stimuli which were taken from either side of the phoneme category boundary. A study by Reed (1989) found that reading–disabled children were less able than good readers to discriminate consonants, although they were no different from the good readers for the discrimination of vowels.

A more recent study by Joanisse, Manis, Keating & Seidenberg (1998) found that only a subset of the dyslexics that they tested showed categorical perception deficits. Dyslexics in this study were labeled (pre–testing) as 'phonological' dyslexics, 'language impaired' dyslexics, or 'delay–type' dyslexics. The testing involved labeling of two sets of contrasts: a 'dug–tug' contrast (VOT continuum) and a 'spy–sky' contrast (F2 onset continuum). Only the 'language impaired' dyslexics showed perception that was significantly less categorical than the normal control group. The authors suggest that this variability in perceptual behaviour among dyslexics could explain the relatively small differences seen between dyslexics and good readers in some of the previously discussed categorical perception studies.

Interestingly, in relation to the aims of the current study, there are a number of studies (two of which have been discussed above) which have actually looked at the use, or weighting, of particular acoustic cues by good and poor readers. As well as looking at word discrimination, Mody et al. (1997) also attempted to replicate Nittrouer's (1992) study of the weighting of cues to a /sɛɪ/–/stɛɪ/ ('say–stay') contrast. Unfortunately, although the authors state that the stimuli which they use are identical to those used by Nittrouer (1992), there is an important difference between the two studies in the design of the stimuli. Nittrouer's study used two continua, both varying along a continuum of silence duration, and each with a different F1 onset configuration. Mody et al., on the other hand, made use of just one continuum, which had a fixed silence duration of 20 ms. (approximately half–way between the two end point gap durations used by Nittrouer), and an F2 onset continuum which varied from a configuration appropriate for /sɛɪ/ (611 Hz) to a configuration appropriate for /stɛɪ/ (211 Hz). The results of Mody et al.'s (1997) study showed a small and non–significant difference in perceptual behaviour between the good and poor readers—the poor readers had a slightly shallower response curve slope that the good readers. The authors compare this result to those of Nittrouer (1992) and conclude that the poor readers did not weight transitional cues more heavily that good readers. This is contrary to what might be expected from Nittrouer's studies if poor readers are delayed in perceptual development. However, while this is one potential explanation of the results, it may not be the only explanation. Because this study made use of only one continuum, there is no opportunity to measure the *shift* in phoneme boundary caused by the addition of a second cue to the contrast, as was done in all of Nittrouer's studies of acoustic cue weighting. Additionally, the fact that the two cues were swapped in the stimulus design relative to those used by Nittrouer (1992) (Nittrouer varied the silence duration on a continuum, while Mody et al. (1997) varied F2 onset configuration along the continuum) means that it is quite difficult to directly compare this study with those of Nittrouer and colleagues.

The second study which undertook to determine the extent to which good and poor readers make use of certain cues is the study by Adlard & Hazan (1998) (discussed above). These authors presented good and poor reading children with speech stimuli in what they call 'combined–cue' and 'single–cue' conditions. The stimuli for the 'combined–cue' condition were designed on a continuum which simultaneously varied a combination of cues (usually two) to a certain contrast: for example, for the 'date–gate' contrast that they used, both the burst frequency *and* the F2 onset transitions were varied from values which cued /d/ to values which cued /g/. For the 'single–cue' condition, stimuli were designed so that only *one* of the cues from the combined–cue condition was varied along the continuum: for the 'date–gate' contrast, therefore, both a continuum varying in burst noises, and a continuum varying in F2 onset transitions were created. The authors note that by comparing listeners' ability to label contrasts which are signalled by different sets of acoustic cues, it should be possible to determine which cue is most important to any listener. The results of the study show little difference in 'categorical–ness' of perception between good and poor readers for the 'combined–cue' conditions. Additionally, perception was generally less categorical for *both* groups in all of the single–cue conditions, although (as found by Nittrouer for fricative–vowel stimuli) the children's perception was more categorical for the /d/–/g/ single–cue contrast when this cue was F2 transitional information rather than burst frequency information. A significant difference in perceptual behaviour between normal and poor readers was, however, found for one of the four contrasts tested: dyslexics were found to respond less categorically than non–dyslexics to a /s/–/z/ contrast cued by changes in fricative duration only.

Finally, Nittrouer herself has examined acoustic cue weighting in good and poor readers. Nittrouer (1999) tested good and poor readers on acoustic cue weighting for four contrasts. The readers were divided on the basis of the reading sub–test of the Wide Range Achievement Test–Revised (WRAT-R; Jastack & Wilkinson 1984), which assigns subjects into 'normal' or 'poor' phonological groups. The four contrasts tested were: /da/–/ta/ (the two cues were burst intensity and vowel onset transitions), /sɛɪ/–/stɛɪ/ (the two cues were silence duration and vowel onset transitions, see Nittrouer 1992), /sa/–/ʃa/ and /su/–/ʃu/ (the two cues were fricative noise frequency and vowel onset transitions, see Nittrouer 1992, Nittrouer 1996*b*, Nittrouer & Miller 1997*b*). Nittrouer (1999) found that poor readers weighted cues differently from good readers in both /s/–/ʃ/ contexts. Specifically, it was found that poor readers weight transitional cues more heavily than fricative cues in labeling these contrasts.

The final studies to be discussed in this section are two investigations by de Gelder & Vroomen (1992, 1998) on audio–visual speech perception. The visual speech perception (or lipreading) aspect of the studies will be discussed in more detail in the following section. In the audio speech perception part of these studies, de Gelder & Vroomen compared categorical perception of a /ba/–/da/ continuum (F2 and F3 onset transitions) by different groups of subjects. However, for de Gelder & Vroomen (1992) unlike the previous studies discussed in this section, the groups of subjects were not good and poor readers, but alphabetic and non–alphabetic (logographic) readers. de Gelder & Vroomen tested Dutch alphabetic readers, Chinese–Dutch bilingual logographic readers, and Chinese–Dutch bilingual and bigraphemic readers (i.e. readers of both an alphabet and logograms). The results showed that logographic readers had significantly less categorical response slopes, and placed their phoneme boundaries in a significantly different place, from the alphabet–only readers. Interestingly, the bigraphemic readers had categorical perception slopes which were intermediate between the alphabetic and logographic readers, leading the authors to suggest that "the possibility that orthographic skills exercise an influence on speech categorisation must be taken seriously" (de Gelder & Vroomen 1992, p. 423). Making use of the same methodology, de Gelder & Vroomen (1998) replicated the results of the 1992 study with good and poor readers. This suggests that the results of the 1992 study were not simply due to the different language experiences of the subjects, but more to their different *meta*language experiences.

It is reasonably clear from all of the above studies that at least some poor readers show some speech perception deficits compared to good readers, and that even non–alphabetic readers may perceive speech differently from alphabetic readers. For the most part, however, even those studies which show a clear deficit in speech perception for poor readers have shown this deficit to be restricted in some way—restricted to a specific aspect of perception (e.g. Snowling et al. 1986), restricted to a specific type of phonemic contrast (e.g. Masterson et al. 1995), restricted to only a portion of poor or dyslexic readers (e.g. Adlard & Hazan 1998, Joanisse et al. 1998), or simply restricted to a very small difference between poor and good readers (e.g. Brady et al. 1983, Werker & Tees 1987). However, for the purposes of this study, the fact that poor readers and logographic readers, who have both been shown to have poor phonemic awareness skills, also possibly have different speech perception strategies to good/alphabetic readers, offers at least some support for Nittrouer's (1996*b*) finding that phonemic awareness and perceptual cue weighting are related.

3.2 Speech vs. non–speech perception

It should be noted at this point that a large number of the studies above have tested poor readers on their perception of *non*–speech as well as speech sounds. The main reason for the inclusion of these tests in the studies is that there is some debate as to "the nature and origin of the perceptual deficit" (Mody et al. 1997, p. 200) in poor readers. There are generally considered to be two hypotheses: the first is that any speech perception problems in poor readers are just that—deficits in perception of *speech*. This account is generally put forward by researchers who hold the view that 'speech is special' (see e.g. Liberman 1996), that is, that speech perception is a function of a system dedicated to speech. The second hypothesis is that poor readers' speech perception difficulties are due to problems in their general auditory capacity—this view is held by researchers who believe that speech perception is a function of a general perceptual system—i.e. a system which operates for perception of all sounds (see Section 1.2 above for a brief discussion of auditory processing). No direct contribution to this debate will be made in this thesis. However, the argument itself could be important to our understanding of the relationship between speech perception and phonemic awareness.

Specifically, the theory that perception of speech is performed by a general auditory system seems to place restrictions on the direction of causality between perception and awareness. It is plausible that some specific aspect of an auditory system could be harnessed by metalinguistic development to enable a child to become aware of phonemes. It seems less plausible, however, that the development of phonemic awareness would have an impact on the functioning of an auditory system that serves for the perception of all, not just speech, sounds. The theory that perception of speech is performed by a dedicated system, on the other hand, seems to allow for the causal relationship between perception and awareness to go either way. It is equally plausible under this second theory that perception could have an impact on phonemic awareness, or that the development of phonemic awareness could affect a perception system dedicated to speech.

A great deal of research into the possible connection between general auditory deficits and literacy has been carried out by Tallal and colleagues (Tallal & Peircy 1973, Tallal & Peircy 1975, Tallal 1980, Tallal & Stark 1981). These authors state that both language impaired and reading impaired children have a perceptual deficit in 'temporal processing.' This deficit is claimed to result in difficulty judging the order in which stimuli have been presented, and in perceiving either brief or rapid acoustic events, such as stop bursts (which are brief, e.g. Tallal 1980), or formant onset transitions (which change rapidly in frequency, e.g. Tallal & Peircy 1975), or stimuli (speech and non–speech) which are presented in rapid succession (e.g. Tallal & Peircy 1973). The primary evidence that reading impaired children suffer from 'temporal auditory processing' deficits comes from

Tallal (1980). In this study, 20 reading impaired children were tested on their ability to perceive rapidly presented tones and stop consonants, and to perform a *temporal order judgment* task on pairs of tones and consonants (a temporal order judgment or TOJ task involves listening to a number of stimuli, and identifying the order in which they were played). Eleven of the subjects performed at the same level as normally developing children, while 9 had difficulty with the tasks, leading Tallal to conclude that reading impairment may coincide with difficulties in perceiving stimuli which are characterised by 'temporal' cues (however, for an examination of Tallal's use of the the term 'temporal' see Mody et al. 1997). Additionally, Tallal found a correlation between temporal order judgment of tones, and non–word reading. This is supported by the results of Reed's (1989) study (discussed above), which found that reading disabled children had more difficulty than good readers in judging the temporal order of stop consonants. Reed (1989) did, however, find that poor readers were no different than good readers for TOJ of vowels rather than consonants.

The evidence for the hypothesis that perceptual differences between poor readers and good readers are speech specific comes from a number of different sources. The study by Mody et al. (1997) discussed above, directly addresses Tallal's claim that speech perception is a capacity of a general auditory system. This study tested good and poor readers on their ability to discriminate between the syllables /ba/, /da/, /sa/, and /ʃa/. In addition, Mody et al. tested the same listeners on their ability to perform temporal order judgements on the same syllables, as well as their ability on discrimination and temporal order judgment tasks involving non–speech stimuli (these stimuli were sine wave analogs of /ba/ and /da/). The results found that while the poor readers had more difficulty than the good readers with certain speech contrasts (see above), there was no difference between good and poor readers in their perception of the non–speech sounds—both groups had more difficulty with this than with speech perception. The authors conclude that "Deficits in speech perception among reading–impaired children are domain specific and phonological rather than general and auditory in origin" (Mody et al. 1997, p. 227).

A second test of Tallal's hypothesis can be found in another study discussed above: Nittrouer's (1999) study. As noted above, Nittrouer tested good and poor readers on their acoustic cue weighting for a number of contrasts. Not mentioned above is the fact that this study also included tests of temporal order judgment of non–speech stimuli (sinusoids at 800 Hz and 1200 Hz). Nittrouer (1999) found that, while poor readers weighted cues differently from good readers in certain contexts, there was no significant difference in temporal order judgment of non–speech between good and poor readers. It should also be noted that in those cases where poor readers differed from good readers in the acoustic cues that they used for perception, it was the *transitional* cues that they weighted more heavily—which is specifically the type of cue that Tallal suggests should be problematic for poor readers. Nittrouer (1999) concludes that poor readers do not suffer from deficits in general auditory perception, but rather have problems with speech perception, and more specifically with the aspect of speech perception which allows them to shift their weighting of acoustic cues.

Other studies, while not necessarily directly addressing the auditory deficit hypothesis, have also found evidence that speech perception is different from non–speech perception. Brady et al. (1983), for instance, tested poor readers' perception of 'environmental' sounds—e.g. phone ringing, car starting—which they subjected to masking in the same way as they had the speech sounds. Although the poor readers were found to perform much worse than the good readers at word repetition in noise, their levels of success were the same as the good readers for identification of environmental sounds, whether in or out of noise. Adlard & Hazan (1998), also found that poor readers, including a subgroup which had been shown to have poor *speech* perception, did not perform any differently than the good readers for non–speech discrimination tests.

Finally, the studies of de Gelder & Vroomen (1992, 1998), which paired tests of audio speech perception with tests of visual speech perception (lipreading), also seem to indicate that speech perception is controlled by a dedicated system. de Gelder & Vroomen (1992) note that

> Visual speech identification performance offers a complementary source of information on speech sound categorisation as vision and audition represent two autonomous but very closely linked input modalities for speech. Adults with normal hearing combine the auditory and visual speech information in normal circumstances, as well as under impoverished conditions (e.g. Massaro, 1987). (p. 415)

In both of these studies, listeners were asked to label a synthetic /ba/–/da/ continuum which was presented simultaneously with a video of a speaker saying /ba/ and /da/. The listeners were also presented with the stimuli in audio–only and video–only conditions. de Gelder & Vroomen (1998) suggest that if the differences in perceptual behaviour between good and poor readers are speech specific, this difference should be apparent in speech reading as well; however, if the perceptual difference is due to a more general auditory deficit, then there should be no reason for poor readers to show a difference in speech reading. As had been found in previous studies of audio–visual speech perception (see McGurk & MacDonald 1976), the addition of a visual speech cue influenced the perception of the speech sounds for both the good and the poor readers. However, poor readers were found to be less categorical than good readers in audio speech perception and less accurate at visual speech perception (de Gelder & Vroomen 1998). Additionally, it was found that normal, non–alphabetic (i.e. logographic) readers had less categorical audio speech perception, and less accurate visual speech perception than alphabetic readers (de Gelder & Vroomen 1992). As the logographic group cannot be said to have any general auditory problems, this result further refutes the general auditory deficit hypothesis.

As noted by Mody et al. (1997), all of the above studies are limited in the extent to which they can test Tallal's 'general auditory deficit' hypothesis, because the "results cannot disprove the hypothesis: They can merely fail to support it where support would be expected" (p. 224). However, the evidence from these studies for a specific perceptual capacity dedicated to speech is both fairly wide–ranging, and fairly convincing. Based on this evidence, then, this study will continue on the assumption that the act of speech perception is performed by a system dedicated to speech, and therefore that the causal relationship between this system, and the system under which metaphonemic awareness develops, could theoretically go in either direction.

3.3 Speech perception and phonological awareness

Clearly the main drawback of the perception and literacy studies discussed above is the fact that none of them explicitly looked for a connection between the perceptual abilities of the subjects and their phonological or phonemic awareness. Studies which explicitly show a correlation between some aspect of speech perception and some aspect of phonological awareness development (including phonemic awareness), therefore, offer slightly more relevant support for Nittrouer's (1996*b*) findings. Unfortunately the number of studies which have directly examined the relationship between these two processes is quite small. A number of studies have shown *potential* correlations—for example, a group of subjects with both shallow categorical perception slopes and poor phoneme awareness skills—without specifically looking for a correlation between the two. One of these is a study by Flege, Walley & Randazza (1992) which examined English speaking adults' and children's perception of native and non–native vowel contrasts. The 'native' contrast was /ɪ/–/i/, and the 'non–native' was /ɪ/ to a vowel which the authors symbolised as /Y/. No age difference was found for the number of vowels identified as /ɪ/, but age differences were found in the slope of the perceptual response curves: 4– to 6–year–old children had shallower slopes than adults. The study then went on to show that this same age group (4– to 6–year–olds) had more difficulty than older children in both a rhyming task and a phoneme segmentation task. Flege et al. suggest, in conclusion, that these two results are related: "The slope differences may. . . have arisen from age–related differences in ability to perform perceptual tasks involving localised sound segments" (p. 2415)—i.e. that the children's inability to segment speech into phonemes influenced their speech perception.

Other studies which have tested both perceptual ability and phonological and/or phonemic awareness ability, without specifically testing a correlation between the two, include Nittrouer's (1999) study, and Joanisse et al.'s (1998) study, both discussed above. Both studies found that the group of children that had performed worse than, or differently to, the good readers on their perception tests, also had poor phonemic awareness. However, neither study attempted to determine to what extent the two processes correlated.

A number of studies, however, which do explicitly examine the relationship between speech perception and phonological and/or phonemic awareness, have been carried out recently by McBride–Chang and colleagues (McBride-Chang 1995*a*, McBride-Chang 1995*b*, McBride-Chang 1996, Manis et al. 1997, McBride-Chang et al. 1997). The main aim of the first of these studies (McBride-Chang 1995*b*) was to pull apart the component skills necessary for what McBride–Chang called phonological awareness—in fact, all of her phonological awareness tasks specifically tapped phonemic, rather than any other level of awareness and will be referred to as phonemic awareness tests from this point on.

McBride–Chang proposes that there should be at least three components to the successful completion of a phonological/phonemic awareness task besides the awareness itself. The first is speech perception—at one level this is important because in order to operate on a stimulus it must first be correctly perceived. At a more complex level, the successful access of the phonological structure of a stimulus requires that the phonological structure itself be correctly organised, which presumably might have required the successful development of the perceptual system. The second component is general cognitive ability—the subject must have the cognitive capacity both to understand the task and to think about the stimuli metalinguistically. Finally, the third component is short–term memory—the stimulus must be held in memory long enough for the metaphonological manipulation to be carried out. McBride-Chang's (1995*b*) study is a test of this proposal. Children from grade 3–4 (age approximately 8–9 years), were given tests of their phonemic awareness as well as their IQ, short–term memory and speech perception. The phonemic awareness tasks were a phoneme deletion task (e.g. "say the word 'melvz' without the 'v' "); a position analysis task (e.g. "what sound comes before/after the 'r' in 'fremps'?"); and a phoneme segmentation task. All used non–words. Three speech perception tests were carried out, all categorical perception, and all identification tasks. The stimuli for these tests were a 'bath–path' contrast (along a VOT continuum), a 'slit–split' contrast (along a gap duration continuum), and a 'ba–wa' contrast (along a continuum varying in length of formant onset transition). The *slope* of subjects' response curves was taken as the measure of categorical perception ability. Using structural equation modeling, which models relationships between processes, McBride-Chang (1995*b*) found a moderate relationship between speech perception and phonemic awareness. Importantly, speech perception was found to be associated with phonemic awareness "even after more complicated verbal abilities such as vocabulary (within general cognitive ability) and verbal short–term memory, have been accounted for" (McBride-Chang 1995*b*, p. 187).

McBride-Chang's (1996) study went on to expand the investigation of the relationship between perception and awareness to encompass reading ability. In this study, subjects were tested on their speech perception and phonemic awareness (in the same way as in the previous study), as well as on a number of other factors, including word reading. Again, using structural equation modeling as above, McBride–Chang tested the possible relationships between perception, awareness, reading, and the other parameters measured. The best model of the subjects' performance was one in which phonemic awareness was highly associated with both word reading and speech perception. Additionally, an alternative model of the relationship between the three processes in which a *diss*ociation was specified between speech perception and phonemic awareness, was found to be a significantly worse model of the subjects' behaviour, further emphasising the relationship between the two processes.

The third study in this set expanded the investigation of speech perception and phonemic awareness to dyslexic readers. Manis et al. (1997) tested dyslexics (from grade 4–10, age approximately 9–15 years), chronological–age–matched good readers, and reading–age–matched children on their categorical perception of a 'bath–path' continuum, and on their phonemic awareness (tested using a position analysis test, as described above). Consistent with the studies described above, the dyslexic subjects were significantly less proficient at the phonemic awareness task than their age–matched controls, but not the–reading–age matched controls. The dyslexics also showed significantly less categorical perception than the age–matched controls, and slightly less categorical perception (although this was not significant) than the reading–age matched controls. Manis et al. (1997) found a significant correlation between results of the phonemic awareness test and the categorical perception test. Interestingly, the authors note that McBride-Chang (1996) had found that "the best fitting model was one in which the relationship between speech perception and reading was mediated by phonological awareness" (Manis et al. 1997, p. 231). However, these authors go on to suggest that "It is possible. . . that causality runs the other way, i.e. that learning to read refines children's representations of speech" (p. 231).

Partly in order to test this possibility, McBride-Chang et al. (1997) carried out a longitudinal study of the development of phonological awareness and reading. The study took place over the course of approximately 18 months and all of the subjects were pre–readers at the beginning of the study. Once again, the subjects were tested on categorical perception of a 'bath–path' contrast. They were also tested on their ability on three phoneme awareness tasks: phoneme synthesis (in which the phonemes in a word are presented segmented and the child is asked to say what the word is—i.e. /k–a–t/ is 'cat'), phoneme elision (which is the same as the phoneme deletion task described above), and a 'sound isolation' task (in which the child is presented with two words, one of which is the rime of the other, e.g. 'pie', 'eye', and asked to identify the sound that is present in one and not the other). As would be predicted from McBride–Chang's previous studies, relatively strong associations were again found between speech perception and phonological awareness. McBride-Chang et al. (1997) also suggest again that the effect of speech perception on word reading may be mediated by its relationship to phonological processing. Additionally, analysis of the longitudinal aspect of the study showed that speech perception, cognitive ability, and verbal short–term memory *together* predicted 26% of the subjects' growth in, and 42% of their final ability in, the phoneme elision/deletion task.

It appears that the relationship between speech perception and phonemic awareness is more robust than that between speech perception and alphabetic reading skills, as these studies by McBride–Chang and colleagues show. This conclusion is supported, though to a lesser extent, by the studies which show a fairly consistent co-occurrence of poor categorical speech perception, or more global acoustic cue weighting strategies, and poor phonemic awareness (i.e. Flege et al. 1992, Joanisse et al. 1998, Nittrouer 1999). Additionally, the results of McBride–Chang's tests which showed that perception was only correlated with reading skill through the relationship of both to phonological or phonemic awareness, may go some way to explaining the limited correlation seen between speech perception ability and reading skill. However, all of these studies provide only partial support for the possibility that a relationship exists between acoustic cue weighting and phonemic awareness. For the evidence which clearly supports the existence of this relationship, we will have to turn to Nittrouer's (1996*b*) study.

3.4 The relationship between acoustic cue weighting and phonemic awareness

The main aim of Nittrouer's (1996*b*) study was to find support for her hypothesis that developmental changes in the weighting of acoustic cues were due to a shift in perceptual strategy from one that was syllable–based to one that was phoneme–based. Nittrouer believed that if she could show that shifts in cue weighting were related to the development of phonemic awareness, which also involved a shift from syllables to phonemes, this would lend credence to her claim. The study *did* find a relationship between perceptual weighting and phonemic awareness. However, while this can be taken as support for Nittrouer's syllable–to–phoneme hypothesis, the study also brings up issues about other aspects of Nittrouer's DWS model. Specifically, Nittrouer hoped that this study would also provide support for the view that shifts in acoustic cue weighting were "based on linguistic experience" (Nittrouer 1996*b*, p. 1061). However, the results of the study raise more questions about the strictly maturational nature of the DWS than they actually answer.

Nittrouer 1996

The experimental subjects for Nittrouer's (1996*b*) study were 7– to 8–year–old children from the following backgrounds: (i) children with what was classed as a significant history of otitis media (i.e. ear infections, significant being defined as having had 6 or more documented infections before the age of 3 years, and/or having had myringotomy tubes inserted before the age of 3 years), (ii) children from a low socioeconomic background (defined as having an annual family income of less than $15,000 US), or (iii) children with *both* a significant history of otitis media, and a low socioeconomic background. Nittrouer chose subjects with these backgrounds because she believed they could be classified as having had diminished linguistic experience (Nittrouer cites studies by Eimas & Clarkson (1986), Gravel & Wallace (1992), Raz & Bryant (1990) and Wallach, Wallach, Dozier & Kaplan (1977), among others, that report language delays, speech perception difficulties, and difficulties with phonemic awareness for these groups). The control children were from mid socioeconomic backgrounds and had no significant histories of otitis media.

All of these children were tested on their phonemic awareness and their acoustic cue weighting. The contrasts for the cue weighting tests were the /sa/–/ʃa/ and /su/–/ʃu/ contrasts from Nittrouer (1992). Measures of both the slope and the separation of the response curves were taken for all subjects.

Two types of task were used to test phonemic awareness. The first was a phoneme deletion task, in which the subject was presented with a nonsense word, and was asked to delete a given phoneme to make a real word (the example given is "say /pɪnt/ without the 't'." Nittrouer 1996*b*, p. 1063). The second task was a modified 'pig latin' task. Playground pig latin involves a modification of words at the onset–rime level, so that the onset of the word is placed after the rime, and the vowel /ɛɪ/ is added: 'star' thus becomes /arstɛɪ/. In the modified version of this task, the first phoneme of the word was moved rather than the whole onset: 'star' thus becomes /tarsɛɪ/ (Nittrouer 1996*b*).

The results showed, first, that all of the experimental groups had significantly less phonemic awareness than the control group. Within the experimental group, the children from low socioeconomic backgrounds (both with and without otitis media) had worse phonemic awareness than the children with histories of otitis media. The results also showed that, for the perceptual tests, children in the experimental groups had shallower response slopes, and more widely separated response curves than the control group. This is taken by Nittrouer as an indication that these children perceive more in terms of syllables, like the younger normally developing children in her previous studies, than in terms of phonemes, like the older children and adults. Again, this perceptual behaviour is more extreme in the children from low socioeconomic backgrounds (both with and without otitis media) than in the children with histories of otitis media.

The most important result for the current study, however, is the finding that the phonemic awareness scores and the degree of transitional vs. fricative cue weighting (measured in terms of the separation of the response curves) are significantly correlated. Specifically, as the separation decreased (indicating, according to Nittrouer, a shift from syllable perception to phoneme perception) success at the phonemic awareness tasks increased.

Nittrouer takes this finding as support for at least one of the claims of the DWS model—that the "developmental weighting shift is related to developmental increases in sensitivity to phonetic structure" (p. 1061)—i.e. is related to a shift from global, syllable–based perception to analytical, phoneme–based perception. Nittrouer then goes on to take the finding that children with histories of otitis media, and those from low socioeconomic backgrounds, have lower phonemic awareness scores and more global speech perception strategies as support for the second hypothesis being tested: the claim that "experience with a native language provides the child with the opportunities to develop both mature perceptual weighting strategies and phonemic awareness" (p. 1068). However, as will be seen in the following section, this is not strictly speaking a conclusion that can be drawn from these results.

Issues

It is clear from Nittrouer's (1996*b*) study that some sort of diminished experience with language may impinge on a child's ability to develop analytical perceptual weighting, and/or good phonemic awareness. What we cannot conclude from this study, however, is that undiminished language experience *must* lead to both analytical speech perception strategies and good phonemic awareness. In the case of phonemic awareness we know that it is rare for maturation alone to lead to phonemic awareness development. We should also not then assume that maturation always leads to analytical speech perception.

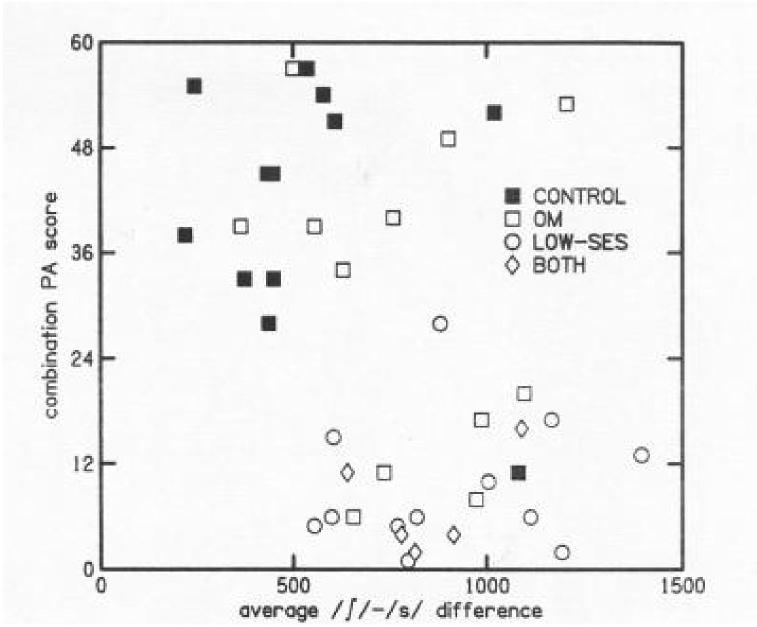

Figure 2.10: **Graph of perception and awareness results from Nittrouer's (1996*b*) study.** Phonemic awareness is displayed on the y-axis. Cue weighting strategy is displayed on the x-axis in terms of separation of response curves (here called **/ʃ/–/s/ difference**): analytical strategies can be found towards the left of the graph, global strategies can be found towards the right of the graph. Figure from Nittrouer (1996*b*, p. 1067) ©American Speech–Language–Hearing Association. Reprinted by permission.

Nittrouer (1996*b*) herself acknowledges that there are two possible explanations for the results of her study. Either "the development of mature perceptual weighting strategies [i.e. less global speech perception] leads to phonemic awareness or... the cognitive demands of improving phonemic awareness forces the child to develop more effective perceptual weighting strategies" (Nittrouer 1996*b*, p. 1067). The cross–sectional design of Nittrouer's (1996*b*) study prevents us from using the results of the study itself to address this issue. However, we can use these results to develop theories regarding the answer to this question.

The relationship between acoustic cue weighting and phonemic awareness is illustrated in Figure 2.10. The first thing to note about the graph is that the distribution of the data points is fairly bimodal: generally all those children with poor phonemic awareness also have very global perceptual strategies—this group includes *all* of the children from low socioeconomic backgrounds, plus some of the children with otitis media, and one child from the control group. The overwhelming majority of the control group, and most of the rest of the otitis media group are in the opposite corner of the graph, with both good phonemic awareness scores and more analytical perceptual strategies.

This type of distribution does put a restriction on the conclusions, or even hypotheses that can be drawn from this study. A more varied distribution—one in which there were a reasonable number of children who did not have both good phonemic awareness and analytical perceptual strategies, or both poor phonemic awareness and global perceptual strategies, would have been more informative. If a number of children had developed one process but not the other, we could hazard a guess that this process develops before the other, and might play a causal role in the development of the other. It should be noted at this point that the bimodal distribution seen in these results is largely due to Nittrouer's choice of subjects. All of the subjects were 7– to 8–years–old, which means that they had received 1–2 years of literacy instruction—enough time for the subjects in the control group at least to have developed very good phonemic awareness. This means that there is little chance of observing any of these subjects' behaviour part–way through their development. An additional difficulty with the choice of subjects is the fact that the experimental groups have been found to have such wide–ranging speech and language problems. This makes it difficult to hypothesise as to which deficit might have caused the other.

However, some information *can* be gleaned from the graph in Figure 2.10. While the distribution of data is *predominantly* bimodal it is not *completely* bimodal. In the top right hand corner of the graph are two data points, representing one control and one otitis media subject. The fact that these data points are at this point on the graph means that these two subjects had good phonemic awareness, but had very global speech perception strategies. Importantly, there are no data points in the opposite corner of the graph, meaning that there were no children in this study who developed analytical speech perception strategies without having developed good phonemic awareness.

What can be hypothesised from this? Nittrouer (1996*b*) suggests that

> The finding that two children in this study demonstrated good phonemic awareness, even though their perceptual weighting strategies were among the less mature [i.e. less analytical], might suggest that discovering syllable–internal structure may create pressure to develop the most effective processing strategies for providing access to that structure. (pp. 1067–1068)

Is it possible that the development of phonemic awareness could impact on perceptual strategies? The studies of McBride–Chang and colleagues (McBride-Chang 1995*b*, McBride-Chang 1996, McBride-Chang et al. 1997, Manis et al. 1997) who found speech perception (along with memory and cognitive ability) to be predictive of later phonemic awareness, would suggest not. The findings from those studies support the more conventional hypothesis that speech perception is a maturational process, the development of which *allows* other processes to develop—processes from speech understanding to metalinguistic awareness. However, while it is not the conventional view, the idea that certain perceptual strategies could develop *because* of the development of phonemic awareness and/or alphabetic literacy has been proposed by other authors, notably Flege et al. (1992) and de Gelder & Vroomen (1992) (discussed above) and Fowler (1991) and Morais & Kolinsky (1994) (discussed at the end of the previous chapter).

The next question to be asked, then, is what impact each of these hypotheses would have on Nittrouer's Developmental Weighting Shift model. Clearly the theory that perceptual weighting changes maturationally, and in doing so allows phonemic awareness to develop, fits most closely with Nittrouer's proposal that the DWS is a developmental or maturational model. In order to accept the theory that shifts in perceptual weighting are caused by the development of phonemic awareness, on the other hand, Nittrouer would have to abandon the maturational aspect of the Developmental Weighting Shift for a view which allows higher, conscious cognitive development to influence lower level, subconscious processes like speech perception.

The fact that there is at least some support for both sides of this argument, and, the additional fact that Nittrouer's (1996*b*) study is inconclusive in this regard, means that the question of possible causality between acoustic cue weighting and phonemic awareness remains to be empirically addressed. The main goal of this thesis, therefore, will be to attempt by means of two studies to begin to answer this question.

4 Hypotheses

The general aim of this thesis is to attempt answer the following question:

- **What is the causal direction (if any) of the relationship between changes in acoustic cue weighting in speech perception, and the development of phonemic awareness?**

The two competing hypotheses for the outcome of the thesis will therefore be:

- **Children's perception will always undergo a shift in acoustic cue weighting from global to analytical strategies before the onset of the development of phonemic awareness. Shifts in acoustic cue weighting will predict later ability in phonemic awareness.**

- **Phonemic awareness will always develop before shifts in acoustic cue weighting take place. Ability in phonemic awareness will predict later shifts in acoustic cue weighting.**

The following chapter will describe in more detail the methodologies that will be adopted in order to evaluate these two hypotheses.

CHAPTER 3

Methodological issues

1 Experimental design

The main goal of this thesis is to determine the *nature* and *direction* of the relationship between acoustic cue weighting and phonemic awareness. In particular, the intention is to discover whether acoustic cue weighting is a pre–requisite skill for the development of phonemic awareness, or whether the development of phonemic awareness impacts on cue weighting strategies.

Nittrouer's (1996*b*) study was unable to address this issue, for two main reasons. The first of these was the cross–sectional design of her study. Cross–sectional paradigms allow for the researcher to determine whether a correlation exists between two processes, but say "little about the direction of causality between the development of these processes" (Nittrouer 1996*b*, p. 1067). An alternative paradigm which does allow for a more extensive evaluation of possible causal directions is a longitudinal study. In this type of study, subjects' behaviour is tracked over an extended period of time, allowing for observations to be made of the gradual development of the processes in question. Although this type of study may not always allow for conclusive claims to be made regarding causality, it does allow such claims to be constrained: if perception consistently develops before awareness, for instance, it is very unlikely that awareness would play a causal role in the development of perceptual strategies. Longitudinal studies also allow for statistical analyses of the predictive relationship between processes—that is, given test results for two processes over a period of time, it is possible to determine to what extent specific results are dependent on the development of a process at a previous point in time.

The second problem with Nittrouer's (1996*b*) study was (as noted in Chapter 2, Section 3.4), the choice of experimental subject groups. Although Nittrouer noted that otitis media and low SES have been shown to affect speech and language development, the exact nature of this effect appears to be diffuse across a number of different areas of speech and language. This means that in examining the perceptual and metaphonemic behaviour of children from these groups, it is impossible to determine which of the two processes might have had a causal effect on the other. The solution to this problem is to make use of an experimental group of *normally* developing children who have not yet developed one of the processes in question. Clearly it is not possible to find a group of normally developing children who also have had no chance to develop perceptually: this was the problem that faced Nittrouer (1996*b*) (who stated that "The ideal experimental paradigm would have included two groups of children from identical backgrounds, but one group would have had all linguistic input withheld for all of their short lives" (p. 1061)). It *is*, however, possible to find normally developing children who have not yet acquired phonemic awareness. Presumably, any child that has not had access to alphabetic print will be unlikely to have developed phonemic awareness (leaving aside the cases of children who may develop phonemic awareness as a result of speech therapy, and the rarer cases of spontaneous phonemic awareness development). Most children in literate societies are trained in reading and reading–readiness from the beginning of formal education: conventional state primary and fee paying schools in the Edinburgh area, for example, begin such training at approximately 4;6–5;6 (years;months) in at least the first year of school (Scottish Office Education Department 1991). There are, however, a number of independent school systems that delay all forms of literacy training and preparation until later: for example, the Steiner and Waldorf school systems do not start literacy training until children are approximately 8 years of age. It is very unlikely that children in these or similar school systems would begin to develop phonemic awareness until after they begin this training. Such children could therefore be considered an appropriate normally developing experimental group when compared to children attending more conventional schools.

This thesis made use of both of these solutions, in two different studies. Experiment 1 was a longitudinal study, and followed the development of acoustic cue weighting and phonemic awareness in a group of normally developing children in their first year of full–time education in conventional state primary or fee–paying schools (referred to as the 'beginning–reading' group). These children were tested three times (referred to as Sessions 1–3), at regular intervals, allowing for assessment of the development of the two processes in relation to each other. Following Nittrouer and colleagues (1987, 1992) a group of adults was also tested on their acoustic cue weighting strategies, for comparison with the children. Experiment 2 was a cross–sectional investigation of children from an independent school which delays reading/reading–readiness training (referred to as the 'reading–training–delayed' group). This group of children were the same age as the children at the *end* of Experiment 1, but had not yet begun any literacy training. It should therefore be possible to compare the cue weighting strategies and phonemic awareness abilities of this (predominantly) non–reading group with their same–age beginning–reading peers from Experiment 1.

2 Stimulus and test design

In Chapter 2, a number of issues regarding Nittrouer's DWS model and the methods used to test acoustic cue weighting were outlined. Issues were also raised regarding the testing of phonemic awareness. In the design of the stimuli for the studies in this thesis, as many of these issues as possible were taken into account.

2.1 Acoustic cue weighting

Synthetic vs. natural stimuli

Nittrouer and colleagues have made use of two basic types of stimuli for their acoustic cue weighting studies: hybrid stimuli—that is, part synthetic and part natural speech—and highly stylised synthetic stimuli. Both of these are potentially problematic for cue weighting studies, as both make use of designs which could render the transitional cues more salient to children than these cues would be in natural speech. In the hybrid stimuli the transitional cues could be considered to be more salient because they are natural speech (the fricative noises were synthetic), while in the all–synthetic stimuli, the transitional cues could be considered more salient due to their highly stylised nature (see Hazan & Rosen 1991).

The stimuli to be used in this study were therefore created by means of a method called 'copy–synthesis' (Hazan & Rosen 1991, see also Liberman 1996). In this method, highly detailed acoustic analyses of natural speech are made, and the resulting values then used to synthesise the stimuli. As noted by Hazan & Rosen (1991, p. 198) "because of the similarity in level of pattern complexity to natural speech [results obtained with this method] would be more representative of natural speech processing than those that are obtained with highly simplified synthetic stimuli." If Nittrouer's interpretation of the changes in cue weighting is correct, listeners should display similar changes for copy–synthesised stimuli to those they display for hybrid or stylised synthetic stimuli.

Limited experimental evidence: phonetic context

The contrast used to test cue weighting in these studies was the contrast between the fricatives /s/ and /ʃ/ in fricative–vowel single–syllable words. As noted in Chapter 2, Section 1.2, Nittrouer and colleagues have suggested that the transitional effect—that is, the separation between the response curves—is proportional to the extent of the transitions. In order to ensure, therefore, that the greatest possible effect of the transitional cues would be seen in the current studies, a vowel context with extensive transitions was required. Of those contexts previously studied by Nittrouer, the one with the most extensive transitions following /s/ and /ʃ/, and which engendered correspondingly greater weight from the children, is /u/. However, there is a difference in the subject population between the current studies and those of Nittrouer and colleagues that makes the use of the /u/ context problematic. While the subjects in Nittrouer's studies spoke Standard American English, the predominant dialect of the subjects in the current studies was Standard Scottish English. For speakers of Scottish English /u/ is generally not a back vowel, as it is in most American dialects, but a front vowel (McClure 1995). This means that /u/ in Standard Scottish English is likely to have transitions following /s/ and /ʃ/ which are more like /i/, which was the context that engendered the smallest transitional effect in Nittrouer & Studdert-Kennedy's (1987) study. Therefore, the current studies made use of a vowel context which has more appropriately extensive transitions following /s/ and /ʃ/ in Scottish English: the back, rounded /o/. This has the additional advantage of being a context which has not been previously studied by Nittrouer. As a result, it served as a small test of the generalisability of the phenomenon of shifts in cue weighting.

Auditory processing

A two–item forced–choice labelling task was chosen as the perceptual testing method for this study. The main reason for choosing a labelling task over a discrimination task was to allow ease of comparison between the results of this study and those of Nittrouer and colleagues' studies (Nittrouer & Studdert-Kennedy 1987, Nittrouer 1992, Nittrouer 1996*b*). Additionally, as was noted in Chapter 1 (Section 1.3), there is some debate as to whether labeling tasks and discrimination tasks tap the same perceptual processes. Simon & Fourcin (1978), for instance, have suggested that only labeling tasks require a *linguistic* decision to be made regarding the stimuli. The hypotheses that will be evaluated in this thesis depend very much on the relationship between children's conscious knowledge of linguistic categories (i.e. phonemic segments) and their speech perception. Therefore, a perceptual task which is presumed to access on–line, subconscious linguistic knowledge, like a labeling task, might be more relevant than one which might be more related to auditory sensitivity.

2.2 *Phonemic awareness*

What constitutes a phonemic awareness task?

McBride-Chang (1995*b*) found that the memory load of a phonemic awareness task, and the general cognitive ability of the test subject have significant effects on phonemic awareness performance. The results of McBride–Chang's study showed that the higher the memory and cognitive demands of the task, the worse will be a subject's performance, regardless of their basic awareness of phonemes. McBride–Chang reported these effects for children aged 8–9 years. As the children in the current studies ranged from 5;2–7;7 care was taken in the testing of these groups, whose memory and cognitive skills were likely to be still less developed than those of the older children tested by McBride–Chang.

In choosing the words for the phonemic awareness test stimuli, a decision was made to use real word stimuli, for memory load reasons. There is some evidence that nonsense words afford a more conclusive test of phonemic awareness than real words: Tunmer & Nesdale (1982), for instance, suggest that "Beginning readers who lack phonemic segmentation skills but read words by sight may. . . resort to a 'grapheme' strategy" (p. 301) in phonemic awareness tasks. Such a strategy, they say, would result in children 'overshooting' the number of phonemes in a word in cases where phonemes are represented by digraphs, e.g. 'tea' has two phonemes: /t/—/i/, but three graphemes: t—e—a. In terms of phonemic awareness *testing*, the possibility that some children may be using a grapheme–strategy causes problems when the number of phonemes in a word is equal to the number of graphemes, as in this case it is impossible to determine if the child has mastered phonemic awareness, or has simply resorted to a spelling strategy. Tunmer & Nesdale (1982) propose that nonsense word testing would eliminate this problem, because children would not have any pre–learned grapheme rules for the stimuli.

However, work on short term memory suggests that new or unfamiliar words (including nonsense words) carry a greater memory load (in particular phonological memory) than familiar words (Gathercole & Baddeley 1998). Therefore, in order to minimise the memory load for the children in the current studies, real words were used to test phonemic awareness. Keeping in mind Tunmer's suggestion that real words encourage the use of graphemic strategies, at least half of the stimuli for each of the phoneme blending and phoneme segmentation tests (25/50 and 26/50) made use of di– and trigraphs (e.g. [ʃip]—'sheep', [katʃ]—'catch'). It should therefore be evident from subjects' responses to these stimuli whether or not they have a tendency to make use of graphemic strategies. Additionally, the use of real words is slightly less of an issue for this particular study than for those that have used older subjects (e.g. McBride-Chang 1995*b*, Tunmer & Nesdale 1982). The subjects in both of the current experiments were predominantly pre–literate or only beginning–literate: it was therefore very unlikely that many of the words to be used for the tests would be in their reading or spelling vocabularies, even for those beginning–readers who were only reading by sight.

There are a large number of possible tasks available for testing phonemic awareness, each of which can also put different demands on a subject's memory and cognitive abilities. Nittrouer (1996*b*), for example, made use of a phoneme deletion and a modified 'pig latin' task. Other studies have made use of different tasks: e.g. phoneme tapping (Liberman et al. 1974), in which the subject is taught to tap out the number of phonemes in a word, and is then asked to tap novel words; substitution (Goswami & Bryant 1990), in which the subject is asked to replace part of a word with another sound or set of sounds (e.g. 'lug' becomes 'fog', 'fli' becomes 'fru'); phoneme oddity tasks, in which the subject is presented with three words, two of which share a phoneme (e.g. 'mop', 'lead', 'whip'), and asked to find the odd one out (in this case 'lead', see Treiman & Zukowski 1996); position analysis, in which a word or non–word is presented, along with a single phoneme from that word, and the subject is asked to say what sound came before or after the presented phoneme in the stimulus word (McBride-Chang 1995*b*).

In addition to requiring that the subject understand the concept of a phonemic unit, most of these tasks also place heavy demands on the subject's cognitive abilities and short term memory. *Position analysis*, for instance, requires the subject to i) segment a word into a sequence of phonemes, ii) hold this sequence in memory long enough to locate the target phoneme in the sequence, iii) understand the concepts of 'before' and 'after', and iv) remember the order of the sequence of phonemes so that the sound that comes before or after the target phoneme can be identified. Likewise, Nittrouer's modified *pig latin* task (Nittrouer 1996*b*) requires a large number of manipulations: the subject must i) segment the word into (at least) two units corresponding to 'initial phoneme' and 'rest of word', ii) hold both units in memory long enough to transpose them, iii) blend the transposed units together and hold this new nonsense word in memory, iv) remember the tag segment /ɛɪ/ ('ay'), and v) blend the tag segment onto the end of the nonsense word formed by the transposition of the original units. Even the relatively straightforward *oddity* task requires that three words be held in memory, and that a complex comparison be made of these words.

In order to minimise the memory and cognitive demands of the tasks in the current studies, two of the three assessments which were used in the current study were tests which require awareness of phonemes, and only *one* basic manipulation of these units. For one of these tests, phoneme segmentation, this manipulation is the division of a word into a sequence of separate phonemes, e.g. 'phone' is /f–o–n/. For the other test, phoneme blending, the manipulation required is the re–synthesis of a number of separate phonemes into a single word, e.g. /f–o–n/ is 'phone'. A third, and slightly more cognitively demanding test was also used: phoneme deletion, which requires the understanding of the concept *subtraction* (removal of part of a whole) as well as the ability to segment phonemes. The cognitive and memory demands of this task were, however, minimised, first by ensuring that the phoneme to be deleted was in the same position within the word for all stimuli, and second by ensuring that both the stimulus word, and the word which was formed when the phoneme was deleted, were real words (e.g. 'snow' with the /s/ removed is 'no').

Finally, in terms of the division suggested by Morais (1991), and to an extent Fowler (1991), all of the tasks used in these studies required explicit manipulation of phonemic segments, rather than implicit recognition of phonemes.

How aware is aware?

Because the first of the two studies in this thesis was a longitudinal study, the children's success at the phonemic awareness tasks was measured predominantly in terms of their progression, rather than in terms of one specific cut–off point. Awareness was therefore considered on a scale from less aware (0% correct) to more aware (100% correct). However, as will be seen in the following section, which describes the graphics that used to represent the results, it is useful to have some fixed point against which to measure the subjects' progression. For the purpose of these studies, this fixed point was the *median* score of the subjects in Experiment 1 at Session 1 of the longitudinal study. In this way, the subjects were measured for success against their own initial ability. Additionally, this fixed point was also used in Experiment 2, so that the ability of this second group of children could be compared against the ability of the group from Experiment 1.

The question of how to gauge awareness applies not only to overall score, but also to the scores given to individual items in a test. The tests themselves were scored in a very basic way: the subjects received one point for every correct response. A correct response for the phoneme blending was simply the correct identification of the segmented word, with all the phonemes in the correct order, thus responding 'room' to the presentation of /r–ʊ–m/ was correct; responding 'arm' was incorrect. The subjects were required to respond to the phoneme segmentation test with segment sounds rather than letter names, and all these sounds were required to be present and in the correct order. Thus, the response /k–l–æ–p/ for the word 'clap' was correct; the response /k–æ–p/ was incorrect, as was /k–æ–l–p/. Additionally, a response in which not all phonemes were segmented was also incorrect, thus /kl–æ–p/ did not receive a point. Incorrect vowel quality was not given a point, however subjects were given a point whether they segmented diphthongs as one sound or two, thus for 'mouse,' both /m–aʊ–s/ and /m–æ–ʊ–s/ were correct. Dialectal variation was taken into account, thus both /t–r–eɪ–n/ and /tʃ–r–eɪ–n/ for 'train' were counted as correct. Additionally, there was some flexibility in terms of whether the subject approached the task from a phonetic or a phonological point of view. For the word 'space', therefore, both /s–p–ɛɪ–s/ and /s–b–ɛɪ–s/ were counted as correct. This last point also applied to the phoneme deletion task: in those cases where the deletion of an initial phonemic segment left two possible results depending on whether a phonetic or phonemic approach was taken, both were considered correct. Thus for the phoneme deletion task, the deletion of the /s/ from the word 'spot' could produce either /p–ɔ–t/ or /b–ɔ–t/ correctly.

3 Preliminary analysis methods

This section will describe in detail the graphics used to represent the results of the cue weighting and phonemic awareness tests, and the methods used to interpret these representations.

3.1 Acoustic cue weighting

Figure 3.1 shows a hypothetical set of response curves for the type of perceptual stimuli which were used in this study.

The data from the cue weighting tests was normalised using a probit transformation. This is a normalizing transformation, which extracts rate–of–change information from data on an S–shaped curve (Cohen & Cohen 1983).

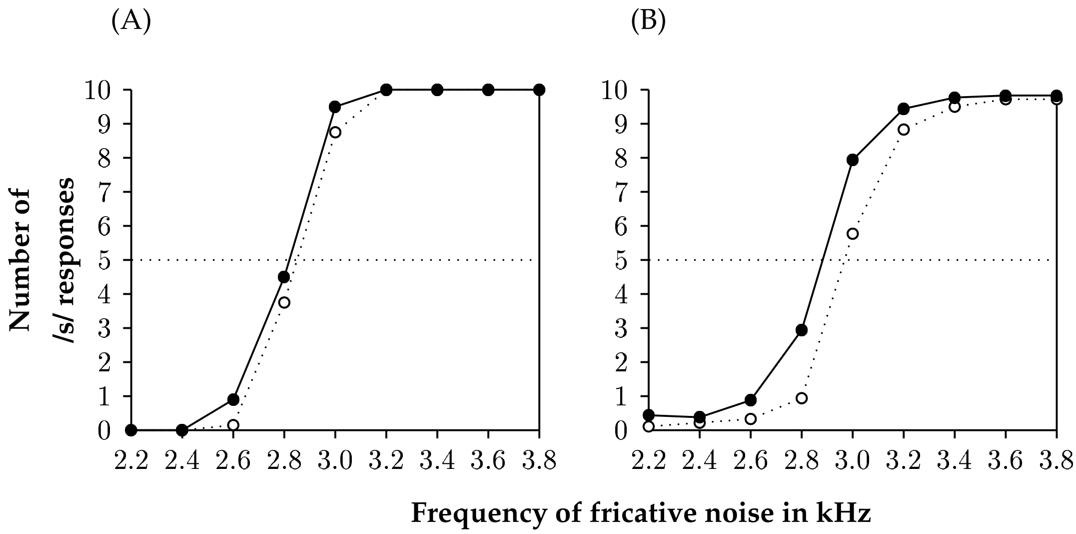

Figure 3.1: **Hypothetical perceptual responses to the /ʃo/–/so/ stimuli used in the current study.** The x-axis shows the continua of fricative noises, ranging in frequency from 2.2kHz (the most /ʃ/–like) to 3.8kHz (the most /s/–like). On both graphs, the solid line represents a listener's /s/ responses to stimuli with /s/–transitions; the dotted line represents the same listener's /s/ responses to stimuli with /ʃ/–transitions. Graph (A) represents responses of a listener with analytical cue weighting strategies. Graph (B) represents the responses of a listener with more global cue weighting strategies.

As noted in Chapter 2, Section 1.1, this transform gives estimates of values which can be used to describe each set of response curves. To re–cap, these values are: the *mean* of the responses, the *slope* of the response curves, and the *separation* of the response curves.

The mean of this transform is equivalent to the point along the fricative continuum (the x-axis) at which the /s/ responses reach 50%. In Figure 3.1 this corresponds to the point (on the x–axis) at which the response curves cross the horizontal dotted line. The mean therefore represents the point along the continuum at which the responses cease to be predominantly /ʃ/ and begin to be predominantly /s/—i.e. the phoneme boundary between /ʃ/ and /s/.

The slope of the response curve is obtained by taking the reciprocal of the standard deviation (Cohen & Cohen 1983). Slope is representative of the rate at which the above change (from predominantly /ʃ/ responses to predominantly /s/ responses) takes place. In Figure 3.1, graph (A) shows a response curve which changes fairly rapidly, while graph (B) shows a response curve with a more gradual slope. The slope of the response curve can therefore be seen to be equivalent to the degree of categorical–ness of the response.

The separation of the response curves is obtained 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 difference between the two continua. In Figure 3.1, graph (A) shows a set of response curves with very little separation between the curves, and graph (B) shows a set of response curves with a larger separation between the response curves. Separation can be seen as representative of the transitional effect—i.e. the extent to which the subject attended to, or weighted, the transitional information.

The use of probit transformations has the primary benefit of allowing direct comparison of the results of this study with those of Nittrouer and colleagues (Nittrouer & Studdert-Kennedy 1987, Nittrouer 1992, Nittrouer 1996*b*), who made use of the same transform. Additionally, probit is a well accepted method of normalising data of this type (Cohen & Cohen 1983) and is well established in the field of speech perception (see Liberman 1996). The use of this transform allows for the relationship between numerous rather complex response curves to be described and analysed by means of three fairly transparent values: the mean, slope, and separation of the response curves.

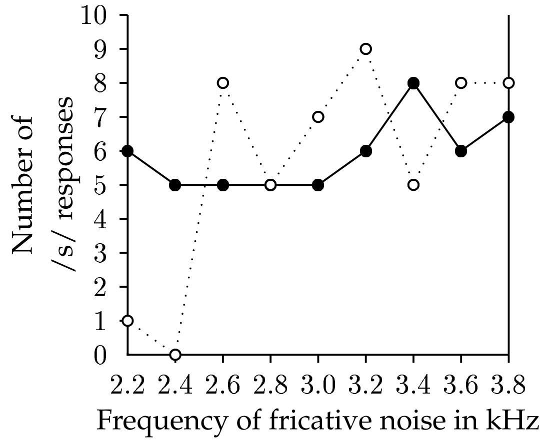

Figure 3.2: **Set of random response curves.** This graph represents the responses of a single listener. The x-axis shows the continua of fricative noises, ranging in frequency from 2.2kHz (the most /ʃ/–like) to 3.8kHz (the most /s/–like). The solid line represents this listener's responses to /s/ responses to stimuli with /s/–transitions; the dotted line represents the same listener's responses to stimuli with /ʃ/–transitions.

It should be noted that probit transformation expects a distribution of data which fairly closely resembles an S–shaped curve (Cohen & Cohen 1983)—only those response curves which conform to this pattern can be reasonably analysed using probit. This restriction had to be borne in mind in the analysis of data from the child subjects, as children's perceptual responses can be quite variable (Hazan & Rosen 1991, Hazan & Barrett 1999, Simon & Fourcin 1978). In order for data from any one subject to be included in analysis, the following criterion had to be met: if the subject's two response curves crossed each other more than twice, those response curves were considered to be non–S–shaped, and that subject's data was not included. A set of random response curves from a child who was tested for, and subsequently eliminated from, Experiment 1 is displayed in Figure 3.2.

3.2 Relationship between perception and awareness

The relationship between perception and awareness, for each session in Experiment 1 and for Experiment 2, will be displayed on a graph such as that shown in Figure 3.3.

In these graphs, phonemic awareness is plotted on the y–axis, in terms of a raw score on one of the phonemic awareness tests (from 0 to 40 or 50, depending on the test). Cue weighting is plotted on the x–axis, in terms of the *separation* (in kHz) between the response curves. It is extremely important to note that an analytical performance on a cue weighting test is characterised, according to

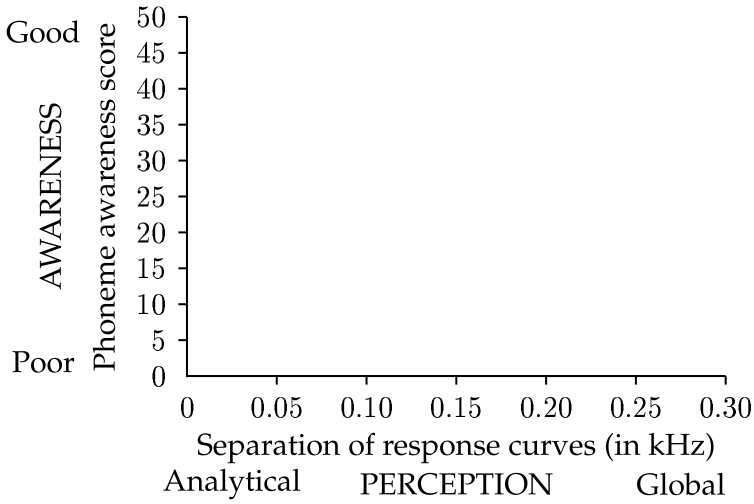

Figure 3.3: **Example of the graph that will be used to display the relationship between perception and awareness.** Awareness is plotted on the y–axis in terms of a raw phonemic awareness score; perception is plotted on the x–axis in terms of the separation (in kHz) between the response curves.

Nittrouer, by a heavier weighting of the fricative noise cues, and thus by *smaller* degrees of separation between response curves. Global cue weighting is therefore found at the right of the graph, moving away from the origin, and analytical cue weighting is found at the left, moving towards the origin and a separation between the response curves of zero.

Because both phonemic awareness development and perceptual weighting changes are gradual processes which take place at varying rates for different children, the children in these studies were not all at the same level of development at each session in either experiment. As a result, each session in both studies can be seen as a small cross–sectional study of the population being tested, inasmuch as at any one session performances ranged from poor to good, and global to analytical. Keeping this in mind, therefore, it is necessary to have a simple means of describing the performance of each child at any session in the study. The simplest method of categorisation is to divide performance in each process into 'poor' and 'good' (for awareness) or 'global' and 'analytical' (for perception): see Figure 3.4. The performance of each subject can therefore be labeled according to the quadrant into which their responses fall.

The division of the graph into quadrants is made at specific points on the x– and y–axes. For cue weighting, the division is made at the point which corresponds to the largest difference between category boundaries displayed by an adult subject—i.e. the most global of the adult responses. The 'analytical' half of the graph can therefore be considered to cover the range of possible adult responses to the stimuli. The division for phonemic awareness is made at the median score for all children on the relevant phonemic awareness test *in the first test session in Experiment 1*. These two division points were maintained from Session 1 throughout Experiment 1 so that the children's performance at subsequent sessions of the study could be tracked in terms of their progression from these initial starting points. The same division points were also used for Experiment 2 so that these subjects' performance could be more easily compared to that of the subjects from Experiment 1.

In addition to examining each of the sessions individually, we can also track the progress of the four groups of subjects across the three sessions in the study. On all three graphs, each of the data points are given one of four symbols—open circle, filled circle, open triangle, filled triangle—each of which corresponds to the position of that data point at *Session 1* of the study. This effectively allows us to give each subject one of four labels corresponding to their initial position. By doing this it is possible to follow the development of acoustic cue weighting and phonemic awareness in each of these four groups of subjects, from that first session through the subsequent sessions in the study.

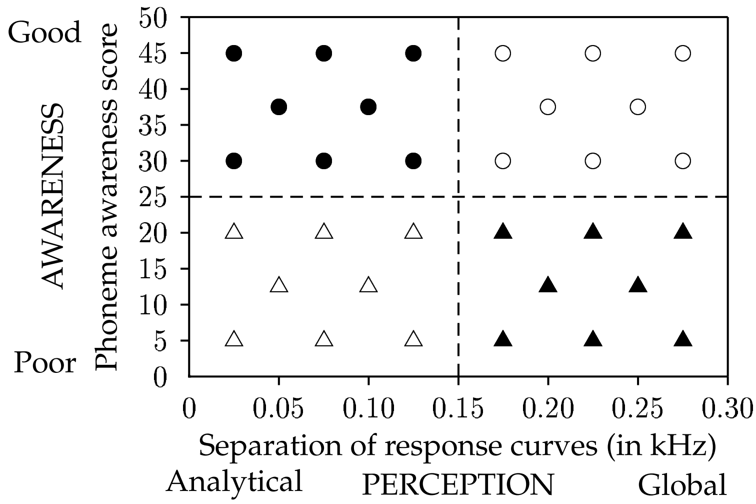

Figure 3.4: **Example of the graph that will be used to display the relationship between perception and awareness.** The graph is divided into quadrants at the median phonemic awareness score (on the y-axis) and the most global of the adults' cue weighting responses (on the x-axis). Each point on the graph represents one subject. The points in each quadrant are given a different symbol so that the performance of each group of subjects can be tracked.

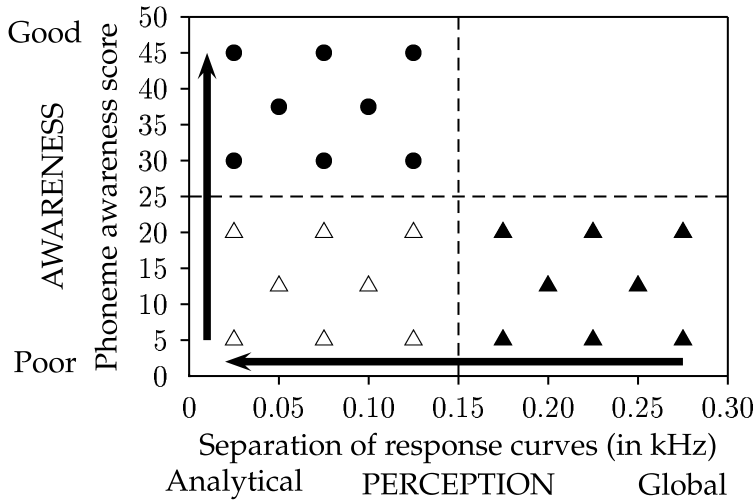

Figure 3.5: **Distribution of data that would be expected if cue weighting strategies change before the development of phonemic awareness.** Note that the arrows indicate the direction that the data points would be expected to move at subsequent sessions.

Predicted relationships

The hypotheses presented in Chapter 2 predict two alternative relationships between cue weighting and phonemic awareness. As noted above, development of either of these is unlikely to be instantaneous for any one subject, or simultaneous across a group of subjects. Keeping this in mind, it should be the case that if data is collected at any time *while both processes are still developing*, the distribution of data should reflect one of these relationships.

The graphs below illustrate these possible distributions. The arrow on each graph indicates the direction that the data points should move at each subsequent session in the study in each case (note that each data point represents one subject).

HYPOTHESIS 1

Hypothesis 1 states that acoustic cue weighting develops first, maturationally, thus facilitating the later development of phonemic awareness. If this is the case, then the data should be expected to be distributed as illustrated in Figure 3.5.

In this relationship some of the children will have developed both good phonemic awareness and analytical cue weighting, while others will have both poor awareness and global cue weighting. There will also be some children who are

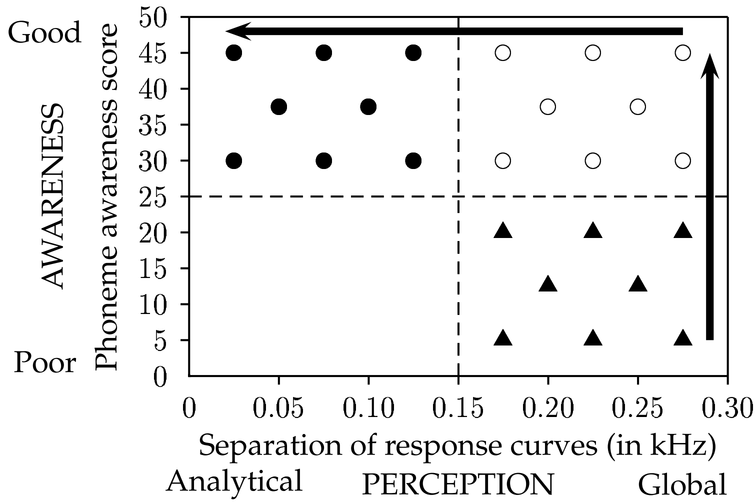

Figure 3.6: **Distribution of data that would be expected if phonemic awareness develops before cue weighting strategies change.** Note that the arrows indicate the direction that the data points would be expected to move at subsequent sessions.

'in transit' between these two groups: children who have already developed analytical cue weighting, but who have not yet developed good awareness. Importantly, we should *not* expect to see any children with extremely good awareness scores who still have global cue weighting.

Over the course of the longitudinal study, the data points should follow the direction of the arrows. That is, the data points representing any children with poor phonemic awareness and global cue weighting should first move left, into the quadrant that corresponds to analytical strategies–plus–poor phonemic awareness. At a subsequent session these same data points should move upwards into the quadrant that corresponds to analytical perceptual strategies–plus–good phonemic awareness.

HYPOTHESIS 2

The second possible relationship between perception and awareness is one in which phonemic awareness develops first, putting pressure on the perceptual system to shift to a strategy which favours phonemes, see Figure 3.6.

If this second hypothesis holds true, there should again be children with both good phonemic awareness and analytical cue weighting, and others with both poor awareness and global cue weighting. Additionally there should also be a third group of children who have begun to develop phonemic awareness, but whose perceptual strategies are not yet analytical. Crucially, there should *not* be any children with strongly analytical cue weighting who still have poor phonemic awareness.

Again, over the course of the longitudinal study, the data points should follow the direction of the arrows. In this case this means that the data points representing any children with poor phonemic awareness and global cue weighting should first move upwards, into the quadrant corresponding to good phonemic awareness–plus–global perceptual strategies. At a later session the data points should then move left, again into the quadrant corresponding to good phonemic awareness–plus–analytical perceptual strategies.

The critical step for each of these two hypotheses, therefore, is the direction of the first movement, which corresponds in each case to the development of the first of the two processes. The other critical point to notice is the area of the graph that is left empty in each situation: the upper right corner in the case where perceptual strategies change first, and the lower left corner in the case where phonemic awareness develops first.

It should be noted that there is a third possible relationship between acoustic cue weighting and phonemic awareness—a linear relationship. In this case one would expect to see only children with both good phonemic awareness and analytical cue weighting, and others with both poor awareness and global cue weighting, at all sessions in the study.

It should also be noted that small amounts of variation in performance can be expected to be seen for both the perceptual tests (Hazan & Rosen 1991) and also the phonemic awareness tests to a certain extent. Because of this, it is unlikely that the results of these tests will fit *exactly* into the four pre–designed quadrants. However, both the general direction of movement, and the location of the empty part of the graph should be apparent from the results, and should give a good indication of which of the two processes develops first.

4 Statistical analyses

In addition to the qualitative analyses permitted by the use of the above graphs, quantitative analysis of the results of these two studies will also be undertaken. Pearson's correlations between cue weighting and phonemic awareness will be measured at each session in Experiment 1 and in Experiment 2. Pearson's correlations will also be assessed between the three measures of phonemic awareness, as well as between the two measures of cue weighting (slope and separation of response curves). ANOVAs will be used to determine the degree of difference in perceptual strategy between the adults, and the children at all stages of Experiment 1. ANOVAs will also be used to determine the degree of difference in the children's performance in both perceptual strategy and phonemic awareness between each stage of Experiment 1 (i.e. between Stage 1 and 2, Stage 2 and 3, Stage 1 and 3), and between each of these stages and the study in Experiment 2. Finally, multiple regression analysis will be used to determine the extent to which variability in a process at one stage can be accounted for by processes at earlier stages.

Chapter 4

Experiment 1

This first experiment was a longitudinal study of changes in acoustic cue weighting, and the development of phonemic awareness in a group of beginning–reading children.

1 Subjects

Eighteen children participated in this study: 8 female and 10 male. An additional 9 children (4 female, 5 male) were also tested, but were not included in any analyses because they failed to meet the perceptual testing criteria (outlined below, see also Section 4 for a discussion of the excluded subjects). All 18 children were tested at Sessions 1 and 2, but three children (1 female, 2 male) dropped out before Session 3, therefore only 15 children were tested at this session.

The children were selected from schools in the Edinburgh area: 12 were from 2 different state primary (non–fee–paying) schools, and 6 were from 2 different classes in a private (fee–paying) school. At Session 1 the children ranged in age from 5;2 to 6;0, with an average age of 5;8. All of the children were native Scottish English speakers. Six of the 18 spoke a second language in addition to Scottish English, to differing degrees of bilingualism (as reported by parents)[1]. No significant differences in performance were found between the bilingual and monolingual children for any of the tests carried out in this study. The results of these 6 children were therefore analysed together with the results of the other 12.

[1]It should be noted that there are some studies which have found that bilingualism has an effect on certain metalinguistic abilities (e.g. Campbell & Sais 1995), however none have shown this effect to extend to the level of phonemic awareness development.

None of the children had a history of chronic otitis media, as this has been shown to delay or interfere with the development of perceptual weighting (see e.g. Nittrouer 1996*b*). For the purposes of this study, chronic otitis media is defined by Nittrouer (1996*b*) as more than 3 ear infections in the first three years of life and/or the implantation of myringotomy tubes. In addition none of the children or their siblings had ever received therapy for expressive language disorders. The reason for avoiding language disordered children was two–fold: first, children with such disorders (phonological disorders in particular) have been shown to have delayed shifts in perceptual weighting (Watson 1997), and second, as noted in Chapter 3, it has been shown that certain speech therapy programmes encourage the development of phonemic awareness skills. The above two criteria were determined by means of parental questionnaires. All of the children had been tested by the school authorities for hearing problems, and all had hearing within a normal range.

Additionally, in order to establish perceptual 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 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 had ever suffered from chronic otitis media, and none had ever received therapy for expressive language disorders.

2 Tests

Altogether, four tests were carried out in this study. These tests were designed to assess: i) acoustic cue weighting, ii) phonemic awareness, iii) general language ability, and iv) reading ability. The language test administered was the Short Form of the standardised British Picture Vocabulary Scale (BPVS; Dunn, Dunn, Whetton & Pintilie 1982). This test, which requires children to point to the correct picture in response to a spoken word, is a test of receptive vocabulary, however, it has been used in a number of studies (e.g. Brady et al. 1983, Nittrouer 1996*b*) as a measure of general language or verbal cognitive ability. The test administered to assess reading ability was the Schonell Graded Word Reading Test (also standardised, Schonell & Goodacre 1971). The materials and testing procedures for both of these tests followed the guidelines for each. Both tests were presented by the author (a non–Scottish English speaker).

The cue weighting and phonemic awareness tests were non–standardised tests designed specifically for this study. The materials and testing procedures for both the cue weighting and the phonemic awareness tests will be outlined in more detail below. Briefly, however, the stimuli for the acoustic cue weighting test made use of a /ʃo/–/so/ ('show–sew') contrast. As in Nittrouer and colleagues' previous studies (Nittrouer & Studdert-Kennedy 1987, Nittrouer 1992, Nittrouer 1996*b*), the two acoustic cues that were manipulated in these stimuli were i) the frequency of fricative noise, and ii) the configuration of the onset transitions of the vowel formants. The fricative noises were designed to vary along a continuum, from a frequency which strongly cued an /ʃ/, to one which strongly cued an /s/. The vowel–onset formant transitions were designed to vary in one of only two configurations: one which strongly cued an /s/ and the other which strongly cued an /ʃ/. Each of the two types of vowel configuration was combined with each of the points on the fricative continuum, resulting in two /ʃo/–/so/ continua, both with the same fricative noises and vowel targets, but one continuum with vowel–initial formant transitions appropriate for having followed an /s/ and one continuum with transitions appropriate for having followed an /ʃ/.

Phonemic awareness was tested by means of three tasks: phoneme blending, phoneme segmentation, and phoneme deletion. The presentation procedures for the three tests were modeled on similar tests (outlined by Hatcher 1994), with appropriate modifications for the specific testing of phonemic awareness, where necessary.

2.1 Cue weighting: materials

The design of the stimuli for the perceptual tests followed that used by Nittrouer (e.g. Nittrouer 1992, Nittrouer 1996*b*), with the modifications discussed in detail in Chapter 3. The most important of these modifications was the use of 'copy–synthesised' rather than hybrid, or stylised synthetic speech.

In both her 1992 and 1996 studies, the vowels in Nittrouer's hybrid fricative–vowel stimuli were excised from 5 *different* utterances of the same syllable. This was done to ensure "that perceptual responses would not be idiosyncratic to any one" vowel utterance (Nittrouer 1992, p. 357). This strategy was adopted for the current study, and adapted to the construction of the synthetic vowels (the synthesis method is described in more detail below). Additionally, in all of her cue weighting studies, Nittrouer presented the perceptual stimuli in isolation (i.e. without a carrier phrase). Again, this method was followed for this study.

The creation of the stimuli for the cue weighting test was undertaken in two parts. The first stage involved the recording of natural tokens of the target words /so/ and /ʃo/ by an adult male speaker of Standard Scottish English, and the subsequent acoustic analysis of these tokens. The second stage involved the copy–synthesis of the two continua of test stimuli: that is, the synthesis of the stimuli using the frequency and durational characteristics of the natural stimuli.

Acoustic analysis

As noted above, the acoustic characteristics of 5 different utterances of the vowels in the words /so/ and /ʃo/ were used to model 5 slightly different stimuli per fricative context. It should be noted that, although the purpose of using 5 different vowel portions was to introduce a small amount of natural variation into the stimuli, it also introduces the possibility that one of the utterances might be widely different in acoustic characteristics, and thus might engender skewed perceptual responses. In order to minimise large variations in articulation of the natural tokens, therefore, the speaker recorded the target words /so/ and /ʃo/ in the carrier phrase 'It's a ___ Bob.' 10 repetitions of each word were recorded, in random order. These recordings were used to determine the precise frequency characteristics of the synthetic stimuli.

Because the utterances were to be presented in isolation, a set of the target words spoken in isolation were also recorded. Four repetitions of each word, in random order, were recorded. The isolated target words in these utterances were found to be longer in duration, on average, than those that had been spoken in the carrier phrase. These recordings were therefore used to determine the durational characteristics of the synthetic stimuli, in particular the rate of frequency change of the vowel formant transitions.

The natural tokens were recorded onto DAT (Sony DTC–60ES) via microphone (Sony ECM–77B) and amplifier (Alice PAK2), in a sound attenuated recording studio. The recordings were then transferred to a computer for analysis: the speech was downsampled to 16 kHz at this point. All acoustic analysis was carried out using Entropic's ***waves+*** software, running under Unix.

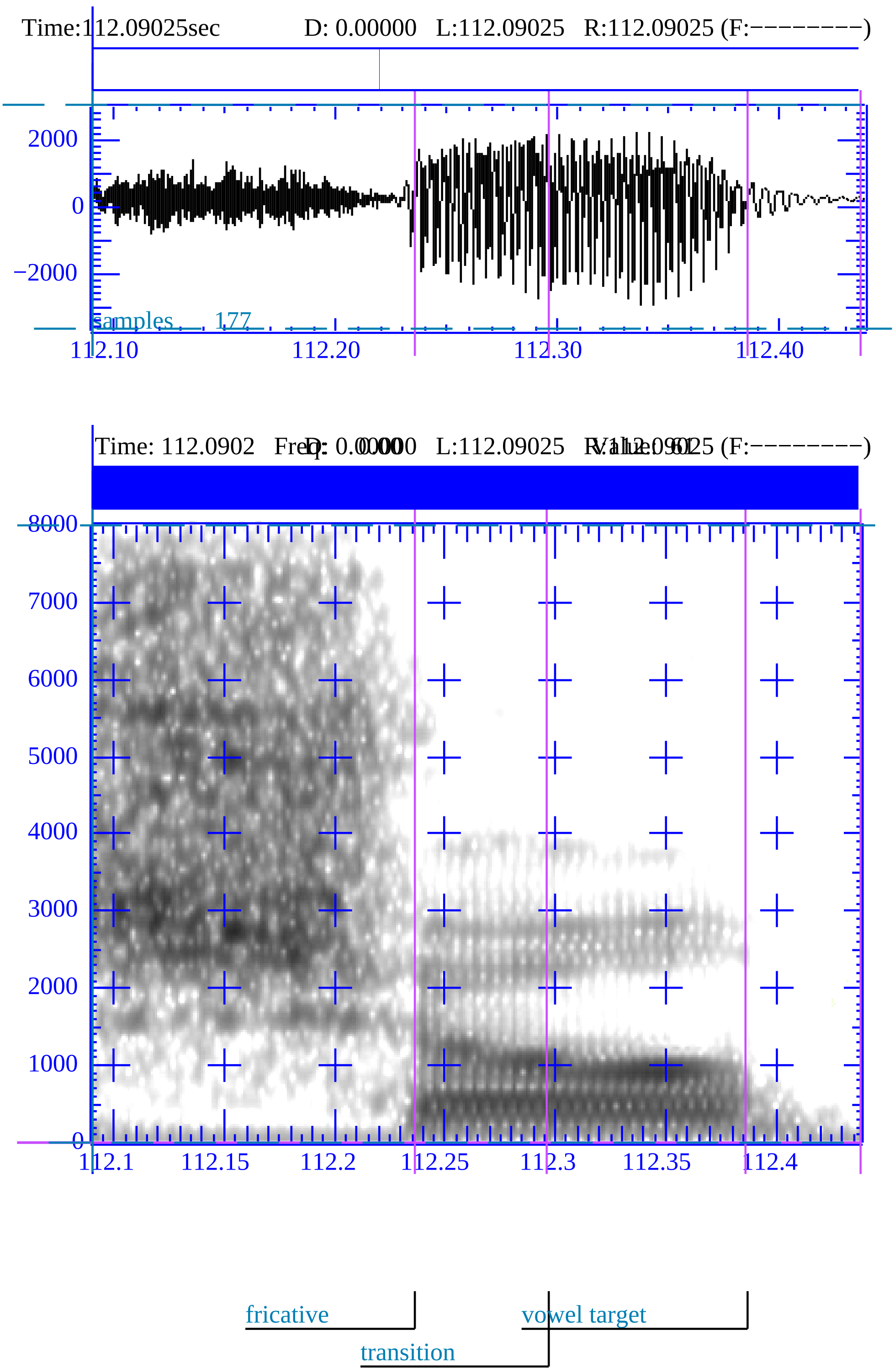

Figure 4.1: **Spectrogram of natural /ʃo/ token.** Duration measurements of the fricative noise, the transition, and the vowel target were made to the points indicated.

Duration (in ms)	Whole word	Fricative	Whole vowel	Transition	Vowel target
Isolation	423	173	250	75	175
Context	321	164	157	60	97

Table 4.1: **Average durations of sections of natural tokens**

All tokens of /so/ and /ʃo/ (both carrier–phrase tokens and isolated tokens), were analysed for: i) overall length of token, ii) length of fricative noise, iii) length of entire vowel (i.e. the length of the vowel from the onset to the offset of voicing), iv) length of the transition only, and v) length of the vowel target (or 'steady–state' portion of the vowel). Transitions were measured from the onset of voicing to the point at which the formant movement levelled off: this section of the vowel was equivalent to the steepest portion of formant change (see Whalen 1981), and was fairly consistently 60ms from the start of voicing (see Figure 4.1). The average values for all of the above length measurements are shown in the Table 4.1. Note that these values are the averages of (i) all 8 isolated tokens and (ii) 10 of the 20 carrier phrase tokens, chosen for their similarity, (the details of this choice will be discussed in more detail below).

Both sets of tokens were also measured for the *overall* pitch (F0) pattern of the entire utterance: that is, the entire phrase, including the target word, in the case of the carrier–phrase tokens, and the target word only, in the case of the isolated tokens. Three measurements were taken—one each at the beginning, middle, and end of voicing in each utterance. The speaker's pitch patterns were found to be highly consistent, over both whole phrases and isolated target words.

The carrier–phrase tokens were analysed for the frequency characteristics of both the fricatives and the vowels. Spectra of each natural fricative noise were examined to determine the frequency of the point of highest amplitude for each. Measurements were also taken of the vowel formant frequencies: the first, second, and third formants (F1, F2, and F3) were measured at 20ms intervals from the beginning of voicing to the end of the vowel. This resulted in 8 frequency sample points per vowel on average, 4 sample points in the transition section, and 4 sample points in the vowel target section.

Synthesis

The synthetic test stimuli were created using *SenSyn*, Sensimetrics' cascade/parallel formant synthesiser (based on Klatt 1980). As noted above, the frequency and durational values used in the creation of the synthetic stimuli were copied from natural tokens of the target words. The frequency values used in the stimuli were actual values taken from 5 of the carrier–phrase natural tokens, while the durational values used to model the stimuli were based on average measurements from all of the isolated natural tokens.

GENERAL The overall duration of each synthetic stimulus was 480ms, with 230ms of fricative noise and 250ms of vowel. It will be noted that the duration of the fricative noise (and thus of the entire stimulus) is slightly longer (47ms) than the average duration of the fricatives in the isolated natural tokens. The length of the fricative noise was taken from Nittrouer (1992), who made the decision to use longer fricative noises in an effort to rule out possible alternative explanations for the differences in cue weighting that she had observed between adults and children. One such alternative was that children's cue weighting strategies were simply the result of a difficulty in making use of fricative cues. Nittrouer's argument was that if this were the case, then the use of longer fricative noises should give children a better opportunity to listen to, and make use of such cues. Although Nittrouer (1992) found that children displayed the same cue weighting strategies for longer as for shorter fricative noises (thus refuting the theory that children have perceptual difficulties with fricative noises), it was decided to maintain longer fricative noises for the current study, in order to fully test the new /ʃo/–/so/ contrast. The duration of the vowel was based on the average length of the isolated natural tokens.

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

VOWELS Ten of the 20 natural carrier–phrase tokens (5 each of /so/ and /ʃo/) were chosen as the models for the frequency characteristics of the synthetic vowels. 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—this ensured that the fricative induced differences in formant frequency at the onset of the vowel did not extend into the vowel target itself.

Recall that, of the 8 frequency measurements taken of the carrier–phrase natural tokens, the first 4 samples spanned the transitional section of the vowel, while the last 4 were all from within the vowel target section. To synthesise vowels with the appropriate duration, the 20ms interval between the 4 frequency measurements taken within the transition was multiplied by 1.25 (the difference between the duration of the transitional section in the carrier–phrase tokens, and the duration of the same section in the isolated tokens). Thus sample 1 was placed at 0ms, sample 2 at 25ms, sample 3 at 50ms, and sample 4 at 75ms. The frequency values were interpolated between these points. The 25ms inter–sample interval was maintained for the first two sample points of the vowel target section (samples 5 and 6), in order to ensure that any tailing off of the transitional movement at the end of the main transition was synthesised at the same rate of change as the main transitional movement. Finally, sample 7 and sample 8 were placed at 185ms and 250ms into the vowel, respectively. In actual fact, the relative placement of the 4 vowel target samples (samples 5–8) was fairly inconsequential, as for the most part the formant frequency values of the samples changed very little, if at all, over the course of these four samples.

The amplitude of voicing was 60 dB from the beginning of the vowel for 185ms, falling to 0 dB from 185ms to 250ms (see Klatt 1980).

COMBINATION Each of the ten synthetic vowels was combined with each of the nine fricative noises, resulting in 90 different stimuli altogether. Because the speaker produced similar intonation patterns for whole phrases and for isolated words, the intonation of the synthetic stimuli was modeled on an average pitch pattern of both. F0 for each token, therefore, began at 160 Hz at 230ms, rose to 180 Hz at 355ms and then fell to 100 Hz at 480ms. Figure 4.2 shows spectrograms of four of the 90 test stimuli.

2.2 *Cue weighting: presentation*

TRAINING The training for the perceptual test involved a number of steps. The first was to introduce the children to the target words /ʃo/ and /so/, and to the pictures that would be used to represent these words: a small boy **show**–ing

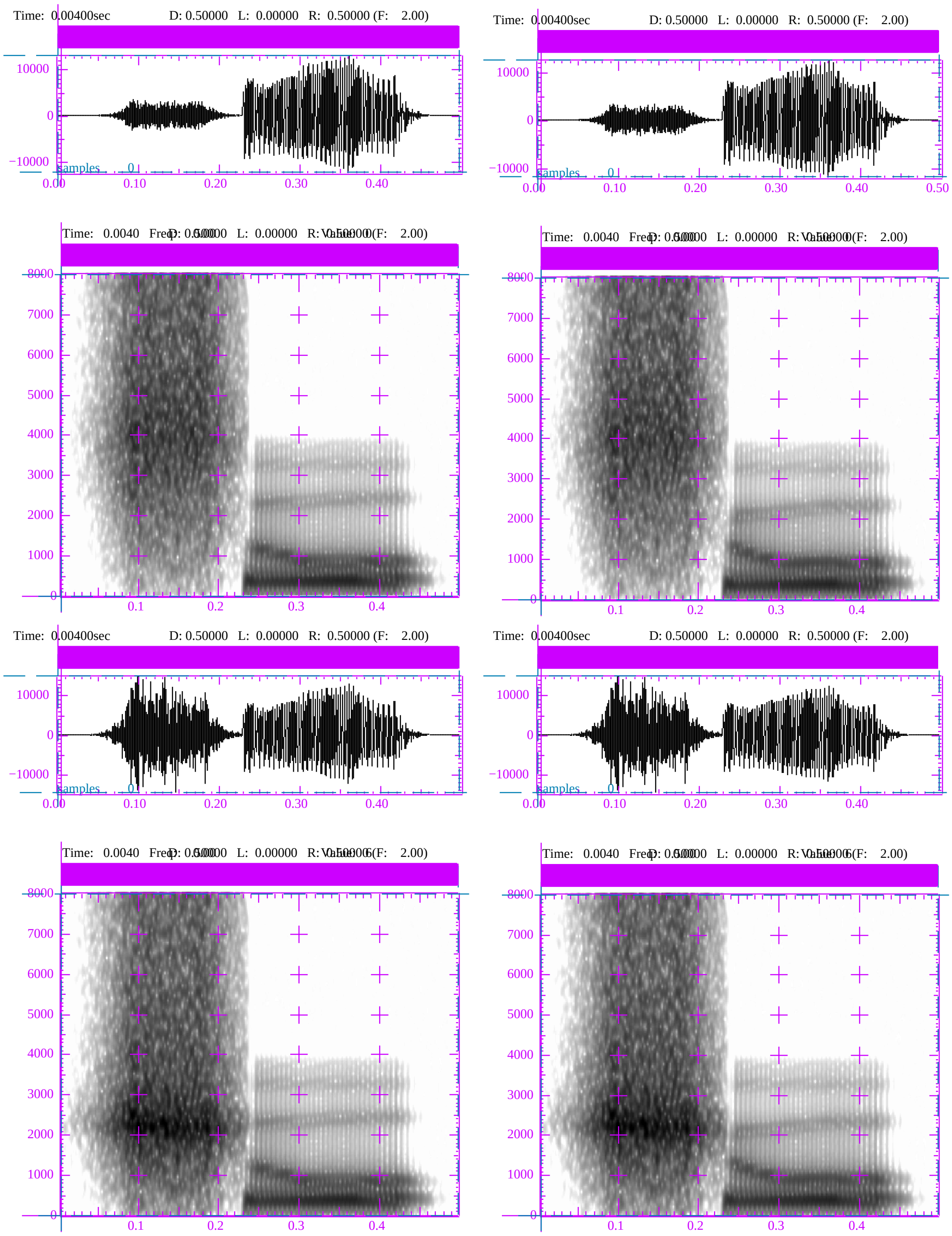

Figure 4.2: **Spectrograms of four example test stimuli.** These spectrograms show four 'endpoint' stimuli: that is, the two most extreme fricative noises, combined with each of the two transitional contexts. The top two spectrograms represent stimuli with the most /s/–like fricative noise, while the bottom two represent stimuli with the most /ʃ/–like fricative noise. The left two spectrograms represent stimuli with /s/–transitions, while the right two spectrograms represent stimuli with /ʃ/–transitions.

someone a hole in a teddy bear, and someone **sew**–ing the hole in the teddy bear. The children were familiarised with both the words and the pictures by means of a short story, which had been recorded by an adult male speaker of Standard Scottish English, and an accompanying picture book (see Appendix A). In order to introduce the children to computer generated speech, a synthesised version of the story was also presented.

The second stage of the training was to introduce the children to the principle of the test. The children were shown the pictures representing the two target words /ʃo/ and /so/. They were then told that they would hear one of the two words corresponding with the pictures, and were instructed to indicate which of the words they had heard by placing a counter on the appropriate picture. The children were then trained on natural tokens of the two words, presented (unrecorded) by the experimenter. Training continued in this way until it was clear that each child understood the task.

The final stage of the training involved a pre–test, which made use of the actual synthetic stimuli. This test served two purposes: first, to ensure that the children understood the testing procedure, and second, to ensure that the children were capable of identifying the synthetic stimuli. The stimuli used in the pre–test were 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 /so/–like stimuli) and the 2.2kHz noise plus vowels with an /ʃ/ transition (the most /ʃo/–like stimuli). There were 10 stimuli in the pre–test (five vowels in each transition condition) which were presented in random order.

TEST The 90 test stimuli were randomised and split into 10 blocks for presentation. Each child was presented with the entire set of 90 stimuli twice at each session, resulting in 180 responses per child, and 10 responses per transition type for each point on the fricative continuum. The children heard the two sets of 90 stimuli on two different days. The 90 stimuli (and the 10 pre–test stimuli) were recorded for presentation in 5 different random orders. Each child heard the 90 stimuli in 2 different random orders at each testing day in the session, and never heard the stimuli in the same 2 random orders across sessions.

Because some children had a tendency to fidget or talk between the stimuli, this study followed the procedure used by Walley & Carrell (1983) and did not fix the interstimulus interval for presentation of the 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 ten, at which point the children were allowed to choose a small prize. Children heard the two sets of 90 stimuli on two different days.

ADULTS The training of the adults involved fewer stages than that of the children. The adults did not listen to either version of the story, and were not trained on unrecorded natural target words. The adults were told that they would hear repetitions of the two words /so/ and /ʃo/ in random order, and were instructed to indicate which they had heard by placing a tick in a box on a form provided. The adults were given the same pre–test as the children, and had to meet the same criterion (see below). Once the adults had completed the pre–test, and it had been determined that they understood the task, they went on to label both sets of 90 stimuli in one sitting. The interstimulus interval for adults was fixed at 3 seconds, and the inter–block interval 5 seconds.

CRITERIA All listeners were required to correctly identify 9 of the 10 pre–test stimuli in order for their results to be included in analyses. For the test proper, all listeners were required to respond correctly to at least 8 of the 10 continuum endpoints (i.e. those stimuli presented in the pre–test). Additionally, if any of the subjects displayed non–S–shaped response curves (i.e. if the subject's two response curves crossed each other more than twice) at either the pre–test or the test stage, that subject's data was not included in analysis. These criteria were used in an attempt to eliminate any listener who was not completely attentive to the task.

2.3 Phoneme awareness: materials

All of the words used in the phonemic awareness tests were selected from the CHILDES database (MacWhinney 1995). All words selected had occurred 5 or more times in the database. To the extent possible within the constraints of the database, the test items were chosen to be balanced within tests for the type of phoneme in a word (i.e. stops, fricatives, nasals, vowels) and the placement of these phonemes within a word (i.e word–initial, –medial, or –final), as these are factors which have been found to affect ability on phoneme awareness tasks (McBride-Chang 1995*b*). The exceptions to this were post–vocalic /r/ and word final /t/, /d/, /s/ and /z/. Post vocalic /r/ was not used at all in the phoneme blending or phoneme segmentation tests. This was done because, although Standard Scottish English is a rhotic dialect of English, some Scottish English speakers do not actually use post vocalic /r/. Additionally, certain words with a word–final /t/, /d/, /s/, or /z/ were not chosen: the criterion for the exclusion of these words was if there was any way in which a child who was unfamiliar with the word could construe these phonemes as past tense or plural morphemes (e.g. /kard/ could be the past tense of the nonsense verb /kar/). Again, this precaution was taken because phoneme/morpheme confounds have also been shown to affect phonemic awareness ability (McBride-Chang 1995*b*). One final factor which has been shown to affect phonemic awareness—specifically, number of phonemes in a word (McBride-Chang 1995*b*)—was manipulated deliberately in these tests in order to maintain a reasonably high level of difficulty throughout the longitudinal testing period. The test and pre–test materials for all three phonemic awareness tests can be found in Appendix B.

Phoneme blending and phoneme segmentation

Both the phoneme blending and phoneme segmentation tests had 50 test items each: 20 three–phoneme words, 20 four–phoneme words and 10 five–phoneme words. Because there is a possibility that manipulation of word–initial consonant clusters and word–final consonant clusters may require different levels of cognitive or metalinguistic ability (McBride-Chang 1995*b*), the four–phoneme words in these tests were balanced for the position of clusters—either CCVC, or CVCC. The five–phoneme words were predominantly CCVCC, with only 3 words in the blending test and 4 four words in the segmentation test that had three–consonant–clusters: CCCVC.

Phoneme deletion

The phoneme deletion test had 40 three– and four–phoneme items, all with two–consonant–clusters, and all with that consonant cluster in word–initial position: CCV(C). The words used for this test were chosen based on the fact that each item remained a valid word after its initial phoneme had been deleted—e.g. ‘snap’ becomes ‘nap’. A number of /s/–initial words had two valid responses, depending on whether a phonemic or a phonetic approach was taken in the segmentation of the initial consonant—e.g. ‘spot’ can become /pɔt/ or /bɔt/. Only /s/–initial words in which both possible responses were valid words were selected for this test.

2.4 *Phonemic awareness: presentation*

The stimuli for all phoneme awareness tests (the training, pre–test and test stimuli) were recorded for presentation to the subjects by an adult male speaker of Standard Scottish English. The speaker was instructed to produce all words clearly, and all individual phonemes without any following vowel, i.e. /s/ rather than /sə/.

Each test was introduced and explained to the children by the experimenter (a non–Scottish English speaker). Testing involved the experimenter asking the children to listen to the recorded stimuli, and to perform the required manipulation (e.g. the experimenter said "Can you break this word up into little bits?" followed by the recorded voice saying "pig.").

Phoneme blending

TRAINING At the first stage of training for the phoneme blending test, the children were introduced to the concept of the task. The children were shown a puppet, and were told that the puppet spoke 'in a funny way,': it said every word 'all broken up into little bits.' The children were then instructed to listen to the puppet, and to guess what they thought it had said. The children were then presented with a set of training words. The first words in this set were two–syllable compound words, which had been segmented at the syllable level: e.g. 'cowboy' was presented as /kaʊ–bɔɪ/. The next words in the set were one syllable words that had been segmented at the onset–rime level: e.g. 'cow' was presented as /k–aʊ/. The final words in the set were three–phoneme words that had been segmented at the phonemic level: e.g. 'pig' was presented as /p–ɪ–g/. The children received corrective feedback throughout this stage of the training.

The next stage of the training was a pre–test, for which no corrective feedback was given. The five items in this test were one syllable words that had been segmented at the onset–rime level. As described in more detail in Chapter 2 (Section 2.1), syllable and onset–rime awareness have been repeatedly found to precede the development of phoneme awareness. It was therefore assumed that children who were unable to manipulate words at this level would not be able to successfully complete the phoneme manipulation in the test proper. If any child was unable to successfully manipulate all five pre–test items, the test was discontinued. The child was not, however, eliminated from further testing or analysis.

TEST The test made use of the puppet introduced in the training session. The children were presented with words segmented at the phoneme level and were instructed (using the same simple terminology as used in the training) to blend the phonemes together into words. No corrective feedback was given. The stimuli for the phoneme blending test were split into 2 balanced sets, each containing 10 three–phoneme words, 10 four–phoneme words and 5–five phoneme words. The stimuli were presented to the children in order of phoneme length (all three–phoneme words, followed by all four–phoneme words, followed by all five–phoneme words). It has been shown that the higher the number of phonemes in a word, the more difficult it becomes for children to successfully manipulate the phonemes (McBride-Chang 1995*b*). Presenting the words in this order, therefore, means that the test should get progressively more difficult. As noted above, one reason for doing this was to maintain a reasonably high level of difficulty for at least some test items throughout the longitudinal study. Additionally, it was assumed that if any child was unsuccessful at a particular stage of the test, they would be unlikely to be able to successfully manipulate later test items. Therefore, if any child was unable to correctly manipulate 5 out of 6 consecutive stimuli, the test was discontinued. The child was not eliminated from further testing or analysis. The children heard the two sets of stimuli on two different days.

Phoneme segmentation

TRAINING For the training for the phoneme segmentation test, the children were given the puppet to which they had been introduced in the phoneme blending test (the phoneme segmentation test always followed the phoneme blending test). The children were then told that they would hear a number of words, and were instructed to try to say the words as the puppet had said them in the previous task: i.e. 'all broken up into little bits' (note that none of the words in the segmentation test were the same as the words in the blending test, although some of the training words were the same). The children were then presented with a set of training words. The first words in this set were two–syllable compound words which the children were encouraged to segment at the syllable level: e.g. 'snowman' becomes /sno–mæn/. The next words were one–syllable CV words which the children were encouraged to segment into onset–rime units: e.g. 'cow' becomes /k–aʊ/. The final words in the set were three–phoneme words, which the children were encouraged to segment at the phoneme level: e.g. 'man' becomes /m–æ–n/. The children received corrective feedback throughout this stage of the training. If a child segmented a word into units which were larger than the target units (e.g. segmented 'man' into onset–rime units instead of phonemes), they were encouraged to 'try to break the word up into even smaller bits.'

The next stage of the training was a pre–test, during which no corrective feedback was given. The five items in this test were one syllable CV words that the children were encouraged to segment into onset–rime units. Again it was assumed that children who were unable to manipulate words at this level would not be able to successfully complete the phoneme manipulation in the test proper. If any child was unable to successfully manipulate all five pre–test items, the test was discontinued. The child was not eliminated from further testing or analysis.

TEST The test made use of the puppet introduced in the training session. The children were presented with words and were instructed (using the same simple terminology as used in the training) to segment the words into phonemes. No corrective feedback was given. Only phoneme sounds (e.g. /k–æ–t/ for 'cat') were accepted as answers. If responses were given as letter names the child was encouraged to respond again, using sounds only. The stimuli for the phoneme segmentation test were split into 2 balanced sets, each containing 10 three–phoneme words, 10 four–phoneme words and 5–five phoneme words. As in the phoneme blending test, the stimuli were presented to the children in order of phoneme length (all three–phoneme words, followed by all four–phoneme words, followed by all five–phoneme words). Again, it was assumed that if any child was unsuccessful at a particular stage of the test, they would be unlikely to be able to successfully manipulate later test items. Therefore, if any child was unable to correctly manipulate 5 out of 6 consecutive stimuli, the test was discontinued. The child was not eliminated from further testing or analysis. The children heard the two sets of stimuli on two different days.

Phoneme deletion

TRAINING Before any training was carried out for this test, the experimenter first established that the children understood the concept of deletion, or 'taking away.' The children were presented with two different coloured counters and were asked to take one away, and say which one was left. Once it was clear that the children understood this concept, they were then trained to apply it to units of speech. The children were introduced to a cartoon drawing of a man with a sack over his shoulder, and were told that this was a thief who liked to steal the first sound from every word he heard. The children were then presented with a set of training words and were asked what would be left if the thief took the first sound in the word away. The sound to be deleted was identified for the child in all cases: e.g. 'what does 'cat' sound like without the /k/?' The set of training words began with single–consonant–initial words (CVC) in which the sound to be deleted was an onset. The next words began with two–consonant clusters: CCV(C), in which the sound to be deleted was just the initial consonant of the cluster. The children received corrective feedback throughout this stage of the training.

The next stage of the training was a pre–test, during which no corrective feedback was given. The five items in this test were CVC words, in which the sound to be deleted was an onset, and the remainder of the word was a rime. Again it was assumed that children who were unable to manipulate words at this level would not be able to successfully complete the phoneme manipulation in the test proper. If any child was unable to successfully manipulate all five pre–test items, the test was discontinued. The child was not eliminated from further testing or analysis.

TEST The test made use of the cartoon character introduced in the training session. The children were presented with CCV(C) words and were instructed (using the same simple terminology as used in the training) to delete the initial phonemes. As in the training, the sound to be deleted was identified in all cases. No corrective feedback was given. The stimuli for the phoneme deletion test were split into 2 sets, each containing 20 words. If any child was unable to correctly manipulate 5 out of 6 consecutive stimuli, the test was discontinued. The child was not eliminated from further testing or analysis. The children heard the two sets of stimuli on two different days.

3 Procedure

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.

The child subjects were tested three times over the course of 7 months, with testing taking place at months 1, 4, and 7. Acoustic cue weighting and phonemic awareness were tested at Sessions 1, 2, and 3 of the study. General language ability and reading skills were tested only at Sessions 1 and 3. The testing for the child subjects was spread out over two days, which were not more than one week apart. The order of testing was as follows:

- Day 1:
  1. BPVS (Sessions 1 and 3)
  2. 1st half of acoustic cue weighting
  3. 1st half of phonemic awareness
- Day 2
  1. 2nd half of acoustic cue weighting
  2. 2nd half of phonemic awareness
  3. Schonell Graded Word Reading (Sessions 1 and 3)

The adult subjects were tested on only one occasion, and only on their acoustic cue weighting strategies. Both halves of the acoustic cue weighting test were presented at the same sitting, with a short break in between.

4 Results

As noted above, 18 children met the perceptual testing criteria. The 9 additional children who were tested but excluded from analysis either did not meet the criteria for correct perceptual responses (9 out of 10 stimuli correctly identified for the pre–test; 8 out of 10 endpoint stimuli correctly identified for the test proper), or had response curves which did not sufficiently approximate S–shaped curves. There are a number of potential explanations for the perceptual behaviour of these 9 children. First, although great care was taken to ensure that all of the children understood the task, a number of children appeared to adopt response 'strategies' which suggested that their responses did not coincide with their percepts. One such strategy involved the regular alternation of responses back and forth between 'sew' and 'show,' while another involved the consistent favouring of one response to the exclusion of the other. Additionally, a number of children appeared to have difficulty maintaining constant attention to the task: these children often displayed delayed responses (and often only after prompting). The longer the delay, the more likely it would be that the child would have forgotten what they had perceived, and be responding simply by guessing. Finally, there is a possibility among all of these children that some may have had undiagnosed ear infections, which (as noted above) would impact on their cue weighting strategies (see e.g. Nittrouer 1996*b*).

At Session 1, the 18 children who met testing criteria had reading ages which ranged from 6;0 and below (0 words read correctly on the Schonell Graded Word Reading Test) to 8;6. At this session these 18 children had Age Equivalents based on BPVS scores which ranged from 4;0 (Confidence Interval 3;4–4;8) to 11;6 (Confidence Interval 10;4–12;9). At Session 3 the remaining 15 children had reading ages which ranged from 6;0 to 8;6. Age Equivalents at this session ranged from 6;3 (Confidence Interval 5;5–7;1) to 12;2 (Confidence Interval 11;0–13;7). As noted above, there were no significant differences between bilinguals and monolinguals for any of the processes tested. Additionally, there were no significant differences between the children attending state primary schools, and those attending fee–paying schools for any of the processes tested.

The cue weighting and phonemic awareness results of this study will first be presented by individual session, and will then be examined longitudinally. All of the statistical analyses were carried out using *SPSS* running under Unix. The raw data for all tests can be found in Appendix C.

4.1 Session 1

Acoustic cue weighting

The graphs in Figure 4.3 show the perceptual response curves for the 8 adults (A), and the 18 children at Session 1 of the study (B). It can be seen that the children's response curves are shallower and more widely separated than those of the adults.

ANOVAs with the perceptual measures of *slope* and *separation* as dependent variables, and *age* as the independent variable, show that there is 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 this session.

An examination of the two measures of acoustic cue weighting shows that there is no significant correlation between the slope and the separation of response curves for either the adults or the children.

Neither the slope nor the separation of the children's response curves correlates with either general language ability, or word reading ability.

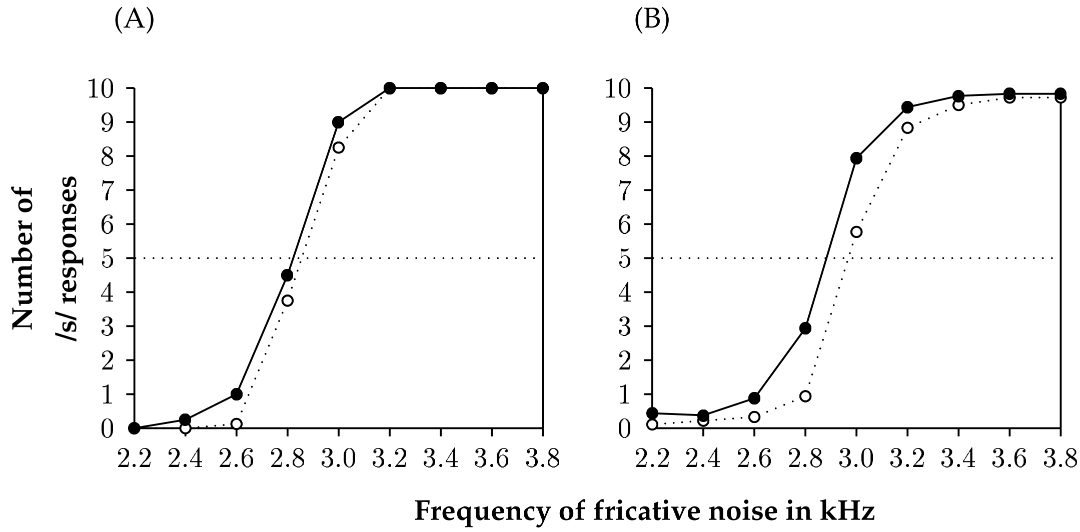

Figure 4.3: **Responses of adults (A) and children at Session 1 (B) to /ʃo/–/so/ continua.** The x-axis shows the continua of fricative noises, ranging in frequency from 2.2kHz (the most /ʃ/–like) to 3.8kHz (the most /s/–like). The solid line represents a listener's /s/ responses to stimuli with /s/–transitions; the dotted line represents the same listener's /s/ responses to stimuli with /ʃ/–transitions.

Phonemic awareness

The mean scores for the phonemic awareness tests at this session are: phoneme blending: 28/50, phoneme segmentation: 25/50, and phoneme deletion: 12/40.

All three measures of phonemic awareness correlate very highly with each other: phoneme blending and phoneme segmentation $[r = .8942, p < .001]$, phoneme blending and phoneme deletion $[r = .5149, p = .02]$, phoneme segmentation and phoneme deletion $[r = .6225, p = .006]$.

Word reading ability correlates highly with both phoneme blending $[r = .6176, p = .006]$ and phoneme segmentation $[r = .7244, p = .001]$, but not with phoneme deletion. General language ability does not correlate with any of the measures of phonemic awareness.

Correlation between cue weighting and phonemic awareness

The graph in Figure 4.4 shows the children's acoustic cue weighting in terms of separation of response curves, and phonemic awareness in terms of phoneme blending at Session 1. The graph is divided into quadrants at the median

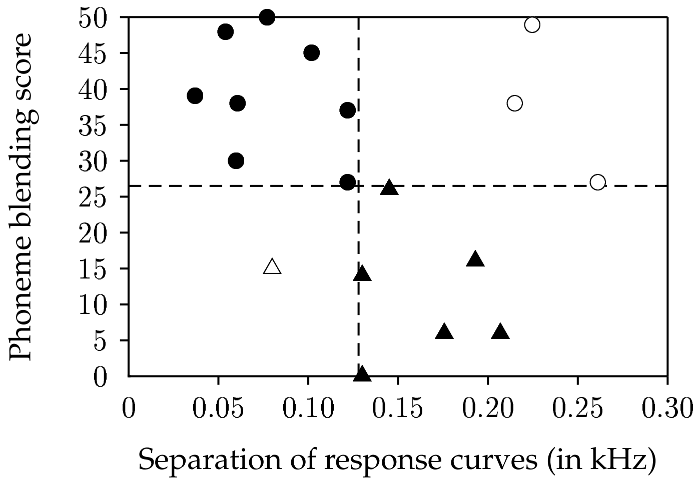

Figure 4.4: **Relationship between perception and awareness: Session 1.** The graph is divided into quadrants at the median phonemic awareness score (on the y-axis) and the most global of the adults' cue weighting responses (on the x-axis). Each point on the graph represents one subject.

phoneme blending score: 26.5/50, and the most global of the adults' cue weighting responses: 0.13kHz separation between the response curves. It can be seen that there are children with both good phonemic awareness and analytical cue weighting strategies, and children with poor phonemic awareness and global cue weighting strategies. It can also be seen that there are a number of children with very good phonemic awareness who have very global cue weighting strategies, but few children who have poor phonemic awareness and very analytical cue weighting strategies.

At this session, separation of response curves correlates with phoneme deletion ability [$r = .6277, p = .005$], but does not correlate with any other measure of phonemic awareness. Slope of response curves does not correlate with any of the three measures of phonemic awareness ability.

4.2 *Session 2*

Acoustic cue weighting

The graphs in Figure 4.5 show the perceptual response curves for the 8 adults (A), and the 18 children at Session 2 of the study (B). Again it can be seen that the children's response curves are still slightly shallower and slightly more widely separated than those of the adults.

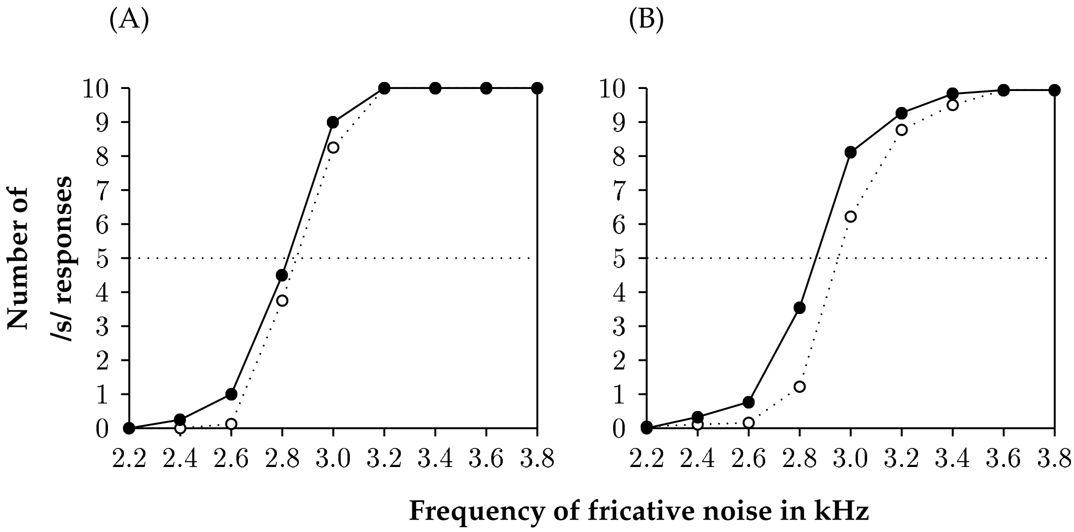

Figure 4.5: **Responses of adults (A) and children at Session 2 (B) to /ʃo/–/so/ continua.** The x-axis shows the continua of fricative noises, ranging in frequency from 2.2kHz (the most /ʃ/–like) to 3.8kHz (the most /s/–like). The solid line represents a listener's /s/ responses to stimuli with /s/–transitions; the dotted line represents the same listener's /s/ responses to stimuli with /ʃ/–transitions.

ANOVAs with the perceptual measures of *slope* and *separation* as dependent variables, and *age* as the independent variable, show that there is again a significant difference in both slope $[F(1,24) = 5.92, p = .02]$ and separation $[F(1,24) = 3.95, p = .05]$ between the adults and the children. This difference appears to be smaller for the separation measure than it was at Session 1 of the study.

An examination of the two measures of acoustic cue weighting shows that there is a significant correlation between the slope and the separation of response curves for the children at this session $[r = .5166, p = .02]$.

General language ability and word reading ability were not tested at this session.

Phonemic awareness

The mean scores for the phonemic awareness tests at this session are: phoneme blending: 39/50, phoneme segmentation: 36/50, and phoneme deletion: 24/40.

Again all three measures of phonemic awareness correlate very highly with each other: phoneme blending and phoneme segmentation $[r = .7898, p < .001]$, phoneme blending and phoneme deletion $[r = .4740, p = .04]$, phoneme segmentation and phoneme deletion $[r = .6677, p = .002]$.

Correlation between cue weighting and phonemic awareness

The graph in Figure 4.6 shows the children's acoustic cue weighting in terms of separation of response curves, and phonemic awareness in terms of phoneme blending at Session 2. The graph is divided into quadrants at the median phoneme blending score from Session 1: 26.6/50, and the most global of the adults' cue weighting responses: 0.13kHz separation between response curves. Again, it can be seen that there are children with both good phonemic awareness and analytical cue weighting strategies, and children with poor phonemic awareness and global cue weighting strategies. It can also be seen that there are children with good phonemic awareness who have global cue weighting strategies, but no children who have poor phonemic awareness and analytical cue weighting strategies.

At this second session, separation of response curves correlates with phoneme blending ability [$r = .7108, p = .001$], and phoneme segmentation ability [$r = .6989, p = .001$] but does not correlate with phoneme deletion. Slope of response curves does not correlate with any of the three measures of phonemic awareness ability.

General language ability and reading ability were not tested at this session.

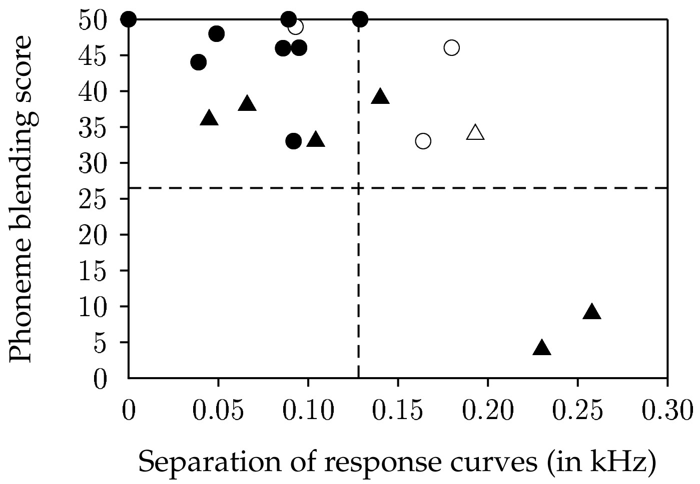

Figure 4.6: **Relationship between perception and awareness: Session 2.** The graph is divided into quadrants at the median phonemic awareness score (on the y-axis) and the most global of the adults' cue weighting responses (on the x-axis). Each point on the graph represents one subject.

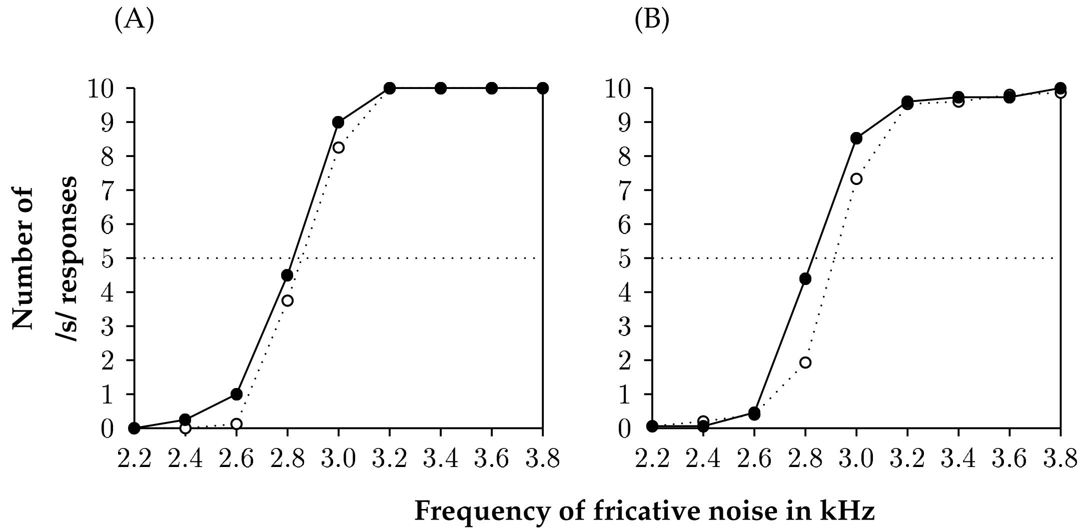

Figure 4.7: **Responses of adults (A) and children at Session 3 (B) to /ʃo/–/so/ continua.** The x-axis shows the continua of fricative noises, ranging in frequency from 2.2kHz (the most /ʃ/–like) to 3.8kHz (the most /s/–like). The solid line represents a listener's /s/ responses to stimuli with /s/–transitions; the dotted line represents the same listener's /s/ responses to stimuli with /ʃ/–transitions.

4.3 Session 3

Acoustic cue weighting

The graphs in Figure 4.7 show the perceptual response curves for the 8 adults (A), and the 15 children who remained at Session 3 of the study (B). At this session there is much less difference between the response curves of the children, and those of the adults.

ANOVAs with the perceptual measures of *slope* and *separation* as dependent variables, and *age* as the independent variable, show that there is no longer any significant difference in slope or separation of response curves between the children at this session and the adults.

An examination of the two measures of acoustic cue weighting shows again that there is no significant correlation between the slope and the separation of response curves for either the adults or the children.

At this session, word reading ability correlates with the separation of the children's response curves [$r = .5941, p = .02$], but not with the slope of the response

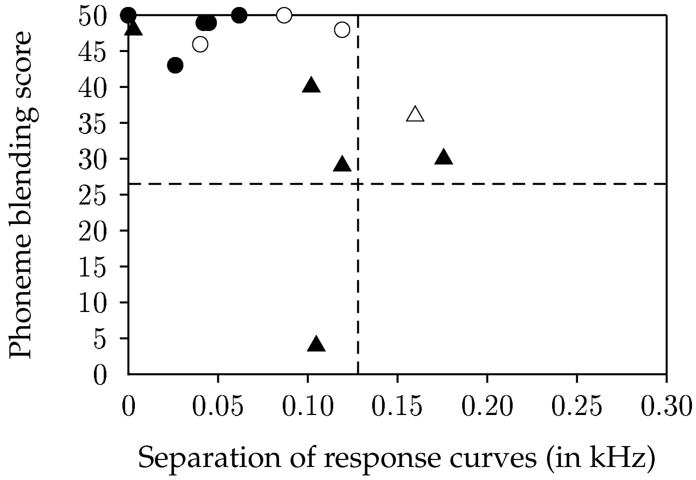

Figure 4.8: **Relationship between perception and awareness: Session 3.** The graph is divided into quadrants at the median phonemic awareness score (on the y-axis) and the most global of the adults' cue weighting responses (on the x-axis). Each point on the graph represents one subject.

curves. Once again, neither the slope nor the separation of the children's response curves correlates with general language ability.

Phonemic awareness

The mean scores for the phonemic awareness tests at this session are: phoneme blending: 41/50, phoneme segmentation: 42/50, and phoneme deletion: 31/40.

All three measures of phonemic awareness correlate very highly with each other again at this session: phoneme blending and phoneme segmentation $[r = .8690, p < .001]$, phoneme blending and phoneme deletion $[r = .7291, p = .002]$, phoneme segmentation and phoneme deletion $[r = .7254, p = .002]$.

At this session word reading ability correlates highly with both phoneme blending $[r = .5824, p = .02]$ and phoneme deletion $[r = .5114, p = .05]$, but not with phoneme segmentation. Once again, general language ability does not correlate with any of the measures of phonemic awareness.

Correlation between cue weighting and phonemic awareness

The graph in Figure 4.8 shows the children's acoustic cue weighting in terms of separation of response curves, and phonemic awareness in terms of phoneme blending at Session 3 of the study. The graph is divided into quadrants at the median phoneme blending score from Session 1 of the study: 26.6/50, and the most global of the adults' cue weighting responses: 0.13kHz separation between response curves. Once again, it can be seen that there are children with both good phonemic awareness and analytical cue weighting strategies, and a few children still with poor phonemic awareness and global cue weighting strategies. It can also be seen that there are children with good phonemic awareness who have global cue weighting strategies, but only one child who displays poor phonemic awareness and slightly analytical cue weighting strategies.

At this session separation of response curves correlates with phoneme blending ability $[r = -.5510, p = .03]$, but does not correlate with any other measure of phonemic awareness. Slope of response curves does not correlate with any of the three measures of phonemic awareness ability.

4.4 Longitudinal

Acoustic cue weighting

ANOVAs with the perceptual measures of *slope* and *separation* as dependent variables, and *session in the study* as the independent variable, show that there is a significant change in cue weighting strategy across all three sessions in the study. Interestingly, while there is a significant difference in separation of response curves between Session 1 and Session 3 $[F(2, 48) = 3.24, p = .04]$, there is no significant difference in slope of response curves between Session 1 and Session 3. Examining the change in perceptual strategy session by session, it was found that there is no significant difference in either slope or separation of response curves between Session 1 and Session 2, or between Session 2 and Session 3, although the difference in separation of response curves approaches significance between Session 2 and Session 3 $[F(1, 31) = 3.47, p = .07]$.

ANOVAs with *separation* as the dependent variable, and the *day of testing*, and the *session in the study* 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 across the sessions in the study, and not by any day–to–day variation in behaviour.

An examination of the two measures of acoustic cue weighting across all sessions in the study shows that there is a certain amount of correlation between the same measure over time, but no correlations at all between the two different measures over time. Slope of response curves at Session 1 correlates significantly with slope at Session 2 [$r = .5379, p = .02$], and approaches a significant correlation with slope at Session 3 [$r = .4772, p = .07$]. Slope at Session 2 correlates significantly with slope at Session 3 [$r = .5807, p = .02$]. Separation of response curves at Session 1 does not correlate with later measures of separation, but separation at Session 2 correlates significantly with separation at Session 3 [$r = .6083, p = .01$].

An examination of the predictive relationship between the two measures of perception using multiple regression analysis, shows that 33% of the variability in measures of slope at Session 3 of the study can be accounted for by measures of slope at Session 2 of the study [$R^2 = .33(F = 6.613), p = .02$]. The addition of slope at Session 1 of the study brings the amount of variability to 35%, however the relationship then becomes non–significant, and slope at Session 1 does not make a unique contribution to final measures of slope.

None of the variability in measures of slope at Session 3 of the study are predicted by any combination of measures of separation at all three sessions in the study.

37% of the variability in separation at Session 3 of the study can be accounted for by measures of separation at Session 2 of the study [$R^2 = .37(F = 3.8), p = .05$]. The addition of separation at Session 1 brings the amount of variability accounted for to 38% [$R^2 = .38(F = 3.8), p = .05$], but separation at Session 1 makes no unique contribution to final measures of separation.

None of the variability in measures of final separation at Session 3 of the study can be accounted for by any combination of measures of slope at any of the three sessions in the study.

Phonemic awareness

ANOVAs with phonemic awareness measures (phoneme *blending*, *segmentation* and *deletion*) as dependent variables, and *session in study* as the independent variable, show that there is a significant change in ability for all three tasks across the three sessions in the study. There is a significant difference between

Session 1 and Session 3 for phoneme blending [$F(2,48) = 3.983, p = .025$], phoneme segmentation [$F(2,48) = 6.865, p = .002$] and phoneme deletion [$F(2,48) = 6.259, p = .004$]. An examination of the change in phonemic awareness ability session by session shows that while there is a significant difference in ability on all three measures between Session 1 and Session 2 (phoneme blending [$F(1,34) = 4.077, p = .05$], phoneme segmentation [$F(1,34) = 5.762, p = .02$] and phoneme deletion [$F(1,34) = 4.922, p = .03$]), there is no significant difference in ability between Sessions 2 and 3 for any of the three measures.

On examining the three measures of phoneme awareness across all sessions in the study, it was found that all measures correlated well ($p < .05$) with all other measures at every session in the study, except for phoneme deletion at Session 1. This measure correlated with all three measures at Session 2, but with none of the other measures at Session 3.

Correlation between cue weighting and phonemic awareness

In examining the relationship between cue weighting and phonemic awareness across all three sessions in this study, we are interested in the extent to which measures of these two processes are able to predict each other. Because phoneme deletion correlated least well with other measures at each session in the study, only phoneme blending and phoneme segmentation will be examined as measures of phonemic awareness ability.

SEPARATION As noted above, multiple regression analysis shows that 37% of the variability of measures of separation at Session 3 can be accounted for by separation at Session 2 [$R^2 = .37 (F = 7.636), p = .01$], and that separation at Session 1 makes no unique contribution to this relationship.

It was also noted above that none of the variability of separation measures at Session 3 can be accounted for by any combination of measures of slope at any of the three sessions in the study.

None of the variability of separation measures at Session 3 can be accounted for by any combination of phoneme segmentation measures at any sessions in the study.

39% of the variability of separation measures at Session 3 can be accounted for by measures of phoneme blending at Session 1 [$R^2 = .39 (F = 8.434), p = .01$]. The addition of phoneme blending at Session 2 brings the amount of variability accounted for to 45% [$R^2 = .45(F = 4.992), p = .02$], but blending at Session 2 does not make a unique contribution to the variability.

72% of the variability of separation measures at Session 3 can be accounted for by a combination of measures of the separation of the response curves at the previous session (Session 2) and importantly, phonemic awareness ability (as measured by the phoneme blending task) at both previous sessions (Sessions 1 and 2) [$R^2 = .72(F = 9.497), p = .002$]. Each of these measures makes a unique contribution to the 72% of variability accounted for by all three: separation at Session 2 accounts for 16.75% of the 72% [$Beta^2 = .1675, p = .007$]; blending at Session 2 accounts for 34.65% of the 72% [$Beta^2 = .3465, p = .01$]; blending at Session 1 accounts for 48.6% of the 72% [$Beta^2 = .4860, p = .003$].

SLOPE As noted above, multiple regression analysis shows that 33% of the variability of measures of slope at Session 3 can be accounted for by slope at Session 2 [$R^2 = .33(F = 6.613), p = .02$], and that slope at Session 1 makes no unique contribution to this relationship.

It was also noted above that none of the variability of slope measures at Session 3 can be accounted for by any combination of measures of separation at any of the three sessions in the study.

None of the variability of slope measures at Session 3 can be accounted for by any combination of phoneme segmentation measures or of phoneme blending measures at any sessions in the study.

BLENDING Multiple regression analysis shows that 81% of the variability of measures of blending at Session 3 can be accounted for by blending at Session 2 [$R^2 = .81(F = 57.815), p < .0001$]. The addition of blending at Session 1 brings the amount of variability accounted for to 84%, but blending at Session 1 makes no unique contribution to this variability.

75% of the variability of measures of blending at Session 3 can be accounted for by measures of segmentation at Session 3 [$R^2 = .75(F = 40.101), p < .0001$].

30% of the variability of measures of blending at Session 3 can be accounted for by measures of separation at Session 3 [$R^2 = .30(F = 5.668), p < .03$].

None of the variability of phoneme blending measures at Session 3 can be accounted for by any combination of measures of slope at any sessions in the study.

SEGMENTATION Multiple regression analysis shows that 83% of the variability of measures of segmentation at Session 3 can be accounted for by segmentation at Session 2 [$R^2 = .83 (F = 63.553), p < .0001$]. Measures of segmentation at Session 1 make no unique contribution to this variability.

75% of the variability of measures of segmentation at Session 3 can be accounted for by measures of blending at Session 3 [$R^2 = .75 (F = 40.101), p < .0001$].

None of the variability of phoneme segmentation measures at Session 3 can be accounted for by any combination of measures of separation at any sessions in the study.

None of the variability of phoneme segmentation measures at Session 3 can be accounted for by any combination of measures of slope alone at any sessions in the study. However, 89% of the variability of phoneme segmentation measures at Session 3 can be accounted for by a combination of segmentation measures at Session 2, and importantly, measures of response curve slope at both previous sessions (Sessions 1 and 2) [$R^2 = .89 (F = 30.482), p < .0001$]. Each of these measures makes a unique contribution to the 89% of variability accounted for by all three: segmentation at Session 2 accounts for 80.71% of the 89% [$Beta^2 = .8071, p < .0001$]; slope at Session 2 accounts for 9.17% of the 89% [$Beta^2 = .0917, p = .04$]; slope at Session 3 accounts for 10.12% of the 89% [$Beta^2 = .1011, p = .03$].

CHAPTER 5

Experiment 2

This second experiment was a cross–sectional study of acoustic cue weighting strategies and phonemic awareness ability in a group of older, reading–training–delayed (and predominantly non–reading) children.

1 Subjects

Eight children participated in this study: 6 female and 2 male. An additional 3 children were also tested, but were not included in any analyses because they failed to meet the perceptual testing criteria (outlined in Chapter 4). The children ranged in age from 6;11 to 7;7, with an average age of 7;3. All of the children were native Scottish English speakers. Two of the 8 spoke a second language in addition to Scottish English, to differing degrees of bilingualism (as reported by parents). As in Experiment 1, neither of the bilingual children performed significantly differently to the monolingual children for any of the tests carried out in this study. The results of these 2 children were therefore analysed together with the results of the other 6.

The children in this study were selected from an independent school in the Edinburgh area which has a policy of delaying all reading, and reading–readiness training until the age of 8 years. None of these children, therefore, had had formal literacy training (some may have had informal exposure at home, although this was not encouraged by the school), and thus it was expected that the group would have minimal, if any, phonemic awareness[1].

For the reasons outlined in Chapter 4, none of the children had a history of chronic otitis media, defined by Nittrouer (1996*b*) as more than 3 ear infections in the first three years of life and/or the implantation of myringotomy tubes. In addition, none of the children or their siblings had ever received therapy for expressive language disorders. The above two criteria were determined by means of parental questionnaires. The children in this group were not tested for hearing problems by school authorities for hearing problems, therefore hearing ability was determined by parental questionnaire: all were reported to have normal hearing.

Additionally, having established the perceptual weighting norms for becoming–literate children in Experiment 1, the results of the reading–training–delayed children from this experiment will be compared with their beginning–reading peers from Experiment 1. For the background of the subjects in Experiment 1 (i.e. sex, age, language background), see Chapter 4.

2 Tests

The same four tests as were carried out in Experiment 1 were also carried out in this study: i) acoustic cue weighting, ii) phonemic awareness, iii) general language ability (tested by means of the BPVS), and iv) reading ability (tested by means of Schonell's Graded Word Reading Test).

Test materials and general testing procedures and criteria were identical to those used in Experiment 1 (see Chapter 4).

3 Procedure

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.

[1]It should be noted that while literacy is not trained until the age of 8 years, this school system begins second and third language training (French and German, in the case of this school) at the first year of school. However, as noted in Chapter 4 (Footnote 1), it was presumed that while this training might have some effect on overall metalinguistic ability, it would not have any effect on phonemic awareness ability. As will become clear in the Results section, this does indeed appear to be the case.

The subjects were tested once, over the course of two days (not more than one week apart). The order of testing was as follows:

- Day 1:
  1. BPVS
  2. 1st half of acoustic cue weighting
  3. 1st half of phonemic awareness
- Day 2
  1. 2nd half of acoustic cue weighting
  2. 2nd half of phonemic awareness
  3. Schonell Graded Word Reading

4 Results

As noted above, 8 children met the perceptual testing criteria. The 3 additional children who were tested but excluded from analysis either did not meet the criteria for correct perceptual responses (9 out of 10 stimuli correctly identified for the pre–test; 8 out of 10 endpoint stimuli correctly identified for the test proper), or had response curves which did not sufficiently approximate S–shaped curves. As explained in Chapter 4, the potential explanations for the perceptual behaviour of these 3 children include the adoption by the subject of response 'strategies', the subject's inability to maintain attention to the task, and the possibility that the subject may have had undiagnosed ear infections.

Six of the 8 children who met testing criteria had reading ages below their chronological age, ranging from 6;0- to 6;7 (only one of these was able to read more than 1 word correctly on the Schonell Graded Word Reading Test). Two of the 8 children were actually found to be reading at or above their chronological age range, with reading ages of 7;7 and 10;2 respectively. The 8 children had Age Equivalents based on BPVS scores which ranged from 5;9 (Confidence Interval 5;0–6;7) to 10;2 (Confidence Interval 9;2–11;4). As noted above, there were no significant differences between bilinguals and monolinguals for any of the processes tested.

All of the statistical analyses were carried out using *SPSS* running under Unix. The raw data for all tests can again be found in Appendix C.

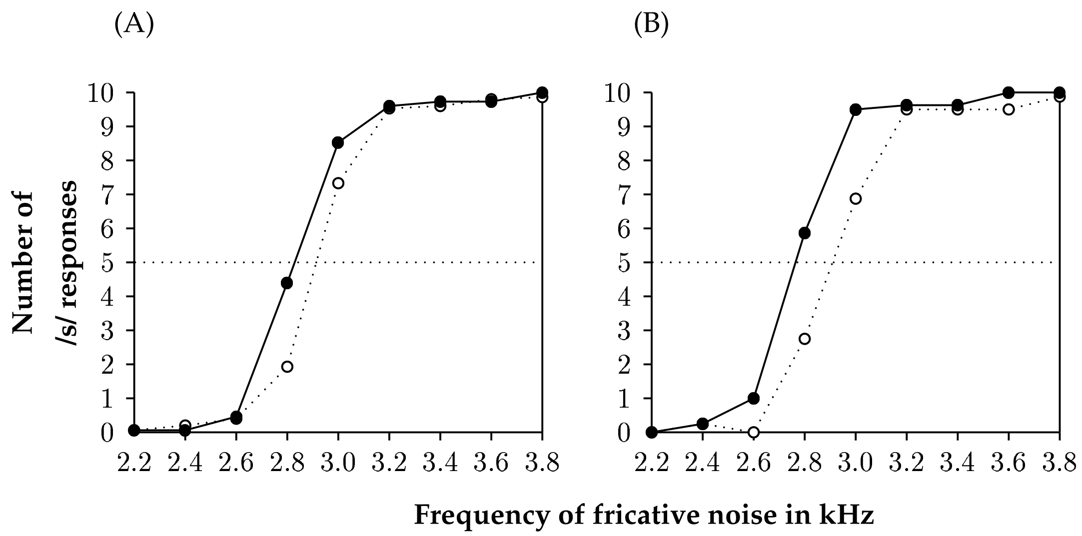

Figure 5.1: **Responses of beginning–reading group (A) and reading–training–delayed group (B) to /ʃo/–/so/ continua.** The x-axis shows the continua of fricative noises, ranging in frequency from 2.2kHz (the most /ʃ/–like) to 3.8kHz (the most /s/–like). The solid line represents a listener's /s/ responses to stimuli with /s/–transitions; the dotted line represents the same listener's /s/ responses to stimuli with /ʃ/–transitions.

4.1 Acoustic cue weighting

The graphs in Figure 5.1 show the perceptual response curves for the 15 beginning–reading children from Session 3 of Experiment 1 (A), and the 8 reading–training–delayed children (B). It can be seen that the response curves of the reading–training–delayed group are more widely separated than those of their beginning–reading peers.

ANOVAs with the perceptual measures of *slope* and *separation* as dependent variables, and *reading training* as the independent variable, show that there is a significant difference in separation $[F(1, 21) = 9.29, p = .006]$ but not slope between the reading–training–delayed children and the beginning–reading children from Session 3 of Experiment 1. Interestingly, this is the first stage at which there is a significant difference in separation between the two groups: a comparison of the reading–training–delayed children to the reading children at Session 1 and again to the same group at Stage 2 shows no significant difference in separation between the two groups. There was no significant difference in slope measures between the reading–training–delayed group and the beginning–readers at any of the three sessions of Experiment 1.

An examination of the two measures of acoustic cue weighting shows that there is no significant correlation between the slope and the separation of response curves for the reading–training–delayed children.

Neither the slope nor the separation of the children's response curves correlates with either general language ability, or word reading ability.

4.2 Phonemic awareness

The mean scores for the phonemic awareness tests for this group were: phoneme blending: 17/50, phoneme segmentation: 14/50, and phoneme deletion: 10/40. It should be noted at this point that the 2 reading children (as would be expected) had good scores on all three phonemic awareness tests. More unexpectedly, a third (non–reading) subject also had good scores on the phoneme blending and phoneme segmentation tasks, but not the phoneme deletion task.

ANOVAs with the phonemic awareness measures of *blending*, *segmentation* and *deletion* as dependent variables, and *reading ability* as the independent variable, show that there is a significant difference in all three between the reading–training–delayed children and the beginning–reading children from Session 3 of Experiment 1 (blending $[F(1,21) = 14.415, p = .001]$, segmentation $[F(1,21) = 18.493, p < .001]$ and deletion $[F(1,21) = 8.472, p = .008]$). This is the first stage at which there is a significant difference in phoneme deletion ability between the two groups: a comparison of the reading–training–delayed children to the reading children at Session 1 and Session 2 shows no significant difference in deletion between the two groups. However, there is a significant difference in both phoneme blending and phoneme segmentation between the reading–training–delayed children and the reading children at Session 2 (blending $[F(1,21) = 13.482, p = .001]$, segmentation $[F(1,21) = 12.131, p = .002]$). There was no significant difference in any measures of phonemic awareness ability between the reading–training–delayed group and the beginning–readers at Session 1 of Experiment 1.

4.3 Correlation between cue weighting and phonemic awareness

The graphs in Figure 5.2 show acoustic cue weighting in terms of separation of response curves, and phonemic awareness in terms of phoneme blending for the beginning–reading children (A) and the reading–training–delayed children (B). The graph is divided into quadrants at the median phoneme blending score from

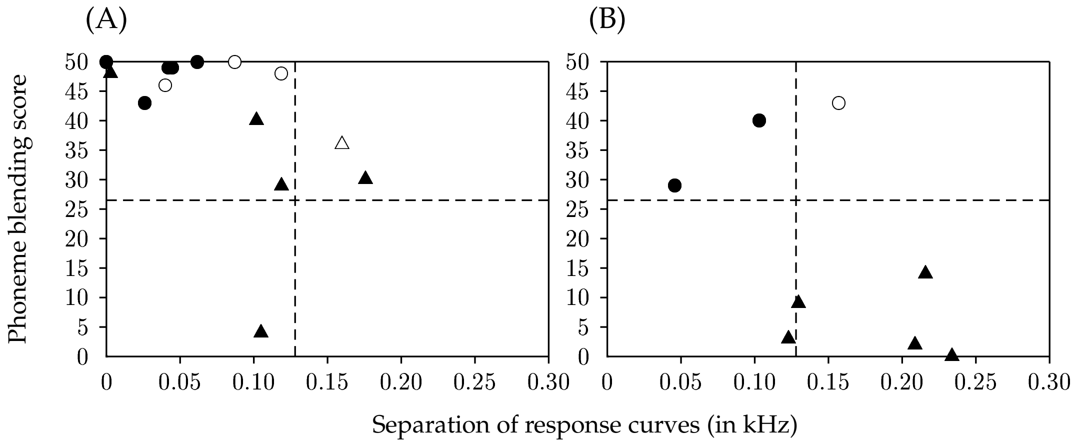

Figure 5.2: **Relationship between perception and awareness: beginning–reading group (A) and reading–training–delayed group (B).** The graph is divided into quadrants at the median phonemic awareness score (on the y-axis) and the most global of the adults' cue weighting responses (on the x-axis). Each point on the graph represents one subject.

Session 1 of Experiment 1: 26.6/50, and the most global of the adults' cue weighting responses: 0.13kHz separation between response curves. It is clear from these graphs that the reading–training–delayed children have both poorer phonemic awareness and more global cue weighting in general than the beginning–readers. Additionally it can be seen that while there are three reading–training–delayed children who have begun to develop some level of phonemic awareness, only two of them have developed analytical cue weighting. Finally, the reading–training–delayed children do not appear to develop analytical cue weighting without having developed good phonemic awareness.

CHAPTER 6

Discussion, conclusions and implications

1 Summary of study

The main aim of this thesis has been to explore the relationship between shifts in acoustic cue weighting and the development of phonemic awareness. A clear correlation between these two processes has been shown to exist by Nittrouer (1996*b*). However, Nittrouer's study did not provide any evidence as to the exact nature of the relationship. In particular, Nittrouer was unable to address the issue of the possible causal direction of the relationship—that is, the possible influence of one process on the development of the other. Additionally, studies to date have provided evidence for both possible causal directions: McBride–Chang and colleagues, (McBride-Chang 1996, Manis et al. 1997, McBride-Chang et al. 1997), for example, have found that early perceptual abilities predict later performance on phonemic awareness tasks, while others (de Gelder & Vroomen 1992, Flege et al. 1992, Morais & Kolinsky 1995) have suggested that the development of phonemic awareness may impact on speech perception strategies. The two studies that form the basis of this thesis, therefore, were designed to address this issue.

The first study (Experiment 1) was a longitudinal study of 18 normally developing school children, aged 5;2 through 6;0 years (average age 5;8) at the beginning of the study. All were in their first year of school at the beginning of the study, and had undergone approximately 6–7 months of reading/reading–readiness training. The children were tested three times, over the course of 7 months. At Sessions 1, 2 and 3 the children were assessed as to their acoustic cue weighting strategies and their phonemic awareness abilities. Additionally at Session 1 and Session 3 the children were also tested on their word reading abilities (Graded Word Reading Test Schonell & Goodacre 1971) and their general language abilities (BPVS: Dunn et al. 1982). The acoustic cue weighting tests made use of a /ʃo/–/so/ ('show–sew') contrast which varied in terms of frequency of fricative noise, and vowel–onset formant transition configuration. Phonemic awareness was tested by means of three tasks, all using real words: phoneme blending, phoneme segmentation and phoneme deletion.

The second study (Experiment 2) was a cross–sectional study of a group of normally developing children, aged 6;11 to 7;7 (average age, 7;3). This group was selected from a school which delays all reading/reading–readiness training until children are approximately 8 years of age: the children in this group had therefore not received any formal literacy instruction at the time of testing. Using the same tests as those used in Experiment 1, the children in this second study were assessed on their acoustic cue weighting strategies, their phonemic awareness skills and their word reading and general language abilities.

While the studies in this thesis were carried out on the premise that a relationship exists between acoustic cue weighting and phonemic awareness, and furthermore were designed to elaborate on the nature of this relationship, the results of these studies also give us some insight into the development of perception and awareness as separate processes. The first two parts of this chapter will therefore examine the new evidence that these studies provide regarding the development of acoustic cue weighting and of phonemic awareness. The second part of this chapter will then go on to determine what can be concluded from these studies regarding the nature and direction of the relationship between the two processes. Finally, this chapter will conclude with an examination of the implications of this study for our understanding of the development of speech perception in general.

2 Acoustic cue weighting

2.1 *Summary of results*

Experiment 1

The results of Experiment 1 show a replication of the shifts in cue weighting observed by Nittrouer and colleagues. Importantly, this is the first time that such shifts have been observed in the same group of children by means of a longitudinal study. The children at the beginning of the longitudinal study have significantly shallower and more separated response curves than do the adults. This significant difference can also be seen between the children at Session 2 and the adults. By Session 3, there is no longer any significant difference between the slope or the separation of the adults' and the children's response curves.

If the perceptual progress of the children in Experiment 1 is examined, we can see that while there is a significant difference in separation of response curves between Session 1 and Session 3, there is no significant difference between slope of response curves at these sessions. Neither slope nor separation of response curves are significantly different between Sessions 1 and 2 or between Sessions 2 and 3, although the separation of the children's response curves approaches significance between Sessions 2 and 3.

Additional support for the phenomenon of shifts in acoustic cue weighting can be seen by comparing day–to–day variation in perceptual behaviour with changes seen between sessions. Because children's perceptual behaviour can be highly variable, it was important to ensure that any differences found between different sessions of the study were due to a genuine shift in perceptual behaviour, and not simply to the sort of variation that might be found from one day of testing to the next. The fact that testing took place over two days, and a full set of perceptual and awareness tests were carried out each day, meant that it was possible to test this. Statistical analysis (ANOVA) showed that the shifts observed in perceptual behaviour were significantly accounted for by differences across the three sessions of the study, and not by any variation that might have occurred from day to day.

Experiment 2

The perceptual behaviour of the children in Experiment 2 shows that the perceptual phenomenon of shifts in cue weighting is not maturational. The reading–training–delayed children in this study have significantly shallower and more separated response curves than do their beginning–reading peers (i.e. the children at Session 3 of Experiment 1). Additionally, the reading–training–delayed group did not perceive the stimuli significantly differently from the children at Session 1 of Experiment 1. The implications of this finding will be discussed in more detail below.

2.2 *Issues*

This section will attempt to address some of the issues raised in Chapter 2 (Section 1.2) regarding acoustic cue weighting. These issues range from those which are reasonably easily addressed—such as whether the type of stimulus used, i.e. synthetic or natural, affects shifts in cue weighting—to those which are clearly outside the scope of one small study—such as whether speech perception is a process which is controlled by a dedicated system, or simply by a general auditory system. The extent to which the results of this thesis can address any of these issues, therefore, will also vary—from providing plausible but largely hypothetical support for a certain viewpoint, to providing clear answers to some of the questions asked in Chapter 2.

Phonetic context

As noted in Chapter 2, one of the more serious problems with Nittrouer's DWS model is the fact that it is based on listeners' perceptual strategies for a very limited number of contrasts. In addition to this, even within those contrasts which *have* been studied, a certain amount of variation has been found in the extent of the transitional effects on which the DWS model is based: that is, the degree to which a listener's responses differ depending on the transitional context of the stimuli. Specifically, the more extensive vowel onset transitions in /u/ and /a/ following /s/ and /ʃ/ engendered a greater transitional effect than did the less extensive transitions from /s/ and /ʃ/ into /i/. While these results led Nittrouer (Nittrouer & Studdert-Kennedy 1987, Nittrouer 1992) to suggest that transitional effects are proportional to the extent of the transitions, this apparent variability in the cue weighting shift phenomenon, along with the lack of further evidence for the phenomenon, weakens claims made regarding the DWS model.

The results of the current studies show that shifts in acoustic cue weighting occur for /ʃo/–/so/ contrasts to a comparable extent to those seen for other contrasts in Nittrouer's studies. However, while these results lend support to the DWS model, they do not address the problem of the limited range of experimental evidence for shifts in cue weighting. While the /ʃo/–/so/ contrast has not specifically been tested by Nittrouer and colleagues, the perception of /ʃ/–/s/ has been well established as engendering cue weighting shifts in the context of certain vowels. As a result, the use of a further /ʃ/–/s/ contrast, albeit with a new vowel context, does not go very far towards expanding our understanding of the phenomenon of cue weighting shifts.

The replication of Nittrouer's results with a previously untested vowel context does, on the other hand, lend a certain amount of support to Nittrouer's argument that the degree of transitional effect is dependent on the extent of the transitions. The results of Experiment 1, with the back vowel /o/, show that reasonably extensive transitions do indeed engender reasonably extensive transitional effects. This supports the contention that different vowel contexts, in giving rise to different transitional effects, are responsible for the observed size of shifts in cue weighting.

The results of this study also address the third, and slightly more abstract, problem brought up in Chapter 2 in relation to the lack of extensive experimental evidence for the DWS—specifically, the claim that children always attend to transitions and adults always attend to 'informative' cues (which, in the case of the contrasts studied thus far by Nittrouer, have been non–transitional cues). The results of Experiment 1 support the theory that young children attend more heavily to transitional information, although given the fact that the perceptual tests used in the study were not drastically different from those which have been used previously by Nittrouer and colleagues, this is possibly unsurprising.

What is more surprising is what the combined results of both Experiment 1 and Experiment 2 tell us about the cues that are weighted more heavily by *older children and adults*. The results of Experiment 1 show the type of perceptual behaviour that has been found numerous times by Nittrouer: the adult subjects' perception was more analytical than the children's, meaning that they weighted transitional cues less heavily and fricative noise cues more heavily than did the children. Additionally, the results of Experiment 1 also showed that over the course of 7 months, the children who had started the study with very global perceptual strategies shifted to more analytical strategies. Again, this finding corresponds well with the results of Nittrouer's cross–sectional studies. Up to this point, therefore, the results of this thesis support Nittrouer's claims that adults and older children always weight cues differently to younger children.

However, the results of Experiment 2 do not fit so straightforwardly into Nittrouer's DWS model, which, it should be recalled, stands for *Developmental* Weighting Shift model. The reading–training–delayed children in this study were at least the same chronological age as both the older children in Nittrouer's studies and the older, beginning–reading children in Experiment 1, and many were in fact older. Despite this, they did not weight the cues available to them any differently than did the *younger* children from Experiment 1: the reading–training–delayed children weighted transitional cues significantly differently to their their beginning–reading peers, but not to the younger beginning–readers.

In terms of the issue in question (i.e. which cues the DWS predicts adults and children will attend to) the results of Experiment 2 indicate that it is *not* the case that older children and adults *always* weight cues differently to younger children. They also indicate that Nittrouer's definition of 'informative' cues as those that are weighted most heavily by older children and adults will have to be refined to take into account the factors that influence adults' cue weighting strategies. A more detailed discussion of these factors and their relationship to perceptual changes will be carried out later in this chapter.

It should be noted that these results also strongly suggest that there is a more fundamental problem with Nittrouer's model of the phenomenon of cue weighting shifts, in particular the developmental stance of the model. We will return to this crucial issue below, when we discuss the full implications of this study on our understanding of speech perception.

Auditory processing

The results of these studies are much more limited in their ability to address the question of whether speech perception is controlled by a speech–specific, or a general auditory system. However, the results do allow for a certain amount of speculation regarding this issue. It was noted in Chapter 2 (Section 3.2) that the auditory theory, i.e. the theory that speech perception is just one capacity of a general auditory system, seems to restrict the possible direction of influence between perceptual weighting and phonemic awareness. It seems unlikely that a system which is designed for perception of all sounds would be impacted upon by the development of a skill such as phonemic awareness, which is highly speech– and language–specific. The reverse, however, can also said to be true—that is, if a listener's speech perception *is* impacted upon by the development of phonemic awareness, then it could be inferred that speech perception is much less likely to be a process which is controlled by general auditory capacities.

The details of the relationship between cue weighting and phonemic awareness will be discussed in greater detail below. However, simply from the results of Experiment 2 it can be seen that changes in acoustic cue weighting do not appear to occur in the absence of phonemic awareness development. This suggests that phonemic awareness does have an impact on perceptual strategies, thus supporting the view that speech perception is controlled by a dedicated speech–specific system.

Synthetic vs. natural stimuli

As noted in Chapter 2 (Section 1.2), Nittrouer herself has acknowledged that one potential problem with the stimuli in her original 1987 and 1992 studies is the way in which they combine natural and synthetic speech. Recall that Nittrouer's stimuli were created by concatenating synthetic fricative noises and natural vowels. This means that it is impossible to determine whether the young children who weighted vowel–onset transitions more heavily than fricative noises did so because they preferred the transitional cues over the non–transitional (fricative) cues, or because they preferred the natural portions of the stimuli to the synthetic portions. Nittrouer & Miller (1997*b*) were able to replicate the results of Nittrouer's original studies using all–synthetic speech. However, because the vowel portions of these new stimuli were highly stylised, it remains difficult to determine whether it was strictly the *transitional* effect of these stimuli that attracted the young children's attention.

The studies in this thesis were designed with this issue in mind. The perceptual test used in both Experiments 1 and 2 made use of wholly synthetic stimuli, which were created by a method called copy–synthesis. This method preserves, to a much greater extent than in Nittrouer & Miller's (1997*b*) study, the configuration of natural speech—in particular, the non–linear change in frequency of the vowel–onset formant transitions. The fact that these all–synthetic stimuli engendered responses which are consistent with the results of Nittrouer's cue weighting studies, indicates that Nittrouer's conclusions regarding her own studies are most likely correct. The perceptual behaviour of the young children in Nittrouer & Studdert-Kennedy's (1987) study, and those that followed it, does appear to be due to the children's perceptual attention to transitional cues, and *not* their perceptual preference for non–synthetic speech.

Slope and separation of response curves

Chapter 2 (Section 1.2) discusses Nittrouer's claim that the *slope* and the *separation* of the response curves from her studies are *both* the result of the same perceptual phenomenon, namely listeners' weighting of acoustic cues. Nittrouer mounts this argument predominantly against the possibility that the cause of the children's shallower response curves is simply that they were inattentive to the perception task. Her argument is based on two main pieces of evidence. The first is simply the fact that age–dependent differences were found for *both* the slope and the separation of the response curves. Nittrouer (1992) argues that if the degree of slope and separation were due to two different phenomena—and specifically if the slope were the result of the children not paying attention to the task—then one would not expect to see both measures changing across age groups. The other piece of evidence comes from one of the studies reported in Nittrouer's (1992) paper. In this study Nittrouer found that for stimuli with *ambiguous* transitional cues, the slopes of young children's response curves were much shallower than adults', however for stimuli with *unambiguous* transitional cues, the slopes of young children's response curves were not significantly different from the adults'. Nittrouer (1992) states that this is an indication that children's shallower response curves are not due to lack of attention: if they were, then there shouldn't have been a difference in children's response curves for stimuli with ambiguous and unambiguous transitional cues.

However, lack of attention is not the only alternative cause of children's shallower responses. Other researchers (e.g. Hazan & Barrett 1999, Simon & Fourcin 1978) have suggested that children's responses to continuously varying stimuli may become gradually more categorical as their speech perception mechanism matures. Hazan & Barrett (1999) have shown that the slopes of children's response curves continue to get steeper even after they have shifted away from heavier weighting of transitional cues, suggesting that the slope and the separation of response curves are indeed dictated by two different phenomena.

The results of the studies in this thesis provide highly convincing support for this view. Examining the slope and the separation of the response curves for each subject group, it becomes clear that there is very little correlation between the steepness of individuals' response curves and the relative placement of their category boundaries. For the adults and the reading–training–delayed children, both of whom were only tested once, no correlation at all was found between slope and separation measures for either group. For the beginning readers from Experiment 1, a significant correlation between slope and separation was found only at Session 2 of the three sessions in the study. In terms of the predictive relationship between these factors, the listeners' response curve slopes were found to be very poor predictors of the listeners' later response curve separation. Similarly, the separation of response curve slopes at early sessions in the study are very poor predictors of later slope of response curves. All of this evidence would seem to indicate that the slope and the separation of response curves should not be treated as two measures of the same aspect of perception, but rather as measures of two different aspects of perception. The question to be asked at this point, then, is what these two different aspects might be.

If we compare the perceptual responses of the reading–training–delayed children to those of the beginning–reading children at each session in the longitudinal study, we find one possible answer to this question. There is no significant difference in either slope or separation of response curves between the reading–training–delayed group and the beginning–reading children at Session 1 of Experiment 1—recall that at this session, the children in the beginning–reading group were younger than those in the reading–training–delayed group, but were predominantly non–reading as well. Moving on to Session 2 of Experiment 1, again we find no significant difference in either slope or separation of response curves between the reading–training–delayed children and the beginning–reading children. However, when we reach Session 3 of the experiment, at which point the beginning–reading children are nearing the same age as the reading–training–delayed children, and are predominantly all reading at a reasonable level, the pattern of responses changes. What we observe is that while there is now a significant difference in the separation of response curves between the Session 3 beginning–readers and the reading–training–delayed group, there remains *no* significant difference in the slope of response curves between beginning–readers and the reading–training–delayed group. At this session, the factor that is the same across the groups is age, which suggests that the slope of response curves is, as suggested by Simon & Fourcin (1978) and Hazan & Barrett (1999), related to the gradual maturation of the speech perception system. By the same token, we can hypothesise that something to do with the factor that is different between the groups—i.e. reading training and/or skill—is responsible for the differences in separation of response curves between the two groups. In fact, as will be discussed further below, it is not strictly literacy level, but a related skill, phonemic awareness, that is related strongly to the separation of response curves in both these groups. However, the evidence presented here makes it clear that the slope and the separation of a listener's response curves are not as highly correlated as Nittrouer has assumed. This suggests, in turn, that speech perception is not a completely homogeneous construct.

3 Phonemic awareness

3.1 Summary of results

Experiment 1

The results of Experiment 1 show the development of phonemic awareness over the course of 7 months. There is a significant difference between Session 1 and Session 3 of the study for all measures of phonemic awareness. Examining these differences session by session, it becomes clear that the greatest change in phonemic awareness ability occurred between Session 1 and Session 2: there is a significant difference between these two sessions for all three measures of phonemic awareness, while there is no significant difference between Session 2 and Session 3 for any of the three measures.

Experiment 2

If we compare the phonemic awareness ability of the reading–training–delayed children in Experiment 2 to that of the children in Experiment 1, a similar, and slightly more detailed picture emerges. There is no significant difference between the reading–training–delayed children and the beginning readers at Session 1, for any phonemic awareness measure. We do, however, find a significant difference between the reading–training–delayed children and the beginning readers at Session 2 for measures of phoneme blending and phoneme segmentation, but not for phoneme deletion. Finally, comparing the reading–training–delayed group to the beginning–readers at Session 3, there is a significant difference between the groups for all three measures of phonemic awareness.

3.2 Issues

Again, in addition to providing information about the ways in which phonemic awareness develops, this study is also able to address a number of the issues raised in Chapter 2 (Section 2.3) regarding phonemic awareness. Because so much more is already known about the development of phonemic awareness than about the development of acoustic cue weighting, this study will not provide much new evidence regarding this process. However, there are some noteworthy results which will be discussed here, before we move on to a discussion of the relationship between the development of this process and changes in acoustic cue weighting.

Phonemic awareness testing

In this section we will determine whether the results of the two studies in this thesis can in any way address the two questions raised in Chapter 2 regarding phonemic awareness testing: namely, what constitutes a phonemic awareness test, and how aware is aware? The studies can in fact only answer these questions in a limited way. In terms of the first question, this is because care was taken in the design of the phonemic awareness stimuli to ensure that they only tested phonemic awareness: all three tasks were designed to tap phoneme, rather than onset–rime or syllable awareness, and all tasks demanded an explicit awareness of phonemes (all required phoneme manipulation) rather than simply an implicit sensitivity towards phonemes.

There are two parts to the second question. The first regards the level of success required on a phonemic awareness test for a subject to be considered aware. Unfortunately, answering this question thoroughly would require large standardising tests of early–school–age children. However, the second half of the question, which regards whether a subject must be equally successful at all phonemic awareness tasks to be considered phonemically aware, can be addressed to a certain extent by the studies in this thesis.

Statistical analysis of the longitudinal study in Experiment 1 shows that at each session all three measures of phonemic awareness were very highly correlated. This would suggest that all three tasks were tapping the same process to a large extent. It would also suggest that one should expect that if a child is phonemically aware, they should perform to the same level for all phonemic awareness measures. However, a closer examination of the results also suggests that assessing phonemic awareness might not be as straightforward as these correlations make it seem. First, phoneme blending and phoneme segmentation were *more* highly correlated with each other than either was with phoneme deletion. Additionally at Sessions 1 and 2, phoneme deletion was more highly correlated with phoneme segmentation than with phoneme blending. This is not entirely surprising, as in order to successfully complete a phoneme deletion task, the subject needs to first be able to segment the phoneme to be deleted from the rest of the word. The fact that phoneme deletion is so highly correlated at each session with the other two factors, however, is slightly misleading.

First, the development of good phoneme segmentation and blending skills appears to precede the development of good phoneme deletion skills: many of the children in Experiment 1 showed good blending and segmentation skills (i.e. scored over 50%) without showing any phoneme deletion skills at all (in terms of the test administered). Contrastively, there was no child in Experiment 1 who had good phoneme deletion skills without also having good phoneme blending and segmentation skills. Again, both of these results are unsurprising, given the nature of the tasks used to test awareness. First, both phoneme blending and phoneme segmentation as basic tasks are less complex than phoneme deletion: they involve only one manipulation, while phoneme deletion requires more than one. Thus one possible explanation for the later development of phoneme deletion skills compared to phoneme blending skills in Experiment 1 is that phoneme deletion is fundamentally more cognitively demanding than the other phonemic awareness tasks.

Additionally, the results of the phoneme deletion tests were generally bimodally distributed—that is, the children either performed at floor level or at ceiling level. The development of phoneme blending and segmentation skills, on the other hand, appears to be much more gradual. The point to note here is that the stimuli in the phoneme blending and phoneme segmentation tasks varied more than those in the phoneme deletion task. The phoneme blending and segmentation stimuli were designed to increase throughout the test in number of phonemes per stimulus, and in number of phonemes per consonant cluster. This meant that there was an increase throughout these tasks in the number of individual phonemes to be held in memory. The phoneme deletion task, on the other hand, consistently had 2 consonants in an initial consonant cluster, and furthermore only ever required the subject to hold 2 items in memory—the initial phoneme, and the 'rest of the word.' This difference in task structure could possibly explain why bimodal results were seen for the phoneme deletion task, and not for the blending and segmentation tasks: although phoneme deletion was a more complex manipulation to master, once it was understood, completing the test did not place any increased demands on the subjects' memory or cognitive ability.

Second, not only does phoneme deletion ability never develop before phoneme blending or segmentation, but early phoneme deletion is the only measure that does not correlate with later measures of phoneme awareness. As described in more detail in Chapter 4, where phoneme blending and segmentation at Session 1 correlate significantly with phoneme deletion at Sessions 2 and 3, and phoneme blending and segmentation at Session 2 also correlate significantly with phoneme deletion at Session 3, the relationship does not seem to work as well the other way around. Specifically, phoneme deletion at Session 1 is the only measure that does not correlate at all with either phoneme blending or segmentation at Session 3. Additionally, the correlations between deletion at Session 1 and blending and segmentation at Session 2, as well as those between deletion at Session 2 and blending and segmentation at Session 3 show the lowest correlation of all the inter–measure relationships. These findings add to our understanding of the delay in development of phoneme deletion relative to blending and segmentation. The results show that the strongest relationship between phoneme deletion and the other two measures is between early measures of blending and segmentation, and later measures of deletion. This would seem to indicate that in addition to being more demanding than blending and segmentation tasks, phoneme deletion may be dependent on the earlier development of these skills.

Keeping this in mind, what can we now say in answer to the questions asked at the beginning of this section? To a certain extent it should be expected that phonemic awareness skills should be transferable—that is, in order to be considered aware, a subject's skill should not be restricted to an ability to only perform one phonemic awareness task. However, as the results above suggest, there may be some phoneme awareness tasks that are cognitively more demanding than others—a child that is in the process of developing cognitively, should not necessarily be expected to be able to complete such tasks as easily as less complex tasks. Additionally it is clear that there may also be tasks which depend for their success on the development of other phonemic awareness skills—again subjects should not be expected to be able to complete such tasks before they have developed the prerequisite skills. Clearly, therefore, as noted by McBride-Chang (1995*b*) there are both cognitive and other metalinguistic skills which are implicated in the completion of phonemic awareness tasks, and not all tasks will require the same skills as others. It should not, therefore, be assumed that all phonemic awareness tasks are equal—just that they are highly related in the level of awareness that they tap.

4 The relationship between acoustic cue weighting and phonemic awareness

At this point in the chapter we turn to the main question that has driven both of the experiments in this thesis: how can one characterise the relationship between shifts in acoustic cue weighting and the development of phonemic awareness? This thesis has been designed, in particular, to determine the *direction* of the relationship between these two processes.

4.1 Does one process develop before the other?

The first step in answering the above question is to determine whether one of the two processes in question consistently develops before the other. This will allow us to constrain our hypotheses regarding causal direction. If, for example, a high score on the phonemic awareness tests always follows the development of analytical perceptual strategies as measured by the perceptual tests, then it is unlikely that phonemic awareness has a causal influence on the development of perceptual strategies. The same can be said if the development of analytical perceptual strategies, as measured by the perceptual test, consistently follows the achievement of high scores on the phonemic awareness tests. These two

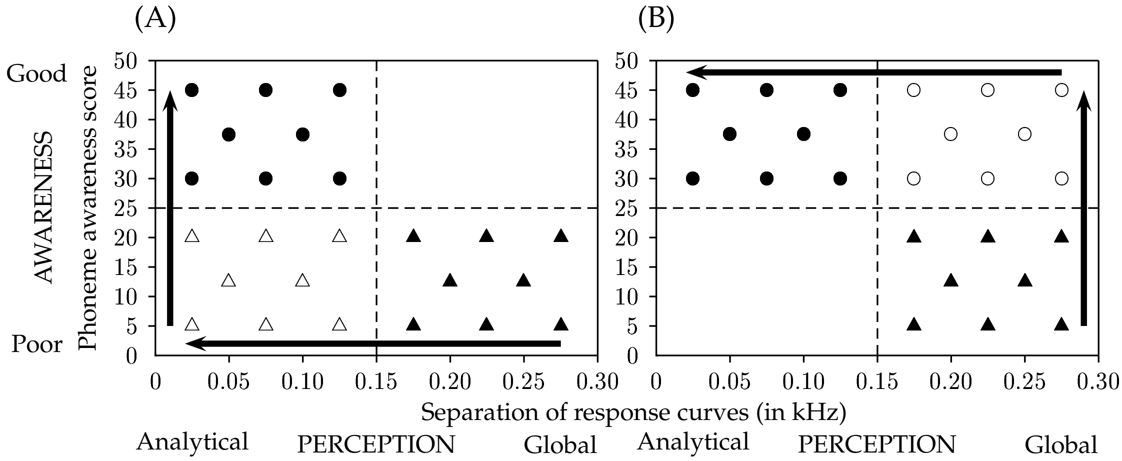

Figure 6.1: **Graphs of two hypothetical relationships between cue weighting and phonemic awareness.** Graph (A) illustrates the situation in which perceptual strategies change before the development of phonemic awareness begins; Graph (B) illustrates the situation in which phonemic awareness develops before perceptual strategies change.

hypothetical situations are illustrated by the graphs in Figure 6.1 (introduced in Chapter 3).

Figure 6.2 shows the results of all three sessions in Experiment 1. Recall that in these graphs perceptual strategy is measured in terms of the degree of separation (in Hz) between the two response curves. This follows the procedure used by Nittrouer (1996*b*) and thus allows for comparisons to be made between the two studies. Additionally, as will be seen later in this chapter, there are other, more theoretical reasons for choosing this as a measure of acoustic cue weighting instead of the slope of subjects' perceptual response curves. The measure of subjects' phonemic awareness is shown on these graphs in terms of their success on the phoneme blending test. This test was chosen simply because it correlates slightly more highly with the separation between the response curves than phoneme segmentation and phoneme deletion.

An examination of each graph in Figure 6.2 separately shows that at each session there were a number of children who had good phonemic awareness but who had not yet developed analytical perceptual strategies. Additionally, at each session, there were *no* children who had developed highly analytical perceptual strategies, but who still had poor phonemic awareness. Just from these individual sets of results, it would appear that phonemic awareness develops before changes in perceptual weighting.

Examining the movement of the data points from Session 1 through Session 3, it becomes clear that the pattern observable at each individual session of the study

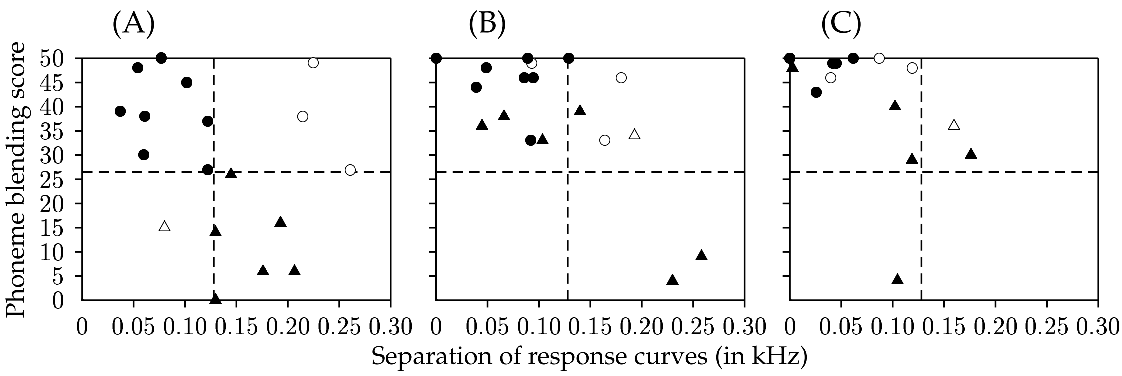

Figure 6.2: **Results of all three sessions of Experiment 1: relationship between cue weighting and phonemic awareness.** Graph (A) shows the results of Session 1; Graph (B) shows the results of Session 2; Graph (C) shows the results of Session 3.

is also the pattern that can be seen throughout the study. Children who had both poor phonemic awareness and global speech perception strategies at Session 1 (filled triangles) developed better phonemic awareness before developing more analytical perceptual strategies at later sessions in the study. Children who had already developed good phonemic awareness at Session 1 (open circles) then went on to develop more analytical perceptual strategies. Once again it is important to note that at no point did any child develop strongly analytical perceptual strategies while still having very poor phonemic awareness skills.

The statistical analyses carried out on these results (given in more detail in Chapter 4) support this view. A significant difference was seen between the Session 1 and the Session 3 results for both perceptual strategy and phonemic awareness. However, when analysed session by session, the change in phonemic awareness ability is greater between Session 1 and Session 2, than between Session 2 and Session 3 (indicated by a significant difference between scores at Session 1 and scores at Session 2, but a very non–significant difference between scores at Session 2 and scores at Session 3). For the changes in perceptual strategy, on the other hand, the difference in cue weighting approaches significance between Session 2 and Session 3, but is not significant at all between Session 1 and Session 2. This again suggests that the major changes in phonemic awareness take place earlier than any major changes in acoustic cue weighting strategy.

If we compare the reading–training–delayed children (Experiment 2: see Figure 6.3) to the reading children at every session in the longitudinal study (Experiment 1: see Figure 6.2), we get further evidence that phonemic awareness develops before shifts in acoustic cue weighting.

The reading–training–delayed children are not significantly different from the beginning–reading children at Session 1 for either phonemic awareness or perceptual strategy. In terms of phonemic awareness, this is not surprising, as many of the beginning–reading group were predominantly non–readers at this session, and thus had not had a chance to develop good phonemic awareness skills.

However, in terms of the perceptual strategies this result is slightly more surprising, in particular if one accepts Nittrouer's theory that shifts in acoustic cue weighting are maturational: the reading–training–delayed children are older than the beginning–readers, and thus should presumably have more 'mature'—i.e. more analytical—perceptual strategies than the beginning readers. As noted above, the results of Experiment 2 call into question the *developmental* or maturational aspect of Nittrouer's model of cue weighting shifts. Again, we will return to this issue later in this chapter.

Moving on to a comparison of the reading–training–delayed group to the beginning–readers at Session 2, we find that while there is still no significant difference between the reading–training–delayed children and the beginning reading children for perceptual strategy, there *is* a significant difference between the two groups for two of the three phonemic awareness tests: blending and segmentation.

Finally, a comparison of the reading–training–delayed group and the beginning–readers at Session 3 shows an increased difference in phonemic awareness ability between the two groups. The reading–training–delayed children are significantly worse than the beginning–reading group for all measures of phonemic awareness: blending, segmentation, and deletion. Importantly, the results also show that at this session, the reading–training–delayed children are significantly different from the beginning–reading children for perceptual strategy: the reading–training–delayed group have significantly more separated response curve slopes than do the beginning–readers.

These results indicate, once again, that the beginning–reading children started to develop phonemic awareness before their perceptual strategies began to change.

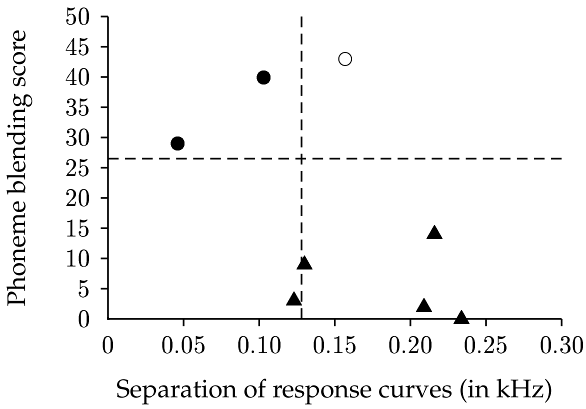

Figure 6.3: **Results of Experiment 2: relationship between cue weighting and phonemic awareness**

4.2 Does one process predict the other?

Having determined that phonemic awareness appears to develop before changes in acoustic cue weighting, the question to be answered at this point, then, is whether the relationship between the two skills is in any way causative. That is, does the process which develops first—phonemic awareness—actually have some causal effect on the process which develops second—analytical perceptual strategy? There are two parts to the answer to this question. First, what happens to cue weighting strategies when phonemic awareness *does* develop? Second, what happens to cue weighting strategies if phonemic awareness *does not* develop?

Statistical analyses of the results of the longitudinal study give a clear indication that phonemic awareness development has some causal influence on changes in cue weighting strategies. Multiple regression analysis shows that 72% of the variance seen in the separation of the response curves at Session 3 can be accounted for by a combination of the separation of the response curves at the previous session (Session 2) *and* importantly, phonemic awareness ability (as measured by the phoneme blending task) at both previous sessions (Sessions 1 and 2) $[R^2 = .72(F = 9.497), p = .002]$. Each of these measures makes a unique contribution to the 72% of variability accounted for by all three: separation at Session 2 accounts for 16.75% of the 72% $[Beta^2 = .1675, p = .007]$; blending at Session 2 accounts for 34.65% of the 72% $[Beta^2 = .3465, p = .01]$; blending at Session 1 accounts for 48.6% of the 72% $[Beta^2 = .4860, p = .003]$.

Additionally, if we examine the relationship from the other direction, it is also clear that it is much less likely that changes in perceptual strategy have a causal effect on the development of phonemic awareness. We have already shown that there was no child who developed analytical cue weighting strategies before developing good phonemic awareness. Multiple regression analysis goes on to show that 81% of the variance seen in phonemic awareness (in terms of phoneme blending) at Session 3 can be accounted for to a very high level of significance by phoneme blending scores at Session 2 $[R^2 = .81(F = 57.815), p < .0001]$. When perceptual weighting (as measured by separation of response curves) at Session 1 and Session 2 are added, however, the percentage of variance accounted for only goes up to 82%, and the level of significance falls $[R^2 = .82(F = 17.393), p = .0002]$. The most important point to note about this last analysis is that neither perceptual strategy at Session 1 nor perceptual strategy at Session 2 make a unique contribution at all to phonemic awareness at Session 3. This means that perceptual strategy (in terms of separation of response curves) does not have any causal effect on the development of phonemic awareness (as measured by phoneme blending).

These comparisons show the conditions under which analytical perceptual strategies *do* develop. However, for a more clear understanding of the conditions under which they *do not* develop, we will have to turn to Experiment 2.

Figure 6.4 once again shows the results of Session 3 of Experiment 1 (Graph A) and of Experiment 2 (Graph B), displayed in terms of separation of response curve slopes and phoneme blending ability. Both of these graphs represent the behaviour of groups of normally developing, same–age children. The only difference between these two groups is in the amount of literacy training that they have been receiving, and thus (presumably) the degree of phonemic awareness that they have developed. A comparison of these two graphs shows clearly that *without* the development of phonemic awareness (as in the case of most of the reading–training–delayed children) acoustic cue weighting strategies do not appear to change, even with a chronological change in age. Statistical analyses support this: as noted above, there is a significant difference in perceptual strategy between the reading–training–delayed children and their beginning reading peers.

What can all of these results together tell us? First, we can say that phonemic awareness develops *before* changes in acoustic cue weighting take place, and that

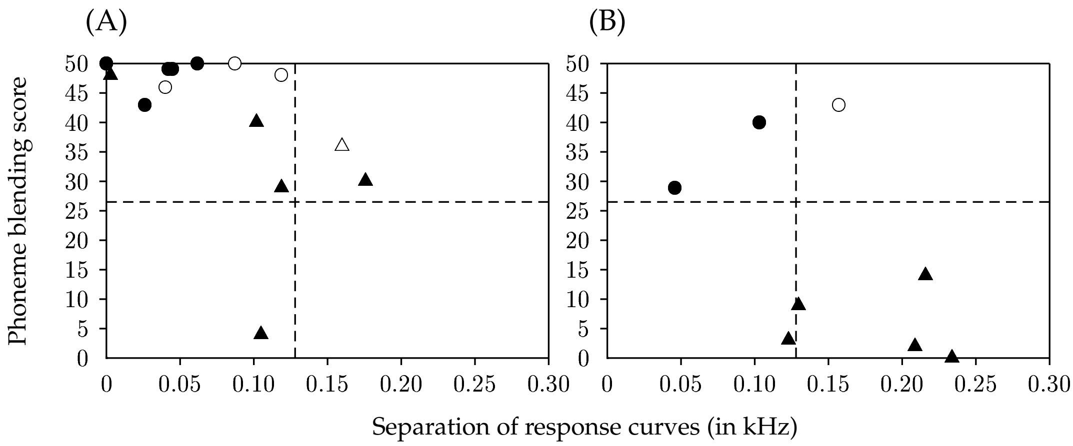

Figure 6.4: **Results of Experiment 1, Session 3 (A) and Experiment 2 (B).**

the reverse—changes in acoustic cue weighting taking place before the development of phonemic awareness skills—never occurs. This means that it is highly unlikely that changes in acoustic cue weighting have a causal effect on the development of phonemic awareness. Second, we are able to say that the development of phonemic awareness is *predictive* of later changes in cue weighting strategies. Finally we can say that changes in acoustic cue weighting do not take place in the absence of phonemic awareness development, even in what Nittrouer had previously classified as older children (e.g. Nittrouer & Studdert-Kennedy 1987, Nittrouer 1992). This last result means that not only are changes in cue weighting strategies predicted by phonemic awareness, but they may be *dependent* on the development of phonemic awareness. Additionally, as noted earlier in this chapter, this last result tells us that changes in cue weighting are not in fact maturational, and suggests that a revision of both the name and the definition of Nittrouer's Developmental Weighting Shift model may be in order.

All of these results and analyses allow us to fairly conclusively answer the question, posed in Chapter 2: **What is the nature and direction of the relationship between changes in acoustic cue weighting and the development of phonemic awareness?**. Two alternative answers to this question were proposed in Chapter 2: (i) Changes in acoustic cue weighting will always occur before the development of phonemic awareness, and shifts in cue weighting will predict later ability in phonemic awareness; and (ii) Phonemic awareness will always develop before shifts in acoustic cue weighting take place, and ability in phonemic awareness will predict later shifts in acoustic cue weighting.

It is clear from the results of these two studies that hypothesis (ii) is the correct answer to the question regarding the nature of the relationship—**changes in acoustic cue weighting are dependent upon (among other things) the development of phonemic awareness skills**.

5 Implications

The rest of this chapter will discuss the implications of the results of the current studies on Nittrouer's Developmental Weighting Shift model, and on our understanding of perceptual development in general.

Having shown the direction of the relationship between perception and awareness, we should now explore the *nature* of this relationship in more depth. In particular, we need, at this point, to reconcile the results of this study, which show that a certain aspect of perception is non–maturational and is influenced by the development of metalinguistic functions, with other studies which have shown the opposite: that perception predicts the development of phonemic awareness. Specifically, the results of this study must be explained in light of the results of the studies carried out by McBride–Chang and colleagues (McBride-Chang 1995*b*, McBride-Chang 1996, Manis et al. 1997, McBride-Chang et al. 1997) on the relationship between phonemic awareness and speech perception. McBride-Chang et al. (1997) found that speech perception, cognitive ability and verbal short term memory together predicted 42% of the variability in phonemic awareness at a later date. How is it possible for phonemic awareness both to predict and to be predicted by speech perception? There could, of course, be a complex feed–back relationship between the two which would explain the results of all of these studies. However, the results of the current study suggest a more straightforward explanation.

As noted in Chapter 2 (Section 3.3), McBride–Chang and colleagues used categorical perception testing in their studies, and measured success on these tests in terms of the 'categorical–ness' or *slope* of the response curves. Both the current study and Nittrouer's (1996*b*) study, on the other hand, examined the correlation between perception and awareness in terms of the *separation* of the perceptual response curves (a measure not available to McBride–Chang and colleagues because of the more conventional method of testing that they used). It does not take a huge leap of the imagination to conclude that the differences seen in these two sets of studies in the relationship between perception and awareness might have something to do with these two different measures of perception.

There are a number of things to note about these two measures, some of which have already been highlighted earlier in this chapter. First, the current studies have shown that slope and separation are not actually very well correlated—the correlation between the two only reached significance at one session in the longitudinal study in Experiment 1; at all other sessions and for the adults and reading–training–delayed children, the relationship fell well below significance.

This in itself would seem to indicate that slope and separation are actually measures of two different aspects of perception. Support for this comes from a comparison of the results of Experiment 1 and Experiment 2. The reading–training–delayed children and the beginning–reading children (from Session 3 of Experiment 1) differed in phonemic awareness ability, but not in age. These two groups were found to differ in degree of separation between response curves, but *not* in slope of response curves. As noted earlier in this chapter, this suggests that while the separation of the response curves is related to literacy (or more probably phonemic awareness) the slope of the response curves is more likely to be a maturational aspect of perception.

All of this could go some way to explaining the differences between the results of this study and those of McBride-Chang et al.'s (1997) study. It is quite possible that before a child can even begin to embark on the process of uncovering phonological structure at a conscious level, their perception (in a more general sense) might have to be reasonably 'mature', and well organised in terms of general phonological structure. However, when the child does embark on the process of becoming phonemically aware, it is possible that the *strategies* that the child uses to perceive phonological structure (i.e. in terms of cue weighting, etc.) might have to change in order to meet the more specific demands of consciously accessing the phonemes in the way that is necessary for alphabetic literacy.

The fact that both slope and separation were measured in this study, as well as phonemic awareness, means that this hypothesis can be empirically tested to a certain extent. We have already seen that the separation of response curves is predicted by earlier abilities on certain phonemic awareness tests, and that separation does not predict later ability in phonemic awareness. What is very interesting to note at this point is that the *reverse* is true of the slope of the response curves: 89% of the variation in measures of phonemic awareness at Session 3 of the longitudinal study (measured this time in terms of phoneme segmentation ability) can be accounted for by the combination of segmentation ability at Session 2, *and* by the slope of response curves at Sessions 1 and 2 $[R^2 = .89(F = 30.48), p < .0001]$. Additionally, slope of response curves at Session 3 are not predicted at all by any earlier measures of phonemic awareness. It is not entirely clear why phoneme *blending* should be the better predictor of acoustic cue weighting strategies while the slope of perceptual response curves predicts phoneme *segmentation* skills, however it is fairly clear that in terms of causal direction, the aspect of speech perception that is measured by the *slope* of a listener's response curves may have some developmental influence on that listener's ability to become aware of phonemes.

These findings go a long way to reconciling the results of this study with the apparently contradictory results of McBride–Chang and colleagues' studies. Additionally, in terms of answering the question posed at the beginning of Chapter 2 about the nature of the relationship, it is clear that we cannot simply say either that **perception impacts on awareness** or that **awareness impacts on perception**, in any global sense—the relationship is more complicated than this. What we can say is that there appears to be both a maturational aspect of speech perception which may have an influence on phonemic awareness skills, and an aspect of perception which has to do with changes in acoustic cue weighting and which only occurs under the influence of phonemic awareness development.

5.2 The Developmental Weighting Shift model

We now turn to an examination of the implications of the results of this thesis on Nittrouer's Developmental Weighting Shift model itself.

There are a number of aspects to Nittrouer's model, all of which have been discussed in detail in Chapter 2. However, to briefly re–cap, they are: i) that a shift occurs between childhood and adulthood in the cues which are weighted most heavily by listeners—specifically, that children give more perceptual weight to transitional cues, while adults give more weight to non–transitional cues—ii) that this shift is related to a more general movement from perception of speech in terms of larger, syllable–sized units in children, to perception of speech in terms of smaller, phoneme–sized units in adults, and iii) that this whole process is due to the increased experience with language that occurs due to maturation. Each of these aspects will be addressed in turn.

First, this study has shown that the phenomenon of shifts in cue weighting is replicable in a number of different situations: for completely synthetic ('copy–synthesised') speech stimuli, and for a new vowel context (/o/). However, this study also showed that these shifts are only replicable for groups of alphabetically literate adults and becoming–literate children. Shifts in cue weighting do not occur for children who are not learning to read, and more specifically, not becoming phonemically aware.

As noted above, Nittrouer and colleagues explain the occurrence of shifts in cue weighting by relating them to theories that children perceive more globally, in terms of larger units, and adults perceive more analytically, in terms of smaller units. Nittrouer argues that if children perceive in terms of units the size of a syllable, then it makes sense for them to attend most to the transitional cues, as these cues give them the most information about the syllable structure as a whole. Similarly, she also claims that if adults do indeed perceive in terms of more phoneme–like units than it makes sense for them to attend most to whatever cue gives them the most information about single segments. Nittrouer's (1996*b*) study found that the children with more global perceptual strategies were also the children who had not yet developed awareness of phonemes (i.e. only had awareness of larger units), while those children with more analytical perceptual strategies were the same children who had developed phoneme awareness. This finding served to reinforce Nittrouer's view that changes in perceptual strategy are due to changes in perceptual unit.

The findings of the studies in this thesis lend further support to this view, while also elaborating on the nature of the change in cue weighting. The cue weighting strategies observed in the responses of all the listeners in both Experiment 1 and Experiment 2 were clearly related to these same subjects' development (or lack) of phonemic awareness. In Experiment 1, perceptual strategy at the end of the study was predicted by phonemic awareness ability at earlier sessions in the study, while in Experiment 2, the reading–training–delayed subjects who had not yet developed phonemic awareness, also had very global perceptual weighting strategies compared to their beginning–reading peers. From these results we can conclude that not only are changes in perceptual strategy related to the development of phonemic awareness, but that the change in metaphonological ability from awareness of larger units (syllables, onset–rime units) to awareness of smaller units (phonemes) actually has an impact on perceptual strategy. This would appear to support Nittrouer's theory that the changes in perceptual strategy that she observed were something to do with a change in phonological units. However, it also casts some doubt on Nittrouer's third claim regarding the DWS.

This last claim, that shifts in acoustic cue weighting are maturational, is particularly problematic for the results of Experiment 2. Both groups of children in this study were chronologically the same age (6–7 years), and thus under Nittrouer's hypothesis should be displaying the same perceptual strategies. More specifically, according to the results found by Nittrouer and colleagues in other cue weighting studies, both groups of 6– to 7–year–olds should display predominantly analytical rather than global perceptual strategies. The results of Experiment 2, however, clearly show that the reading–training–delayed, predominantly phonemically unaware group displays *global* rather than analytical perceptual strategies. As noted above, the conclusion that must be drawn from these results is that Nittrouer's Developmental Weighting Shift model is not in fact a model of a maturational process. It appears, instead, that changes in acoustic cue weighting require some sort of catalyst, which, in the case of these experiments, is alphabetic literacy instruction, and more specifically the onset of phonemic awareness that occurs as a result of this instruction. Without such a catalyst, shifts in cue weighting do not appear to occur.

Clearly then, Nittrouer's Developmental Weighting Shift model is reasonably robust—at least for /ʃ/–/s/ contrasts—and holds up well to variations in stimulus design. Additionally, it does appear from the predictive relationship found between cue weighting and phonemic awareness development that Nittrouer's hypothesis regarding the DWS as a reflection of changes in phonological unit may indeed be valid. However, the claim that the DWS model describes a maturational process is strongly refuted by the findings of these studies. This last point has potential knock-on effects for more general theories of perception, and possibly of phonological development in general, as will be seen in the following section.

5.3 Perception: development, testing and units

At the very end of the first chapter of this thesis, a number of very broad issues were raised concerning the way in which we currently view speech perception, and in particular perceptual development. As these issues were raised, it was stated that addressing them would be outside the scope of this thesis. It was also stated, however, that these issues would form the framework within which the thesis would be constructed, and as such, the results of the thesis should have some, if not direct, then more general implications for these issues.

The first of these issues revolves around the view of speech perception as a 'black box'—i.e. a largely unspecified system which takes in the acoustic signal, and which outputs some sort of phonological pattern—upon which the development of the rest of speech and language communication depends. This view can be seen, for example, in the studies of perception and literacy discussed in

Chapter 2, Section 3.1, in which perception was almost always considered to be a process which allowed for the later development of literacy skills, among other things. This view depends on a number of assumptions about the way in which perception operates. First, speech perception must be assumed to be maturational—thus the speech perception system of the average adult experimental subject will be representative of the way in which a fully functioning and 'mature' perceptual system should work. Second, it must be assumed that speech perception cannot be affected by the development of higher cognitive processes, and certainly not by the development of non–necessary processes like alphabetic literacy and phonemic awareness. It is quite clear from the results of this study, however, that while some aspects of perceptual development are very likely to be maturational, not *all* perceptual changes are maturational. Furthermore, this study shows that at least one aspect of perceptual change is not only affected by, but possibly induced by the development of a non–universal and not directly innate part of language.

The finding that different aspects of perception appear to be influenced to different extents both by maturation and by the development of other cognitive processes, brings us to the second point to be discussed in this section. This issue concerns the view that speech perception is a unitary construct which develops and operates as a homogeneous system. Many studies, in particular those that do not directly investigate perception itself, tend to treat speech perception as a single cognitive process—thus ability on a categorical perception test, for example, is seen as a reflection of speech perception ability as a whole. This view can again be seen in many of the studies of literacy and perception discussed in Chapter 2, Section 3.1.

Again, the lack of correlation or predictive relationship in this study between the two measures of speech perception—slope and separation of response curves—points to a need to consider speech perception less as a single, undifferentiated process, and more as a complex of abilities. This is not the first time that such a suggestion has been made—McBride-Chang et al. (1997) states that one "issue of particular concern for future studies is the idea of speech perception as a global construct" (p. 629)—however, the evidence from the current study provides convincing evidence for this more heterogeneous view of perception. Furthermore, while the results of the study do not allow us to make conclusive claims about the exact nature of such a complex, it would seem appropriate to consider that some of the changes in speech perception observed in studies to this point reflect a general maturation of overall perceptual *ability* (such as 'categorical–ness' of responses), while others reflect changes in perceptual *strategy* (such as shifts in acoustic cue weighting).

We now turn to the issue of the phoneme as the fundamental unit of perception. The results of this thesis raise some questions regarding the strength of this claim. As noted in Chapter 1, most theories of speech perception, whether acoustic or gestural, propose the phoneme or phonemic feature as the basic unit of a normal perceptual system. Under these theories, speech is made up of *phonemic* units which are joined together by *coarticulation*. Although the acoustic theorists and the gesturalists disagree as to the purpose of coarticulation—to the acoustic theorists it is unnecessary 'noise' which occurs at phoneme boundaries, while to gesturalists it is a fundamental consequence of the rapid sequential articulation of gestures which helps to inform the perceptual system—for both groups it is the phoneme that is perceived. There is in fact some evidence (from studies of speech errors and speech deficits, for example) that phonemes do exist at some perceptual level.

Other researchers, including Nittrouer and colleagues, have gone on to propose that perceptual systems, and other systems which require phonological organisation, do not start out with phonemes as the basic unit, but instead start out with a unit which is much larger (i.e. a syllable) and shift over the course of childhood to the smaller, phonemic unit. The finding of a relationship between acoustic cue weighting and the movement from syllable and onset–rime awareness to phoneme awareness in Nittrouer's (1996*b*) study appears to support this theory. The evidence from this thesis goes on to provide more experimental support for the existence of a relationship between perceptual cue weighting and phonemic awareness, and thus possibly for a change in unit at the perceptual level. However, the results of this thesis may also prove problematic for the theory that phonemes are the *default* endpoint of perceptual maturation.

In order to understand the potential problems caused by the results of this thesis, we must first accept Nittrouer's proposal that shifts in cue weighting reflect changes in the unit of perceptual organisation from a syllable to a phoneme. We must then go on to accept that the relationship between shifts in cue weighting and the development of phonemic awareness is due to the fact that *both* are moving from syllable to phoneme organisation. Given both of these premises, and given that the development of phonemic awareness is not necessary for successful communication, then the findings of this thesis would seem to indicate that the shift to phoneme–based organisation in perception is *also* not a necessary step for successful communication. Simply put, if shifts in perceptual organisation are caused by phonemic awareness, and phonemic awareness is a higher cognitive skill which is not developed by a large proportion of people (i.e. people who have not had access to literacy training, or some other relevant training), then shifts to phoneme–based perceptual strategies should also not occur in a large proportion of people. This calls into question the *fundamental* role of phonemes in perception that is assumed by theorists, and also the role of coarticulatory cues as secondary in perception. Under this alternative hypothesis, the results of the studies of speech errors, etc. referred to above would be due, not to the subjects' mature perceptual systems, but rather to the fact that they were probably alphabetic literates.

The evidence provided by this thesis is not sufficient to support a claim that phonemes exist simply as a results of alphabetic literacy. In fact, it would be unwise to make such a claim at all, given the fact that in order for the alphabet itself to have been invented, phonemes must exist at some level of organisation. However, beyond a better understanding of the nature of the relationship between shifts in cue weighting and the development of phonemic awareness, what *should* be taken away from this thesis is the fact that higher cognitive processes like literacy and metalinguistic awareness *do* have an influence on more fundamental processes like perception. The results of both Experiment 1 and Experiment 2 show that alphabetic literacy, and the phonemic awareness that occurs as a result of literacy development, have an effect on perceptual strategies. What this means is that the perceptual behaviour of the 'average adult subject' referred to above should probably *not* be taken as representative of perceptual ability in a strictly mature state. Rather, this type of subject will have perceptual skills which are the result of *both* maturation *and* the effects of phonemic awareness. This should also be borne in mind in the testing of child subjects, most of whom will have been brought up within literate environments, and will be attending schools where literacy is taught from an early age.

The fact that perceptual behaviour has been shown to be affected by the development of higher cognitive skills means that we cannot assume that the perceptual abilities of most experimental subjects (who are generally selected from highly literate populations) have been shaped simply by maturational processes. By the same token, future research must seek to determine to what extent higher order knowledge like literacy impinges on *other* lower level linguistic behaviour. If the true course of development of speech perception, and other linguistic systems, is to be effectively mapped, then literacy and possibly other higher cognitive processes, *must* be taken into account.

Appendix A

Perception test materials

1 Text of story presented before perceptual testing:

This story is about a boy just about your age, whose name is Callum. This story is also about Callum's teddy bear, whose name is Mr Bear. Callum and Mr Bear go everywhere together. One day Callum noticed that Mr Bear had a hole in his tummy! His stuffing was starting to escape! "What am I going to do?" thought Callum. "If all Mr Bear's stuffing escapes, he'll go all flat and floppy!"

Callum decided to go to the kitchen to show Mr Bear to his dad. "Can you sew him for me?" Callum asked his dad. But dad was too busy making the tea.

So, Callum went upstairs to show Mr Bear to his mum. "Can you sew him for me?" Callum asked his mum. But mum was too busy doing the ironing.

So, Callum went back downstairs to show Mr Bear to his big sister. "Can you sew him for me?" Callum asked his sister. But his sister was too busy talking on the phone.

Just then, the doorbell rang. It was Callum's friend Mairi. "What's the matter Callum?" Mairi asked, "You look sad. Mr Bear doesn't look too happy either." "He isn't happy," said Callum, "He's got a hole in his tummy, and his stuffing is starting to escape, and I tried to show him to everyone, but they're all too busy to sew him up." "I know what to do," said Mairi, "Let's show Mr Bear to my dad. He's a tailor and he knows how to sew lots of things, even teddy bears."

So Mairi and Callum took Mr Bear to show Mairi's dad, and right enough, Mairi's dad knew just what to do to sew Mr Bear back together.

2 Pictures used in perceptual testing

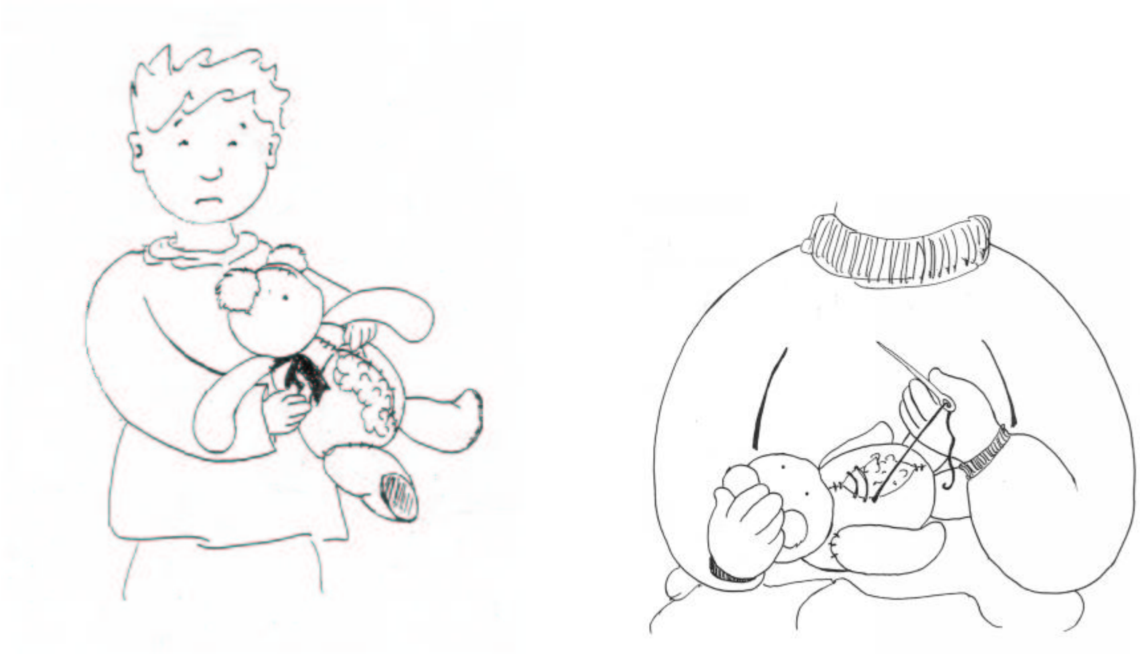

Figure A.1: **Pictures of "show" and "sew"** These line drawings accompanied the above story, and were used in the perceptual tests for the children.

Appendix B

Phonemic awareness test materials

Phoneme blending

Training:	cow–boy
	snow–man
	sun–shine
	sh–eep
	c–ow
	s–ing
	b–a–th
	p–e–n
	f–a–ce
	p–i–g
Pre-test:	m–ilk
	f–ish
	b–ear
	j–uice
	d–oll

Phoneme blending

Test (a): d–o–g
c–u–p
r–u–n
f–i–sh
r–ai–n
sh–o–p
m–a–n
c–a–tch
l–o–ve
c–a–t
f–r–o–g
s–n–a–ke
l–a–m–p
c–l–o–ck
t–a–s–te
f–l–a–g
s–t–u–ck
t–oa–s–t
s–t–o–p
n–e–s–t
f–r–ie–n–d
s–l–e–p–t
t–r–u–n–k
g–r–ou–n–d
s–t–r–ee–t

Phoneme blending

Test (b): m–oo–n
d–u–ck
l–ea–f
g–a–me
l–e–g
t–e–n
s–o–ck
s–i–t
t–i–me
r–oo–m
l–e–f–t
f–e–n–ce
t–e–n–t
l–a–s–t
d–r–u–m
j–u–m–p
g–l–a–ss
s–t–i–ck
d–r–o–p
s–a–n–d
s–t–a–n–d
p–l–a–n–t
s–t–a–m–p
s–c–r–a–tch
s–p–l–a–sh

Phoneme segmentation

Training:	cowboy
	snowman
	sunshine
	cow
	man
	sun
Pre-test:	tea
	zoo
	toy
	shoe
	day

Phoneme segmentation

Test (a): red
cake
peck
book
cheese
phone
shut
comb
soup
mess
stone
clap
bump
vest
snail
chest
soft
space
green
hand
skunk
crunch
trust
stretch
blast

Phoneme segmentation

Test (b): nose
make
laugh
bus
mouse
cage
bed
sun
name
can
post
block
crab
dust
spoon
desk
glove
fast
belt
step
spend
crisp
scrub
screen
strap

Phoneme deletion

Training:	cat
	bus
	hat
	seat
	glass
	Scott
Pre-test:	leg
	cup
	sit
	farm
	sand

Phoneme deletion

Test (a):	snow
	ground
	slow
	break
	snap
	bread
	blow
	crib
	great
	drag
	bright
	price
	clean
	clock
	slid
	brush
	plate
	grow
	glove
	play

Phoneme deletion

Test (b): cloud
snail
broom
sleep
sled
clap
fly
place
spot
store
switch
crash
spill
smile
clip
bring
block
draw
black
slip

APPENDIX C

Raw data

Note that the following abbreviations are used in the tables in this section:

Slp. /s/:	slope of /s/–transition response curves
Slp. /ʃ/:	slope of /ʃ/–transition response curves
Mn. /s/:	mean of /s/–transition response curves
Mn. /ʃ/:	mean of /ʃ/–transition response curves
Avg. slp.:	average slope of two response curves
Sep. 1:	separation of response curves in absolute difference between means
Sep. 2:	separation of response curves in kHz
Bln.:	phoneme blending score (/50)
Seg.:	phoneme segmentation score (/50)
Del.:	phoneme deletion score (/40)
Schonell:	raw score on Schonell's Graded Word Reading Test
BPVS:	raw score on British Picture Vocabulary Scale
RA:	reading age (from Schonell)
AE:	age equivalent (from BPVS)

Beginning–reading children: **Stage 1**

Subject	Slp. /s/	Mn. /s/	Slp. /ʃ/	Mn. /ʃ/	Avg. slp.	Sep. 1	Sep. 2	Bln.	Seg.	Del.	Schonell	RA	BPVS	AE
R1	-1.86	4.40	-2.02	4.01	-1.94	0.38550	0.077100	50	44	18	21	7;4	14	6;3
R2	-0.53	5.13	-0.58	4.00	-0.55	1.12296	0.224592	49	40	17	13	6;11	13	5;9
R3	-1.00	3.75	-0.84	3.78	-0.92	-0.27190	0.054380	48	49	37	36	8;6	13	5;9
R4	-1.75	4.71	-1.94	4.20	-1.84	0.50874	0.101748	45	30	4	8	6;7	18	8;3
R5	-1.66	3.85	-1.74	3.66	-1.70	0.18628	0.037256	39	27	28	4	6;4	22	10;11
R6	-0.60	4.01	-0.58	3.40	-0.59	0.61042	0.122084	37	36	24	6	6;6	12	5;4
R7	-0.58	4.96	-0.75	3.89	-0.66	1.07262	0.214524	38	23	0	9	6;8	23	11;6
R8	-1.78	4.61	-1.71	3.89	-1.75	0.72460	0.144920	26	25	21	14	6;11	12	5;4
R9	-0.61	5.50	-1.94	4.20	-1.27	1.30368	0.260736	27	30	0	26	7;7	17	7;9
R10	-1.20	5.19	-1.66	4.89	-1.43	0.29948	0.059896	30	25	0	4	6;4	13	5;9
R11	-1.68	4.90	-0.77	4.26	-1.23	0.63888	0.127776	14	3	1	1	6;0-	10	4;5
R12	-0.97	4.57	-1.38	4.18	-1.17	0.39108	0.078216	15	20	29	1	6;0-	22	10;11
R13	-1.26	5.25	-1.78	4.22	-1.52	1.03516	0.207032	6	13	0	0	6;0-	15	6;8
R14	-1.32	4.66	-0.47	3.69	-0.90	0.96839	0.193678	16	6	0	3	6;2	16	7;2
R15	-0.44	4.07	-0.53	3.43	-0.48	0.64917	0.129834	0	0	0	0	6;0-	11	4;10
R16	-0.54	4.80	-0.48	4.50	-0.51	0.30435	0.060870	38	37	30	7	6;7	12	5;4
R17	-2.07	3.91	-1.93	3.30	-2.00	0.61219	0.122438	27	34	2	9	6;8	9	4;0
R18	-0.51	5.58	-0.81	4.70	-0.66	0.88437	0.176874	6	4	0	1	6;0-	10	4;5

Table C.1: Raw data

Beginning–reading children: **Stage 2**

Subject	Slp. /s/	Mn. /s/	Slp. /ʃ/	Mn. /ʃ/	Avg. slp.	Sep. 1	Sep. 2	Bln.	Seg.	Del.	Schonell	RA	BPVS	AE
R1	-2.02	4.01	-2.02	4.01	-2.02	0	0	50	46	40	n/a	n/a	n/a	n/a
R2	-1.20	3.81	-1.07	3.34	-1.14	0.46389	0.092778	49	42	38	n/a	n/a	n/a	n/a
R3	-1.33	4.46	-2.12	3.82	-1.72	0.64536	0.129072	50	49	40	n/a	n/a	n/a	n/a
R4	-1.11	4.84	-1.86	4.40	-1.48	0.44464	0.088928	50	34	6	n/a	n/a	n/a	n/a
R5	-1.11	4.22	-1.67	3.98	-1.39	0.24321	0.048642	48	41	39	n/a	n/a	n/a	n/a
R6	-0.60	3.47	-0.81	2.99	-0.71	0.47579	0.095158	46	49	39	n/a	n/a	n/a	n/a
R7	-0.56	4.75	-0.63	3.85	-0.59	0.90084	0.180168	46	25	25	n/a	n/a	n/a	n/a
R8	-1.87	4.31	-1.36	4.09	-1.61	0.22525	0.045050	36	42	39	n/a	n/a	n/a	n/a
R9	-1.00	4.89	-1.65	4.07	-1.33	0.81968	0.163936	33	36	36	n/a	n/a	n/a	n/a
R10	-1.08	5.23	-1.62	5.04	-1.35	0.19624	0.039248	44	40	0	n/a	n/a	n/a	n/a
R11	-0.80	5.23	-1.68	4.90	-1.24	0.32845	0.065690	38	42	2	n/a	n/a	n/a	n/a
R12	-1.45	4.65	-0.97	3.68	-1.21	0.96657	0.193314	34	43	39	n/a	n/a	n/a	n/a
R13	-1.75	4.71	-0.64	4.19	-1.19	0.51757	0.103514	33	27	4	n/a	n/a	n/a	n/a
R14	-1.39	4.88	-1.38	4.18	-1.38	0.69776	0.139552	39	28	0	n/a	n/a	n/a	n/a
R15	-0.92	5.15	-1.35	4.00	-1.13	1.14786	0.229572	4	0	0	n/a	n/a	n/a	n/a
R16	-1.85	4.41	-0.63	3.98	-1.24	0.43017	0.086034	46	41	40	n/a	n/a	n/a	n/a
R17	-1.09	4.87	-0.93	4.41	-1.01	0.46026	0.092052	33	43	38	n/a	n/a	n/a	n/a
R18	-0.96	6.04	-0.93	4.74	-0.95	1.29377	0.258754	9	15	0	n/a	n/a	n/a	n/a

Table C.2: Raw data

Beginning–reading children: **Stage 3**

Subject	Slp. /s/	Mn. /s/	Slp. /ʃ/	Mn. /ʃ/	Avg. slp.	Sep. 1	Sep. 2	Bln.	Seg.	Del.	Schonell	RA	BPVS	AE
R1	-1.94	4.20	-1.94	4.20	-1.94	0	0	50	47	40	33	8;3	14	6;3
R2	-1.08	4.20	-0.56	3.77	-0.82	0.43513	0.087026	50	50	39	21	7;4	16	7;2
R3	-1.61	4.61	-1.90	4.30	-1.75	0.30888	0.061776	50	50	40	37	8;6	17	7;9
R4	-1.78	4.61	-1.78	4.61	-1.78	0	0	50	48	40	11	6;10	21	10;2
R5	-1.78	4.60	-1.86	4.40	-1.82	0.20862	0.041724	49	44	38	8	6;7	24	12;2
R6	-0.70	3.66	-0.76	3.43	-0.73	0.22476	0.044952	49	50	40	33	8;3	14	6;3
R7	-0.91	4.82	-0.48	4.22	-0.70	0.59477	0.118954	48	36	38	18	7;2	24	12;2
R8	-0.95	4.91	-0.55	4.89	-0.75	0.01459	0.002918	48	46	39	28	7;9	14	6;3
R9	-0.70	4.40	-1.94	4.20	-1.32	0.19784	0.039568	46	46	40	30	8;0	16	7;2
R10	-0.72	4.70	-0.85	4.57	-0.79	0.12847	0.025694	43	43	33	12	6;10	10	4;5
R11	-1.54	5.50	-1.62	4.99	-1.58	0.51216	0.102432	40	47	3	7	6;7	12	5;4
R12	-1.68	4.90	-1.98	4.10	-1.83	0.79868	0.159736	36	49	39	5	6;5	19	8;11
R13	-0.68	5.42	-1.32	4.54	-1.00	0.88040	0.176080	30	43	38	2	6;0	16	7;2
R14	-0.73	5.30	-1.01	4.70	-0.87	0.59514	0.119028	29	25	0	4	6;4	17	7;9
R15	-1.35	5.09	-1.02	4.57	-1.19	0.52817	0.105634	4	2	0	7	6;7	14	6;3
R16	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
R17	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
R18	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a

Table C.3: Raw data

Reading–training–delayed children

Subject	Slp. /s/	Mn. /s/	Slp. /ʃ/	Mn. /ʃ/	Avg. slp.	Sep. 1	Sep. 2	Bln.	Seg.	Del.	Schonell	RA	BPVS	AE
N1	-1.10	5.29	-1.46	4.51	-1.28	0.78357	0.156714	43	40	39	59	10;2	21	10;2
N2	-1.49	5.32	-1.63	5.09	-1.56	0.22895	0.045790	29	35	0	1	6;0-	17	7;9
N3	-1.19	5.03	-0.66	3.98	-0.93	1.04454	0.208908	2	0	0	0	6;0-	16	7;2
N4	-0.75	4.61	-0.55	4.00	-0.65	0.61666	0.123332	3	0	0	0	6;0-	16	7;2
N5	-1.71	4.80	-0.67	3.72	-1.19	1.08143	0.216286	14	1	0	0	6;0-	16	7;2
N6	-0.94	5.00	-0.86	3.84	-0.90	1.16851	0.233702	0	0	0	0	6;0-	14	6;3
N7	-1.62	5.13	-1.78	4.61	-1.70	0.51457	0.102914	40	31	40	25	7;7	13	5;9
N8	-0.66	5.28	-0.89	4.65	-0.78	0.63672	0.127344	9	6	1	7	6;7	19	8;11

Adults

Subject	Slp. /s/	Mn. /s/	Slp. /ʃ/	Mn. /ʃ/	Avg. slp.	Sep. 1	Sep. 2	Bln.	Seg.	Del.	Schonell	RA	BPVS	AE
A1	-1.94	4.20	-1.98	4.10	-1.96	0.09588	0.019176	n/a	n/a	n/a	n/a	n/a	n/a	n/a
A2	-1.71	4.81	-1.61	4.61	-1.66	0.20108	0.040216	n/a	n/a	n/a	n/a	n/a	n/a	n/a
A3	-1.19	5.75	-1.63	5.09	-1.41	0.66249	0.132498	n/a	n/a	n/a	n/a	n/a	n/a	n/a
A4	-1.23	4.72	-1.51	4.70	-1.37	0.02000	0.004000	n/a	n/a	n/a	n/a	n/a	n/a	n/a
A5	-1.37	5.18	-1.58	5.30	-1.47	-0.11746	0.023492	n/a	n/a	n/a	n/a	n/a	n/a	n/a
A6	-1.08	4.74	-1.78	4.22	-1.43	0.51812	0.103624	n/a	n/a	n/a	n/a	n/a	n/a	n/a
A7	-1.71	4.81	-1.85	4.41	-1.78	0.39712	0.079424	n/a	n/a	n/a	n/a	n/a	n/a	n/a
A8	-1.50	5.69	-1.58	5.27	-1.54	0.41393	0.082786	n/a	n/a	n/a	n/a	n/a	n/a	n/a

Table C.4: Raw data

References

Adlard, A. & Hazan, V. (1998), 'Speech perception in children with specific reading difficulties (dyslexia)', *Quarterly Journal of Experimental Psychology* **51A**(1), 153–177.

Bertelson, P. & de Gelder, B. (1991), The emergence of phonological awareness: Comparative approaches, *in* I. G. Mattingly & M. Studdert-Kennedy, eds, 'Modularity and the Motor Theory of Speech Perception: Proceedings of a Conferance to Honor A. M. Liberman', Lawrence Erlbaum, Hillsdale, NJ, pp. 393–411.

Bertelson, P., de Gelder, B., Tfouni, L. V. & Morais, J. (1989), 'Metaphonological abilities of adult literates: New evidence of heterogeneity.', *European Journal of Cognitive Psychology* **1**, 239–250.

Bertoncini, J., Bjeljac-Babic, R., Jusczyk, P. W., Kennedy, L. J. & Mehler, J. (1988), 'An investigation of young infants' perceptual representations of speech sounds', *Journal of Experimental Psychology* **117**, 21–33.

Bertoncini, J., Floccia, C., Nazzi, T. & Mehler, J. (1995), 'Morae and syllables: Rhythmical basis of speech representations in neonates', *Language and Speech* **38**(4), 311–329.

Bertoncini, J. & Mehler, J. (1981), 'Syllables as units in infant speech perception', *Infant Behavior and Development* **4**, 247–260.

Best, C. T., Morrongiello, B. & Robson, R. (1981), 'Perceptual equivalence of acoustic cues in speech and non–speech perception', *Perception & Psychophysics* **29**(3), 191–211.

Bjeljac-Babic, R., Bertoncini, J. & Mehler, J. (1993), 'How do 4–day–old infants categorize multisyllabic utterances?', *Developmental Psychology* **29**, 711–721.

Blumstein, S. E. (1981), Phonological aspects of aphasia, *in* M. T. Sarno, ed., 'Acquired Aphasia', Academic Press, New York, pp. 129–155.

Blumstein, S. E. & Stevens, K. N. (1980), 'Perceptual invariance and onset spectra for stop consonants in different vowel environments', *Journal of the Acoustical Society of America* **67**(2), 648–662.

Bradley, L. & Bryant, P. E. (1983), 'Categorising sounds and learning to read: A causal connection', *Nature* **301**, 419–421.

Brady, S., Shankweiler, D. & Mann, V. (1983), 'Speech perception and memory coding in relation to reading ability', *Journal of Experimental Child Psychology* **35**, 345–367.

Bryant, P. (1998), 'Sensitivity to onset and rhyme does predict young children's reading: A comment on Muter, Hulme, Snowling and Taylor (1997)', *Journal of Experimental Child Psychology* **71**, 29–37.

Bryant, P. E., MacLean, M., Bradley, L. & Crossland, J. (1990), 'Rhyme, alliteration, phoneme detection and learning to read', *Developmental Psychology* **26**(3), 1–10.

Campbell, R. & Sais, E. (1995), 'Accelerated metalinguistic (phonological) awareness in bilingual children', *British Journal of Developmental Psychology* **13**, 61–68.

Carlisle, J. F. (1991), Questioning the psychological reality of onset–rime as a level of phonological awareness, *in* S. S. Brady & D. P. Shankweiler, eds, 'Phonological Processes in Literacy: A Tribute to Isabelle Y. Liberman', LEA, Hillsdale, NJ, pp. 85–95.

Christensen, L. A. & Humes, L. E. (1997), 'Identification of multidimensional stimuli containing speech cues and the effects of training', *Journal of the Acoustical Society of America* **102**(4), 2297–2310.

Cohen, J. & Cohen, P. (1983), *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences*, second edn, LEA, Hillsdale, NJ.

Conrad, R. (1971), 'The chronology of the development of covert speech in children', *Developmental Psychology* **5**, 398–505.

Cowan, N., Leavitt, L. A., Massaro, D. W. & Kent, R. D. (1982), 'A fluent backward talker', *Journal of Speech and Hearing Research* **25**, 48–53.

Cutler, A. (1996), Prosody and the word boundary problem, *in* J. L. Morgan & K. Demuth, eds, 'Signal to Syntax: Bootstrapping from Speech to Grammar in Early Acquisition', Lawrence Erlbaum, Mahwah, NJ, pp. 87–99.

Daniloff, R. & Hammarberg, R. (1973), 'On defining coarticulation', *Journal of Phonetics* **1**, 239–248.

Davis, B. L. & MacNeilage, P. F. (1995), 'The articulatory basis of babbling', *Journal of Speech and Hearing Research* **38**(6), 1199–1211.

de Gelder, B. & Vroomen, J. (1992), Auditory and visual speech perception in alphabetic and non–alphabetic Chinese–Dutch bilinguals, *in* R. J. Harris, ed., 'Cognitive Processing in Bilinguals', Elsevier Science Publishers B. V., pp. 413–426.

de Gelder, B. & Vroomen, J. (1998), 'Impaired speech perception in poor readers: Evidence from hearing and speech reading', *Brain and Language* **64**, 269–281.

Delattre, P. C., Liberman, A. L. & Cooper, F. S. (1955), 'Acoustic loci and transitional cues for consonants', *Journal of the Acoustical Society of America* **27**(4), 769–773.

Dorman, M. F., Studdert-Kennedy, M. & Raphael, L. J. (1977), 'Stop–consonant recognition: Release bursts and formant transitions as functionally equivalent, context–dependent cues', *Perception & Psychophysics* **22**(2), 109–122.

Dunn, L. M., Dunn, L. M., Whetton, C. & Pintilie, D. (1982), 'British Picture Vocabulary Test'.

Eimas, P. D. (1974), 'Auditory and linguistic processing of cues for place of articulation by infants', *Perception & Psychophysics* **16**, 513–521.

Eimas, P. D. (1975), 'Auditory and phonetic coding of the cues for speech: Discrimination of the [r–l] distinction by young infants', *Perception & Psychophysics* **18**, 341–347.

Eimas, P. D. & Clarkson, R. L. (1986), Speech perception in children: Are there effects of otitis media?, *in* J. F. Kavanagh, ed., 'Otitis Media and Child Development', York Press, Parkton, MD, pp. 139–159.

Eimas, P. D. & Miller, J. L. (1980), 'Contextual effects in infant speech perception', *Science* **209**, 1140–1141.

Eimas, P. D., Siqueland, E. R., Jusczyk, P. W. & Vigorito, J. (1971), 'Speech perception in infants', *Science* **171**, 303–306.

Elliott, L. L., Hammer, M. A., Scholl, M. E. & Wasowicz, J. M. (1989), 'Age differences in discrimination of simulated single–formant frequency transitions', *Perception & Psychophysics* **46**(2), 181–186.

Ferguson, C. A. & Farwell, C. B. (1975), 'Words and sounds in early language acquisition', *Language* **51**, 419–439.

Fitch, H. L., Halwes, T., Erickson, D. M. & Liberman, A. L. (1980), 'Perceptual equivalence of two acoustic cues for stop–consonant manner', *Perception & Psychophysics* **27**(4), 343–350.

Flege, J. E., Walley, A. C. & Randazza, L. (1992), 'A developmental study of native and non–native vowel perception', *Journal of the Acoustical Society of America* **92**(4), 2415.

Fowler, A. E. (1991), How early phonological development might set the stage for phoneme awareness, *in* S. S. Brady & D. P. Shankweiler, eds, 'Phonological Processes in Literacy: A Tribute to Isabelle Y. Liberman', LEA, Hillsdale, NJ, pp. 97–117.

Fowler, C. A. (1986*a*), 'An event approach to the study of speech perception from a direct–realist perspective', *Journal of Phonetics* **14**, 3–28.

Fowler, C. A. (1986*b*), 'Reply to commentators', *Journal of Phonetics* **14**, 149–170.

Fowler, C. A. (1994), 'Invariants, specifiers, cues: An investigation of locus equations as information for place of articulation', *Perception & Psychophysics* **55**(6), 597–610.

Fowler, C. A. & Rosenblum, L. D. (1991), The perception of phonetic gestures, *in* I. G. Mattingly & M. Studdert-Kennedy, eds, 'Modularity and the Motor Theory of Speech Perception: Proceedings of a Conferance to Honor A. M. Liberman', LEA, Hillsdale, NJ, pp. 33–59.

Fox, B. & Routh, D. K. (1975), 'Analysing spoken language into words, syllables, and phonemes: A developmental study', *Journal of Psycholinguistic Research* **4**, 331–342.

Gathercole, S. E. & Baddeley, A. D. (1998), 'The role of phonological memory in vocabulary acquisition: A study of young children learning new names', *British Journal of Psychology* **51A**(1), 153–177.

Godfrey, J. J., Syrdal-Lasky, A. K. amd Millay, K. K. & Knox, C. M. (1981), 'Performance of dyslexic children on speech perception tests', *Journal of Experimental Child Psychology* **32**, 401–424.

Gordon, P. C., Eberhardt, J. L. & Rueckl, J. G. (1993), 'Attentional modulation of the phonetic significance of acoustic cues', *Cognitive Psychology* **25**, 1–42.

Goswami, U. C. & Bryant, P. E. (1990), *Phonological Skills and Learning to Read*, Erlbaum, Hove.

Gravel, J. S. & Wallace, I. F. (1992), 'Listening and language at 4 years of age', *Journal of Speech and Hearing Research* **35**, 588–595.

Greenlee, M. (1980), 'Learning the phonetic cues to the voiced voiceless distinction: A comparison of child and adult speech perception', *Journal of Child Language* **7**, 459–468.

Gussenhoven, C. & Jacobs, H. (1998), *Understanding Phonology*, Arnold Publishers, London.

Hanson, V. L. (1991), Phonological processing without sound, *in* S. S. Brady & D. P. Shankweiler, eds, 'Phonological Processes in Literacy: A Tribute to I. Y. Liberman', LEA, Hillsdale, NJ, pp. 153–161.

Hatcher, P. J. (1994), *Sound Linkage: An Integral Programme For Overcoming Reading Difficulties*, Whurr, London.

Hazan, V. & Barrett, S. (1999), The development of phoneme categorisation in children aged 6 to 12, *in* J. J. Ohala, Y. Hasegawa, M. Ohala, D. Granville & A. C. Bailey, eds, 'Proceedings of the XIVth International Congress of Phonetic Sciences', Vol. 3, San Francisco, pp. 2493–2496.

Hazan, V. & Rosen, S. (1991), 'Individual variability in the perception of cues to place contrasts in initial stops', *Perception & Psychophysics* **49**(2), 187–200.

Howell, J., Hill, A., Dean, E. & Waters, D. (1993), Increasing metalinguistic awareness to assisst phonological change, *in* D. J. Messer & G. J. Turner, eds, 'Critical Influences on Child Language Acquisition and Development', St. Martins, New York, NY.

Hulme, C., Muter, V. & Snowling, M. (1998), 'Segmentation does predict early progress in learning to read better than rhyme: A reply to Bryant', *Journal of Experimental Child Psychology* **71**, 39–44.

Innes, S. (1995), Facilitation of phonological awareness in three and four year old children, Master's thesis, Queen Margaret College.

Jastack, S. & Wilkinson, G. S. (1984), 'Wide Range Achievement Test—Revised', Wilmington, DE: Jastack Associates.

Joanisse, M. F., Manis, F. R., Keating, P. & Seidenberg, M. S. (1998), 'Language deficits in dyslexic children: Speech perception, phonology and morphology', *UCLA Working Papers in Linguistics* **96**, 135–161.

Jusczyk, P. W. (1997), *The Discovery of Spoken Language*, MIT Press, Cambridge, Massachusetts.

Jusczyk, P. W. & Derrah, C. (1987), 'Representation of speech sounds by young infants', *Developmental Psychology* **23**, 648–654.

Kingston, J. & Diehl, R. L. (1995), Intermediate properties in the perception of distinctive feature values, *in* B. Connell & A. Arvaniti, eds, 'Phonology and Phonetic Evidence: Papers in Laboratory Phonology IV', Cambridge University Press.

Kirtley, C., Bryant, P., MacLean, M. & Bradley, L. (1989), 'Rhyme, rime and the onset of reading', *Journal of Experimental Psychology* **48**, 224–245.

Klatt, D. (1980), 'Software for a cascade/parallel formant synthesizer', *Journal of the Acoustical Society of America* **67**, 971–995.

Krause, S. E. (1982), 'Vowel duration as a perceptual cue to postvocalic consonant voicing in young children and adults', *Journal of the Acoustical Society of America* **71**(4), 990–995.

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

Ladefoged, P. (1993), *A Course in Phonetics*, 3rd edn, Harcourt Brace, Fort Worth, TX.

Laver, J. & Trudgill, P. (1979), Phonetic and linguistic markers in speech, *in* K. R. Scherer & H. Giles, eds, 'Social Markers in Speech', Cambridge University Press, pp. 1–32.

Liberman, A. M. (1996), *Speech: a Special Code*, MIT Press.

Liberman, A. M., Cooper, F. S., Shankweiler, D. & Studdert-Kennedy, M. (1967), 'Perception of the speech code', *Psychological Review* **74**, 431–461.

Liberman, A. M., Delattre, P. C. & Cooper, F. S. (1958), 'Some cues for the distinction between voiced and voiceless stops in initial position', *Language and Speech* **1**, 153–166.

Liberman, A. M., Dellatre, P. C., Cooper, F. S. & Gerstman, L. J. (1954), 'The role of consonant–vowel transitions in the perception of the stop and nasal consonants', *Psychological Monographs: General and Applied* **68**(8), 1–13.

Liberman, A. M., Dellatre, P. & Cooper, F. S. (1952), 'The role of selected stimulus–variables in the perception of the unvoiced stop consonants', *American Journal of Psychology* **65**, 497–516.

Liberman, A. M., Dellatre, P., Gerstman, L. J. & Cooper, F. S. (1956), 'Tempo of frequency change as a cue for distinguishing classes of speech sounds', *Journal of Experimental Psychology* **52**(2), 127–137.

Liberman, A. M., Harris, K. S., Eimas, P. D., Lisker, L. & Bastian, J. (1961), 'An effect of learning on speech perception: The discrimination of durations of silence with and without phonemic significance', *Language and Speech* **4**, 175–195.

Liberman, A. M., Harris, K. S., Hoffman, H. S. & Griffith, B. C. (1957), 'The discrimination of speech sounds within and across phoneme boundaries', *Journal of Experimental Psychology* **54**, 358–368.

Liberman, A. M. & Mattingly, I. G. (1985), 'The motor theory of speech perception revised', *Cognition* **21**, 1–36.

Liberman, I. (1973), 'Segmentation of the spoken word and reading acquisition', *Haskins Reports on Speech Research* **SR–33**, 157–166.

Liberman, I. Y., Shankweiler, D., Fischer, F. W. & Carter, B. (1974), 'Explicit syllable amd phoneme segmentation in the young child', *Journal of Experimental Child Psychology* **18**, 201–12.

Liberman, I. Y., Shankweiler, D., Liberman, A. M., Fowler, C. & Fischer, F. W. (1978), Phonetic segmentation and recoding in the beginning reader, *in* A. S. Reber & D. L. Scarborough, eds, 'Toward a Psychology of Reading', LEA, New York.

Lindblom, B. (1989), Some remarks on the origin of the "phonetic code", *in* C. von Euler, I. Lundberg & G. Lennerstrand, eds, 'Brain and Reading', Macmillan, Basingstoke, pp. 27–44.

Lindblom, B., MacNeilage, P. & Studdert-Kennedy, M. (1983), Self–organising processes and the explanation of language universals, *in* B. Butterworth, B. Comrie & O. Dahl, eds, 'Explanations for Language Universals', Mouton, The Hague, pp. 181–203.

Lotto, A. J. & Kluender, K. R. (1998), 'General contrast effects in speech perception: Effect of preceding liquid on stop consonant identification', *Perception & Psychophysics* **60**(4), 602–619.

Lundberg, I. (1991), Phonemic awareness can be developed without reading instruction, *in* S. S. Brady & D. P. Shankweiler, eds, 'Phonological Processes in Literacy: A Tribute to Isabelle Y. Liberman', LEA, Hillsdale, NJ, pp. 47–53.

Lundberg, I., Frost, J. & Petersen, O. P. (1988), 'Effects of an extensive program for stimulating phonological awareness in children', *Reading Research Quarterly* **23**, 263–284.

Lundberg, I., Olofsson, r. & Wall, S. (1980), 'Reading and spelling skills in the first school years predicted from phoneme awareness skills in kindergarten', *Scandinavian Journal of Psychlogy* **21**, 159–173.

MacNeilage, P. F. (1994), 'Prologomena to a theory of the sound pattern of the first spoken language', *Phonetica* **51**, 184–194.

MacNeilage, P. F. & Davis, B. L. (1990), Acquisition of speech production: Frames then content, *in* M. Jeannerod, ed., 'Attention and Performance XIII: Motor Representation and Control', LEA, Hillsdale, NJ.

MacNeilage, P. F. & Davis, B. L. (1993), Motor explanations of babbling and early speech patterns, *in* B. de Boysson-Bardies, S. de Schonen, P. Jusczyk, P. MacNeilage & J. Morton, eds, 'Developmental Neurocognition: Speech and Face Processing', Kluwer Academic, Dordrecht.

MacWhinney, B. (1995), *The CHILDES Project: Tools for Analysing Talk*, second edn, LEA, Mahwah, NJ.

Manis, F. R., McBride-Chang, C., Seidenberg, M. S., Keating, P., Doi, L. M., Munson, B. & Petersen, A. (1997), 'Are speech perception deficits associated with developmental dyslexia?', *Journal of Experimental Child Psychology* **66**, 211–235.

Mann, V. (1980), 'Influence of preceding liquid on stop–consonant perception', *Perception & Psychophysics* **28**, 407–412.

Mann, V. A. (1986), 'Phonological awareness: The role of reading experience', *Cognition* **24**, 65–92.

Mann, V. A. (1991), Are we taking too narrow a view of the conditions for development of phonological awareness?, *in* S. S. Brady & D. P. Shankweiler, eds, 'Phonological Processes in Literacy: A Tribute to Isabelle Y. Liberman', LEA, Hillsdale, NJ, pp. 55–64.

Masterson, J., Hazan, V. & Wijayatilake, L. (1995), 'Phonemic processing problems in developmental phonological dyslexia', *Cognitive Neuropsychology* **12**(3), 233–259.

McBride-Chang, C. (1995*a*), 'Phonological processing, speech perception, and reading disability: An integrative review', *Educational Psychologist* **30**(3), 109–121.

McBride-Chang, C. (1995*b*), 'What is phonological awareness', *Journal of Educational Psychology* **87**(2), 179–192.

McBride-Chang, C. (1996), 'Models of speech perception and phonological processing in reading', *Child Development* **67**, 1836–1856.

McBride-Chang, C., Chang, L. & Wagner, R. (1997), 'Growth modeling of phonological awareness', *Journal of Educational Psychology* **89**(4), 621–630.

McClure, J. D. (1995), The vowels of Scottish English—formants and features, *in* J. Windsor Lewis, ed., 'Studies in General and English Phonetics', Routledge, London.

McGurk, H. & MacDonald, J. (1976), 'Hearing lips and seeing voices', *Nature* **264**, 746–748.

Mehler, J., Dupoux, E., Nazzi, T. & Dehaene-Lambertz, G. (1996), Coping with linguistic diversity: The infant's viewpoint, *in* J. L. Morgan & K. Demuth, eds, 'Signal To Syntax: Bootstrapping from Speech to Grammar in Early Acquisition', Mahwah, NJ.

Menn, L. (1983), Development of articulatory, phonetic and phonological capabilities, *in* B. Butterworth, ed., 'Language Production', Vol. 2, Academic Press, London, pp. 3–50.

Miller, J. L. & Eimas, P. D. (1981), 'Studies on the categorization of speech by infants', *Cognition* **13**, 135–165.

Mody, M., Studdert-Kennedy, M. & Brady, S. (1997), 'Speech perception deficits in poor readers: Auditory processing or phonological coding?', *Journal of Experimental Child Psychology* **64**, 199–231.

Morais, J. (1991), Constraints on the development of phonemic awareness, *in* S. S. Brady & D. P. Shankweiler, eds, 'Phonological Processes in Literacy: A Tribute to Isabelle Y. Liberman', LEA, Hillsdale, NJ, pp. 5–27.

Morais, J., Bertelson, P., Cary, L. & Alegria, J. (1986), 'Literacy training and speech segmentation', *Cognition* **24**, 45–64.

Morais, J., Cary, L., Alegria, J. & Bertelson, P. (1979), 'Does awareness of speech as a sequence of phones arise spontaneously?', *Cognition* **7**, 323–331.

Morais, J. & Kolinsky, R. (1994), 'Perception and awareness in phonological processing: the case of the phoneme', *Cognition* **50**, 287–297.

Morais, J. & Kolinsky, R. (1995), The consequences of phonemic awareness, *in* B. de Gelder & J. Morais, eds, 'Speech and Reading: A Comparative Approach', LEA, UK.

Morrongiello, B. A., Robson, R. C., Best, C. T. & Clifton, R. K. (1984), 'Trading relations in the perception of speech by five–year–old children', *Journal of Experimental Child Psychology* **37**, 231–250.

Muter, V., Hulme, C., Snowling, M. & Taylor, S. (1997), 'Segmentation, not rhyming, predicts early progress in learning to read', *Journal of Experimental Child Psychology* **65**, 370–396.

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

Nittrouer, S. (1993), 'The emergence of mature gestural patterns is not uniform—evidence from an acoustic study', *Journal of Speech and Hearing Research* **36**(5), 959–972.

Nittrouer, S. (1995), 'Children learn separate aspects of speech production at different rates: Evidence from spectral moments', *Journal of the Acoustical Society of America* **97**(1), 520–530.

Nittrouer, S. (1996*a*), 'Discriminability and perceptual weighting of some acoustic cues to speech perception by 3-year-olds', *Journal of Speech and Hearing Research* **39**(2), 278–297.

Nittrouer, S. (1996*b*), 'The relation between speech perception and phonemic awareness: Evidence from low–SES children and children with chronic OM', *Journal of Speech and Hearing Research* **39**(5), 1059–1070.

Nittrouer, S. (1999), 'Do temporal processing deficits cause phonological processing problems?', *Journal of Speech, Language, and Hearing Research* **42**, 925–942.

Nittrouer, S. & Crowther, C. S. (1998), 'Examining the role of auditory sensitivity in the developmental weighting shift', *Journal of Speech, Language, and Hearing Research* **41**, 809–818.

Nittrouer, S., Crowther, C. S. & Miller, M. E. (1998), 'The relative weighting of acoustic properties in the perception of [s]+stop clusters by children and adults', *Perception & Psychophysics* **60**(1), 51–64.

Nittrouer, S., Manning, C. & Meyer, G. (1993), 'The perceptual weighting of acoustic cues changes with linguistic experience', *Journal of the Acoustical Society of America* **94**(3), S1865.

Nittrouer, S. & Miller, M. E. (1997*a*), 'Developmental weighting shifts for noise components of fricative–vowel syllables', *Journal of the Acoustical Society of America* **102**(1), 572–580.

Nittrouer, S. & Miller, M. E. (1997*b*), 'Predicting developmental shifts in perceptual weighting schemes', *Journal of the Acoustical Society of America* **101**(4), 2253–2266.

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

Nittrouer, S., Studdert-Kennedy, M. & McGowan, R. (1989), 'The emergence of phonetic segments: Evidence from the spectral structure of fricative–vowel syllables spoken by children and adults', *Journal of Speech and Hearing Research* **32**, 120–132.

Nittrouer, S., Studdert-Kennedy, M. & Neely, S. T. (1996), 'How children learn to organize their speech gestures: Further evidence from fricative-vowel syllables', *Journal of Speech and Hearing Research* **39**(2), 379–389.

Ohde, R. N. (1994), 'The development of the perception of cues to the [m]–[n] distinction in CV syllables', *Journal of the Acoustical Society of America* **96**(2), 675–686.

Ohde, R. N. & Haley, K. L. (1997), 'Stop–consonant and vowel perception in 3–and 4–year–old children', *Journal of the Acoustical Society of America* **102**(6), 3711–3722.

Ohde, R. N., Haley, K. L. & McMahon, C. W. (1996), 'A developmental study of vowel perception from brief synthetic consonant–vowel syllables', *Journal of the Acoustical Society of America* **100**(6), 3813–3824.

Öhman, S. (1966), 'Coarticulation in VCV utterances: Spectrographic measures', *Journal of the Acoustical Society of America* **39**, 159–168.

Oney, B. & Goldman, S. R. (1984), 'Recoding and comprehension skills in Turkish and English: Effects of the regularity of grapheme–phoneme correspondences', *Journal of Educational Psychology* **76**, 447–568.

Pardo, J. S. & Fowler, C. A. (1997), 'Perceiving the causes of coarticulatory acoustic variation: Consonant voicing and vowel pitch', *Perception & Psychophysics* **59**(7), 1141–1152.

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

Polka, L. & Werker, J. F. (1994), 'Developmental changes in perception of non–native vowel contrasts', *Journal of Experimental Psychology: Human Perception and Performance* **20**, 421–435.

Porpodas, C. D. (1989), The phonological factor in reading and speaking of Greek, *in* P. G. Aaron & R. M. Joshi, eds, 'Reading and Writing Disorders in Different Orthographic Systems', Kluwer Academic Publishers, Dordrecht, pp. 177–190.

Raz, I. S. & Bryant, P. (1990), 'Social background, phonological awareness and children's reading', *British Journal of Developmental Psychology* **8**, 209–225.

Read, C. A., Zhang, Y., Nie, H. & Ding, B. (1986), 'The ability to manipulate speech sounds depends on knowing alphabetic reading', *Cognition* **24**, 31–44.

Reed, M. A. (1989), 'Speech perception and the discrimination of brief auditory cues in reading disabled children', *Journal of Experimental Child Psychology* **48**, 270–292.

Repp, B. H., Liberman, A. M., Eccardt, T. & Pesetsky, D. (1978), 'Perceptual integration of acoustic cues for stop, fricative and affricate manner', *Journal of Experimental Psychology: Human Perception and Performance* **4**, 621–637.

Rosen, S. & Howell, P. (1987), Auditory articulatory, and learning explanations of categorical perception in speech, *in* S. Harnad, ed., 'Categorical Perception: The Groundwork of Cognition', Cambridge University Press, pp. 113–160.

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

Scottish Office Education Department (1991), 'English Language 5–14', *Curriculum and Assessment in Scotland: National Guidelines* .

Shankweiler, D., Liberman, I. Y., Mark, L. S., Fowler, C. A. & Fischer, F. W. (1979), 'The speech code and learning to read', *Journal of Experimental Psychology: Human Learning and Memory* **5**, 531–545.

Shattuck-Huffnagel, S. (1983), Sublexical units and suprasegmental structure in speech production, *in* P. F. MacNeilage, ed., 'The Production of Speech', Springer–Verlag, New York, pp. 109–136.

Shattuck-Huffnagel, S. (1987), The role of word–onset consonants in speech planning: New evidence from speech error patterns, *in* E. Keller & M. Gopnik, eds, 'Motor and Sensory Processes of Language', LEA, Hillsdale, NJ, pp. 17–51.

Simon, C. & Fourcin, A. J. (1978), 'Cross–language study of speech pattern learning', *Journal of the Acoustical Society of America* **63**(3), 925–935.

Snowling, M., Goulandris, N., Bowlby, M. & Howell, P. (1986), 'Segmentation and speech perception in relation to reading skill: A developmental analysis', *Journal of Experimental Child Psychology* **41**, 468–507.

Stevens, K. & Blumstein, S. (1981), The search for invariant acoustic correlates of phonetic features, *in* P. D. Eimas & J. L. Miller, eds, 'Perspectives on the Study of Speech', LEA, Hillsdale, NJ, pp. 1–38.

Stevens, K. N. (1980), 'Acoustic correlates of some phonetic categories', *Journal of the Acoustical Society of America* **68**(3), 836–842.

Stevens, K. N. (1985), Evidence for the role of acoustic boundaries in the perception of speech sounds, *in* V. Fromkin, ed., 'Phonetic Linguistics: Essays in Honor of Peter Ladefoged', Academic Press, New York.

Stevens, K. N. & Klatt, D. H. (1974), 'Role of formant transitions in the voiced–voiceless distinction for stops', *Journal of the Acoustical Society of America* **55**(3), 653–659.

Streeter, L. A. (1976), 'Language perception of 2–month–old infants shows effects of both innate mechanisms and experience', *Nature* **259**, 39–41.

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

Studdert-Kennedy, M. (1991), Language development from an evolutionary perspective, *in* N. A. Krasnegor, D. M. Rumbaugh, R. L. Schiefelbusch & M. Studdert-Kennedy, eds, 'Biological and Behavioral Determinants of Language Development', LEA, Hillsdale, NJ.

Sussman, H. M., Fruchter, D. & Cable, A. (1995), 'Locus equations derived from compensatory articulation', *Journal of the Acoustical Society of America* **97**, 3112–3124.

Sussman, H. M. & Shore, J. (1996), 'Locus equations as phonetic descriptors of consonantal place of articulation', *Perception & Psychophysics* **58**(6), 936–946.

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

Sussman, J. E. & Carney, A. E. (1989), 'Effects of transition length on the perception of stop consonants by children and adults', *Journal of Speech and Hearing Research* **32**(1), 151–160.

Tallal, P. (1980), 'Auditory temporal perception, phonics and reading disabilities in children', *Brain and Language* **9**, 182–198.

Tallal, P. & Peircy, M. (1973), 'Developmental aphasia: Impaired rate of nonverbal processing as a function of sensory modality', *Neuropsychologia* **11**, 389–395.

Tallal, P. & Peircy, M. (1975), 'Developmental aphasia: The perception of brief vowels and extended stop consonants', *Neuropsychologia* **13**, 69–74.

Tallal, P. & Stark, R. E. (1981), 'Speech acoustic cue discrimination abilities of normally developing and language–impaired children', *Journal of the Acoustical Society of America* **69**, 568–574.

Trehub, S. E. (1976), 'The discrimination of foreign speech contrasts by infants and adults', *Child Development* **47**, 466–472.

Treiman, R. & Baron, J. (1981), Segmental analysis ability: Development and relation to reading ability, *in* T. G. Waller & G. E. McKinnon, eds, 'Reading Research: Advances in Theory and Practice', Vol. 2, Academic Press, New York, pp. 159–198.

Treiman, R. & Zukowski, A. (1991), Levels of phonological awareness, *in* S. S. Brady & D. P. Shankweiler, eds, 'Phonological Processes in Literacy: A Tribute to Isabelle Y. Liberman', LEA, Hillsdale, NJ, pp. 67–83.

Treiman, R. & Zukowski, A. (1996), 'Children's sensitivity to syllables, onsets, rimes and phonemes', *Journal of Experimental Child Psychology* **61**, 193–215.

Tunmer, W. E. & Nesdale, A. R. (1982), 'The effects of digraphs and pseudowords on phonemic segmentation in young children', *Applied Psycholinguistics* **3**, 299–311.

Vihman, M. M. (1992), Early syllables and the construction of phonology, *in* C. A. Ferguson, L. Menn & C. Stoel-Gammon, eds, 'Phonological Development: Models, Research Implications', York Press, Timonium, MD, pp. 393–422.

Vihman, M. M. (1993), 'Vocal motor schemes, variation and the production–perception link', *Journal of Phonetics* **21**, 163–169.

Vihman, M. M. (1996), *Phonological Development: The Origins of Language in the Child*, Blackwell, Oxford.

Vihman, M. M., Macken, M. A., Miller, R., Simmons, H. & Miller, J. (1985), 'From babbling to speech: A reassessment of the continuity issue', *Language* **67**, 297–319.

Vihman, M. M. & Velleman, S. L. (1989), 'Phonological reorganization: A case study', *Language and Speech* **32**, 149–170.

Vihman, M. M., Velleman, S. L. & McCune, . (1994), How abstract is child phonology? Towards an integration of linguistic and psychological approaches, *in* M. Yava¸s, ed., 'First and Second Language Acquisition', Singular Publishing Group, San Diego.

Wallach, L., Wallach, M. A., Dozier, M. G. & Kaplan, N. E. (1977), 'Poor children learning to read do not have trouble with auditory discrimination but do have trouble with phoneme recognition', *Journal of Educational Psychology* **69**, 36–39.

Walley, A. C. & Carrell, T. D. (1983), 'Onset spectra and formant transitions in the adult's and child's perception of place of articulation in stop consonants', *Journal of the Acoustical Society of America* **73**, 1011–1022.

Walley, A. C., Smith, L. B. & Jusczyk, P. W. (1986), 'The role of phonemes and syllables on the perceived similarity of speech sounds for children', *Menory and Cognition* **14**(3), 220–229.

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

Wardrip-Fruin, C. (1985), 'The effect of signal degredation on the status of cues to voicing in utterance–final stop consonants', *Journal of the Acoustical Society of America* **77**(5), 1907–1912.

Wardrip-Fruin, C. & Peach, S. (1984), 'Developmental aspects of the perception of acoustic cues in determining the voicing feature of final stop consonants', *Language and 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.

Werker, J. F. & Lalonde, C. E. (1988), 'Cross–language speech perception: Initial capabilities and developmental change', *Developmental Psychology* **24**, 627–683.

Werker, J. F. & Tees, R. C. (1984), 'Cross–language speech perception: Evidence for perceptual reorganization during the first year of life', *Infant Behaviour and Development* **7**, 49–63.

Werker, J. F. & Tees, R. C. (1987), 'Speech perception in severely disabled and average reading children', *Canadian Journal of Psychology* **41**(1), 48–61.

Whalen, D. H. (1981), 'Effects of vocalic formant transitions and vowel quality on the english [s]–[š] boundary', *Journal of the Acoustical Society of America* **69**(1), 275–282.

Whalen, D. H. (1991), 'Perception of the english /s/–/ʃ/ distinction relies on fricative noises and transitions, not brief spectral slices', *Journal of the Acoustical society of America* **90**(4), 1776–1784.

Wimmer, H. & Goswami, U. (1994), 'The influence of orthographic consistency on reading development: Word recognition in English and German children', *Cognition* **51**, 91–103.