THE PERCEPTION AND PRODUCTION OF STRESS AND INTONATION BY CHILDREN WITH COCHLEAR IMPLANTS

ROSEMARY O'HALPIN
University College London
Department of Phonetics & Linguistics

A dissertation submitted to the University of London for the
Degree of Doctor of Philosophy

2010

ABSTRACT

Users of current cochlear implants have limited access to pitch information and hence to intonation in speech. This seems likely to have an important impact on prosodic perception. This thesis examines the perception and production of the prosody of stress in children with cochlear implants. The interdependence of perceptual cues to stress (pitch, timing and loudness) in English is well documented and each of these is considered in analyses of both perception and production. The subject group comprised 17 implanted (CI) children aged 5;7 to 16;11 and using ACE or SPEAK processing strategies. The aims are to establish

(i) the extent to which stress and intonation are conveyed to CI children in synthesised bisyllables (BAba vs. baBA) involving controlled changes in F_0, duration and amplitude (Experiment I), and in natural speech involving compound vs. phrase stress and focus (Experiment II).

(ii) when pitch cues are missing or are inaudible to the listeners, do other cues such as loudness or timing contribute to the perception of stress and intonation?

(iii) whether CI subjects make appropriate use of F_0, duration and amplitude to convey linguistic focus in speech production (Experiment III).

Results of Experiment I showed that seven of the subjects were unable to reliably hear pitch differences of 0.84 octaves. Most of the remaining subjects required a large (approx 0.5 octave) difference to reliably hear a pitch change. Performance of the CI children was poorer than that of a normal hearing group of children presented with an acoustic cochlear implant simulation. Some of the CI children who could not discriminate F_0 differences in Experiment I nevertheless scored above chance in tests involving focus in natural speech in Experiment II. Similarly, some CI subjects who were above chance in the production of appropriate F_0 contours in Experiment III could not hear F_0 differences of 0.84 octaves. These results suggest that CI children may not necessarily rely on F_0 cues to stress, and in the absence of F_0 or amplitude cues, duration may provide an alternative cue.

ACKNOWLEDGEMENTS

For guidance and direction I am indebted to my supervisor Dr Andrew Faulkner, who has been accessible and helpful with every aspect of my research, and who has given careful criticism of various drafts of the thesis. I am grateful to Professor Stuart Rosen for constructive comments at different stages of this dissertation; to my specialist adviser Ms Laura Viani, consultant ENT surgeon and director of the Cochlear Implant Programme at Beaumont Hospital, Dublin for her encouragement and support; and to Dr Evelyn Abberton for helpful suggestions at the early stages of this project. I am also grateful to the Health Research Board for a Health Services Research Fellowship which partly funded this research.

My thanks are also due to Dr Yi Xu for providing a custom-written PRAAT script for F_0 extraction and measurements, and for helpful discussions on prosodic issues; Dr Gary Norman for audiological and mapping details for the children with implants; Steve Nevard for setting up the audio recordings and Dave Cushing for technical assistance at UCL; Jill House for suggestions regarding intonation issues; Dr Michael O’Kelly for help with statistics and comments on a draft of the thesis; Professor Neil Smith and Professor Valerie Hazan for feedback on earlier drafts of some of the chapters.

For arranging the use of soundproof facilities in their respective locations I am grateful to Anne Marie Gallagher and her colleagues at Beaumont Hospital, Dr Jesudas Dayalan (Clonmel), and Nick Devery and Bernie Lowry (Tullamore).

I must also thank Michael Ashby, Dr Volker Dellwo, Dr Yu-ching Kuo, Anne Parker and Dr Celia Wolf, at UCL for their assistance; my colleagues in the Cochlear Implant Programme and at Beaumont Hospital for all their support; Mary and Billy Kelly, Sian Kelly, Dr. Kate O’Malley and Patrick O’Halpin for their assistance; Julia Boyle at Beaumont Hospital for arranging secondment to carry out this research; Dr Anne Cody and Patricia Cranley at the Health Research Board and Gemma Heath at the Royal College of Surgeons in Ireland for their assistance; Barry O’Halpin for creating pictures for the experiments; Nuala Scott and Patricia Vila for formatting the final drafts of the manuscript.

I am very grateful to the children, their families, and the talkers who participated enthusiastically in this study, and to the visiting teachers of the children with cochlear implants who were very helpful.

Finally, my thanks to my family, relatives and friends, especially to Patrick and Barry for their support and encouragement in the final stages of this project.

TABLE OF CONTENTS

CHAPTER THREE – EXPERIMENT II: SENSITIVITY TO VARIATIONS IN STRESS AND INTONATION IN NATURAL SPEECH STIMULI

107

CHAPTER FOUR – THE PRODUCTION OF FOCUS BY CI AND NH TALKERS: ACOUSTIC MEASUREMENTS OF F_0, AMPLITUDE AND DURATION **159**

LIST OF FIGURES

LIST OF TABLES

LIST OF APPENDICES

CHAPTER ONE

BACKGROUND & REVIEW OF THE LITERATURE

1.1 Introduction

Most research on the design and assessment of cochlear implant speech processing strategies has focussed on vowel and consonant perception in English, and little attention has been given to pitch and intonational aspects of speech. There have, however, been a few studies of pitch perception for speech in lexical tone languages such as Mandarin and Cantonese, where pitch determines meaning in otherwise identical syllables.

The limitations of current speech processing strategies in delivering adequate pitch information to implant users are well documented. In the electrode array in the cochlea, the entire speech frequency range has to be spread over a limited number of channels resulting in poor spectral resolution compared to normal hearing. One consequence of this limited spectral resolution is that the primary auditory cues to pitch used by normal hearing listeners are unavailable. It appears that implant users rely on relatively weak cues to pitch that are carried in the temporal modulation patterns.

Overview of the thesis

The current study investigates the perception and production of intonation and stress contrasts by early and later implanted children ranging between 5;7 and 17;4 years using two commonly used speech processing strategies (i.e. ACE and SPEAK in multi-channel implants. Normal hearing children of a matching age range are included in the perception experiments for comparison.

The hypotheses and theoretical basis for the experiments and analyses are discussed in detail in *Chapter One* (see sections 1.1 – 1.10). The relevance of these theoretical issues to the perception and production experiments is discussed in section 1.11.

In *Chapter Two* an adaptive 2 down-1 up staircase is used in a controlled experiment to establish the smallest discriminable F_0 (fundamental frequency), duration and amplitude differences between stressed and unstressed syllables (*Experiment I*). Non-meaningful synthesised pairs of /baba/ stimuli are presented with similar or different stress positions in a same/different task procedure. The advantage of this type of task is that no linguistic demands are made on the children, and performance depends on hearing ability**.** The synthesised stimuli are also presented in an acoustic simulation of a cochlear implant to the group of normal hearing children.

In *Chapter Three* recorded natural speech stimuli are presented with picture prompts in two different tasks requiring linguistic as well as hearing ability (*Experiment II*). In one task subjects are asked to discriminate differences in lexical stress in compounds and noun phrases such as *blackboard vs. black board.* In a second task subjects are required to identify the focus word in final and non-final focus position in two element phrases such as *a BLUE book* vs. *a blue BOOK* or three element declarative sentences such as *the BOY is painting a boat* vs. *the boy is painting a BOAT.* The advantage of the recorded stimuli is that there is consistency in how the stimuli are delivered to each subject, and the same inter or intra speaker differences remain constant throughout.

In *Chapter Four* acoustic analysis of the production of F_0, duration and amplitude is carried out for multiple repetitions of elicited focus in three element sentences (*Experiment III*) from the children with cochlear implants as well as four normal hearing subjects. These three element sentences are the same as those presented in the perception tasks in Experiment II. A question and answer sequence is used with picture prompts to elicit semi-spontaneous speech which ensures that the task is understood by children across the age range. A limited set of familiar vocabulary items is elicited in declarative sentence by picture prompts which avoid unexpected linguistic complexities such as embedded language or inference that might arise in completely spontaneous conversations.

However, even if appropriate adjustments of one or a combination of acoustic cues (i.e. F_0, duration, or amplitude) are made by individual implanted children in the focus words/syllables in Experiment III, what matters ultimately is whether they manage to convey focus on the appropriate word to a listener. For this reason auditory judgements by an experienced listener (i.e. the present investigator) of the CI subjects' appropriate production of focus are included in the analyses of the data in Experiment III.

1.1.1 Limited previous research

To date there has been very little previous systematic research into the perception and production of stress and intonation in English by children with cochlear implants. Intonation is involved in many aspects of language, including grammar, semantics, pragmatics, affect, and interaction. Yet the perception of pitch is difficult for implant users and it is possible that this, perhaps combined with other factors, can hinder the development of language.

A few prosodic aspects of English, however, have been investigated for implanted children. These include pitch discrimination in a study of voice similarity and talker discrimination (Cleary, Pisoni and Kirk, 2005), and weak syllable processing (Titterington, Henry, Kramer, Toner and Stevenson, 2006). More attention has been given to pitch perception and production in Chinese tone languages such as Mandarin and Cantonese (Barry and Blamey, 2004; Barry, Blamey, Martin, Lees, Tang, Ming and van Hasselt, 2002a; Barry, Blamey and Martin, 2002b; Ciocca, Francis, Aisha and Wong, 2002; Peng, Tomblin, Cheung, Lin and Wang, 2004; Xu, Li, Hao, Chen, Xue and Han, 2004) where pitch determines meaning in otherwise identical syllables. Apart from the study of weak syllable processing by Titterington et al. detailed investigation of intonational issues has not yet been carried out for English speaking children with cochlear implants. Most of the developmental literature on intonational contrasts such as lexical stress and focus in normal hearing children is based on British (Wells, Peppé and Goulandris, 2004; Cutler and Swinney, 1987; Dankovičová, Piggott, Wells and Peppé, 2004) or American populations (Atkinson-King, 1973; Vogel and Raimy, 2002). There have been no large scale normative studies of intonation skills of children using Southern Hiberno English (SHE) but there have been a few reports on discrimination of compound vs. phrase pairs, questions, statements, commands and emotional prosody in 8;0 year old normal hearing children (Doherty, Fitzsimons, Assenbauer and Staunton, 1999) and production of contrastive stress by an 8;0 year old hearing child and hearing aid users (O'Halpin, 1993, 1997).

The current study investigates the perception of stress and intonation in lexical stress and focus by a group Southern Hiberno English speaking children with cochlear implants and a normal hearing group within the same age range. The production of focus by the implanted children will also be examined and the wide age range (5;0 –

17;0) of the normal hearing and implanted children should provide additional information on the development of intonation skills in children beyond age 12;0 or 13;0 years. This older age group has not received much attention in the general acquisition literature. For normal hearing listeners there are a number of interdependent perceptual cues to stress and intonation (pitch, timing, loudness). Experimental evidence shows that pitch makes syllables stand out and seem more prominent to listeners. However, given the limitations of pitch information available through current speech processors it is possible that cochlear implant users rely more on timing and loudness cues. These issues are investigated for a group of implanted children in controlled perception experiments using synthesised and natural speech stimuli.

1.1.2 The hypotheses and framework for the current study

It seems to be widely believed that F_0 (fundamental frequency) is the most important cue to stress although there is some evidence that this may vary according to individual subjects, the context of the data, or how it is elicited. Whether F_0 is the primary cue in signalling intonation contrasts remains to be determined (sections 1.2 and 1.4) for normal hearing subjects but the issue is further complicated for children using cochlear implants. Coding of F_0 (or the perceptual correlate pitch) is limited in cochlear implants (see section 1.7) and implanted children may only have access to duration and amplitude cues. To date very little attention has been given to the perception and production of linguistic stress and intonation contrasts (e.g. compound vs. noun phrase or focus) in English speaking children with implants. It has yet to be established whether the perception and production of intonation: -

(i) are directly linked to the implanted children's ability to hear F_0 and intonation development depends on their auditory skills.

or

(ii) are not directly linked to any one cue and intonation develops as an abstract phonological system which is not necessarily perceived and produced by the same cues.

The hypotheses in (i) and (ii) above will be discussed in more detail below.

(i) *F_0 is a necessary cue to stress and intonation*

If F_0 is a necessary cue to stress and intonation implanted children will need good access to pitch cues (perceptual correlate of F_0) in order to hear these contrasts. In order to produce intonation contrasts they will need to be able to hear them in their ambient environment. If these children do not have access to F_0, the intonation contrasts will not be accessible to them and consequently they will not develop abstract phonological representations in the same way as normal hearing children. In other words they will not be able to hear the F_0 patterns associated with pragmatic contrasts such as given vs. new or focussed words, or grammatical contrasts such as compounds vs. noun phrases. Because they have no prior knowledge or stored representation of how intonation conveys these contrasts they will never learn to produce them appropriately. The tendency for exaggerated pitch contrasts or rising pitch for encouragement used by adults in speech directed at children during the early stages of prosodic development will not be accessible to implanted children and will put them at a disadvantage compared to normal hearing children (section 1.3).

However, F_0 cues may not be completely inaccessible to implant users, and experiments with implanted children using Chinese tones (section 1.8) and with English speaking implanted adults (section 1.9) have indicated that if there is a big enough F_0 difference between pairs of stimuli this might be perceived by some implant users. If this is the case, the exaggerated pitch changes typical in the speech of adults to children might be more accessible to implanted children during early prosodic development and will help them develop some phonological awareness of stress and intonation contrasts cued by F_0. However, a number of studies indicate that implanted children and adults often have difficulty hearing F_0 differences of less than half an octave as found in everyday speech. In any case, as children using implants grow older they will be unable to hear the more subtle pitch changes used in everyday adult speech which will hinder further development of intonation skills needed to interpret and convey more advanced linguistic contrasts (e.g. pragmatic, semantic, grammatical, interactive). All of the possibilities set out above follow from the hypothesis in (i) above that input (i.e. perception of F_0) is directly linked to output (production of F_0) and that intonation development depends on implanted children's ability to hear F_0 differences.

(ii) F_0 *is <u>not</u> a necessary cue to stress and intonation*

In contrast with all of this, if F_0 plays a less important role in the perception and production of intonation, implanted children will be able to rely on other cues such as duration and amplitude. This puts them at much less of a disadvantage during the early stages of prosodic development. There are other adjustments in prosodic cues besides pitch in the speech of adults such as extra lengthening, longer pauses and changes in loudness which can facilitate prosodic development. In addition, paralinguistic cues such as eye contact, gestures, jumping up and down and reaching which will draw attention to certain features such as response required or not required, rhythm or focus. In this way implanted children can perceive stress, intonation and other contrasts using whatever cues are available to them and develop an abstract prosodic and linguistic system which is independent of their ability to hear a particular cue. Studies of young normal hearing children suggest that the production of linguistic stress and intonation does not necessarily develop in parallel with perception (section 1.3), and that sometimes children can produce focus, for example, in their own speech before they can interpret some aspects of focus in the speech of others. This is attributed to a physiological reflex associated with semantic interest in a word which in turn generates tension and increases F_0. It is possible that implanted children, having acquired an abstract representation of prominence or a key word, can try to convey focus by producing appropriate increases or changes in F_0 as a physiological reflex without being able to hear these F_0 changes when produced by others. This would support the hypothesis in (ii) above that intonation contrasts such as focus develop as abstract phonological systems which are not necessarily perceived or produced by the same cues.

1.2 Linguistic aspects of stress and intonation in English

English is described as a stress language where each word in citation form has one main stress which may shift in continuous speech to maintain regularity (Roach, 1982; Cruttenden, 1997; Fujimura and Erickson, 1997). In English, word stress or lexical stress is not fixed and is generally not predictable except with reference to a complex set of rules. However, there are some cases where word stress can be used to indicate differences in lexical meaning or grammatical class such as deFER versus DIFFer or INsult (noun) versus inSULT (verb). In addition, compound word combinations have the primary stress on the first element such as BLACKboard as opposed to blackBOARD (Cruttenden, 1997).

For normal hearing listeners the perceptual parameters of stress (pitch, timing and loudness) make certain syllables stand out to listeners (Cruttenden, 1997; Crystal, 1969; Faure, Hirst and Chacouloff, 1980; Ladd, 1980; Borden, Harris and Raphael, 1994). In any stretch of speech a speaker can impose rhythmical structure on an utterance and make a particular stressed syllable prominent by pitch movement or accent (Ladd, 1980, 1996). There can be more than one accented syllable in an utterance and the pattern of pitch changes in a stretch of speech is referred to as intonation (Ladd, 1996; Fujimura and Erickson, 1997; Cruttenden, 1997; Ladefoged, 2001). However, Rahilly (1998) suggests that an agreed phonological approach needs to be developed to gain better insight into regional and sociolinguistic variation. For example in Belfast English intonation (BfE) tone-groups (i.e. intonation groups) are defined on the basis of pause and not by perceivable pitch change as for British English (Rahilly, 1997). Rahilly (p.115) considers the generally accepted view of a single nucleus per tone-group problematic and prefers to use the term ‘prominence’. The BfE data suggest that there can be more than one peak of prominence within each pause-defined unit, and the author has also used this approach in a study of deafened speakers of BfE (Rahilly, 1991). See 1.4.4 for further discussion of regional variation.

At the linguistic level various oppositions are found in the literature between broad and narrow focus, given and new or contrastive information, or a speaker may wish to emphasise a particular word for grammatical purposes. However, the distinction between new and contrastive information is not always clear in the literature. For example, according to Halliday new information may be ‘cumulative to or contrastive with what has preceded’ (Couper-Kuhlen, 1986, p.125), and if for some reason we focus on old information this too can be described as contrastive (Cruttenden, 1997, pp. 82-84). For example, in a sentence such as *the boy is painting a boat* used in the present study contrast can be *implicit* in a particular context

the BOY ***(and not the girl, man, woman..)*** *is painting a boat*

or *explicit* where a speaker highlights or brings the word *BOY* into focus in response to a question such a

Is the GIRL painting a boat?

No, the BOY is painting a boat

It has also been suggested that when new and contrastive items occur together there is a difference in the pitch configuration with a steeper fall or a higher pitch or key on the contrastive item (Chafe, 1974; Brown, Curry and Kenworthy, 1980; Brazil, Coulthard and Johns, 1980). On the other hand, according to Ladd (1980, 1996) contrastive stress may simply be a process of deaccenting or boosting of old or new information respectively. The development of autosegmental-metrical (AM) theory (Pierrehumbert, 1980; Beckman and Pierrehumbert, 1986) brought together levels (tone-sequences) and configurations (contours) in a system which represented the intonation contour as a string of pitch accents and boundary or phrasal tones in prosodic domains of varying sizes. Different pitch accent types (e.g. H* L L%) were identified which corresponded to nuclear tones (e.g. a fall*)* in the British tradition (Ladd, 1996, p.82). For further discussion of these and related issues beyond the scope of the current investigation see Ladd and Shepman (2003) and references therein.

Ladd (1996) is critical of earlier systems which tried to map acoustic correlates such as F_0, duration and intensity to new, contrastive or given information and states that subsequent approaches have taken the view that words can be in focus for various reasons and are marked by pitch accents. More recently Xu and Xu (2005) take the view that focus is a communicative function which is realised in *parallel* rather than alternating with other ‘F_0- controlling functions’ (p. 293) as assumed in the American autosegmental-metrical (AM) and the British nuclear tone theories. According to Xu and Xu the location of local F_0 peaks is not determined by focus itself but by articulatory mechanisms, and the characteristics of F_0 peaks on stressed syllables are determined by narrow focus with pitch adjustment such as ‘expansion under focus, compression after focus, and little or no change before focus' (p.186). In other words there is an increase in the size of the peak (generally accompanied by increases in duration and amplitude) on the stressed focus word, the pre-focus F_0 peaks remain unchanged, and the post-focus F_0 peaks are lower than in neutral conditions. The sharp drop in F_0 following the focus word which is treated differently by the British (high-fall nuclear accent) and American AM theories (two separate levels i.e. transition from accentual H* or LH* to the phrasal level L-) is regarded by Xu and Xu as simply a consequence of the pitch adjustments described above and intrinsic to focus (p.187).

Gussenhoven (2006) discusses types of focus in English and challenges traditional single oppositions or 'semantic contrasts' mentioned earlier such as broad and narrow, old and new, or neutral and contrastive. He lists various focus meanings or types which are signalled by pitch accents in the intonation contour such as 'presentational focus' (corresponding overtly or implicitly to an answer to a question), 'corrective' focus which is commonly referred to as 'narrow' or 'contrastive' (a rejection of an alternative), 'reactivating' focus (commonly referred to as 'old' information), or 'countersupposition' focus (a correction of information detected in the hearer's discourse).

The linguistic aspects of stress and intonation in English discussed above will be taken into consideration for normal hearing subjects and cochlear implant users in the discussion of acoustic measurements the production of focus in Chapter Four.

1.2.1 The theoretical basis for auditory judgements of stress and intonation in the present study

The British tone group (O'Connor and Arnold, 1973) theory specifies a single nucleus on the last accented syllable which consists of a glide, obtrusion, or movement in pitch which makes it more perceptually prominent than other stressed syllables. Some authors refer to the placement of extra prominence on a stressed syllable as *tonicity, sentence stress* or *nuclear stress* (Crystal 1969, 1987; Wells and Local, 1993). Difficulties arise when a pre-final accented or stressed syllable is made prominent for reason of focus or contrast. A 'fixed' nucleus on the last accented syllable then becomes downgraded and then we might have superordinate and subordinate nuclei

(Couper-Kuhlen, 1986). However, not all varieties of English conform to the notion of a single nucleus, for example, experiments with Belfast English (Rahilly, 1991, 1997) and Scottish English (Brown et al., 1980) found more than one prominent syllable in their tone groups and that tone boundaries were signalled by pause and not by pitch movement. The notion of a single nucleus has also been problematic in the analysis of speech produced by deaf children with established rhythmic problems such as inappropriate pausing, and inability to make a distinction between stressed and unstressed syllables (O'Halpin, 1993, 1997, 2001). The autosegmental metrical (AM) approach (Beckman and Pierrehumbert, 1986) represents the intonation contour as a series of pitch accents (H* or L* tones), and the nucleus is simply treated as the last accented syllable in the intonation phrase even when earlier syllables are in focus. Pitch accents become prominent when a speaker wishes to convey new information and focus (Ladd, 1996), and this approach suits the analysis of the production data in the current study where focus is elicited on target pitch accented words. If the focus occurs early in the sentence the following pitch accents may become deaccented.

The auditory judgement of focus, for example, on target words in different focus positions is concerned with whether implanted and normal hearing subjects have succeeded in conveying focus to a trained listener. Given the limitations of cochlear implants (section 1.7) in delivering adequate pitch information the main issue addressed in this particular investigation is whether or how these children convey focus to a listener. It is also of interest whether the target focus words are ambiguous or contrastive enough especially in final sentence position where other discourse factors such as turn delimitation come into play (see section 1.3.2.2). Once we have established whether these children can convey focus we need to see how they compare with normal hearing children in their own linguistic environment (i.e. different varieties of Southern Hiberno English) as well as other varieties of English, but this is beyond the scope of the present study as normative studies for hearing adults and children have yet to be carried out.

1.3 Developmental issues in the perception and production of stress and intonation

1.3.1 The early years

1.3.1.1 Perception

According to Jusczyk (1997, 2002) word segmentation skills developed in the second half of the first year lay the foundation for the development of a lexicon and of language acquisition generally. Before they can segment words from fluent speech, normal-hearing infants learn about the predominant rhythmic properties and stress and intonation patterns in their native language from the input they receive. By a process of 'prosodic bootstrapping' (Jusczyk, 1997, p.157), clausal units and phrase boundaries in the input are marked off, putting the infant in a position to extract the underlying syntactic organisation of an utterance at a later stage. Jusczyk (1997, 2002) cites perceptual experiments (Cutler and Norris, 1988; Cutler and Carter, 1987; Jusczyk, Cutler and Redanz, 1993) which indicate that there is a trochaic bias (strong followed by weak) in hearing English-learning infants. Another study (Jusczyk, Houston and Newsome, 1999) cited by Jusczyk (2002, p.13) suggests that by 9 months a preference for stressed versus unstressed syllables is shown and that by 10.5 months words beginning with unstressed syllables can be segmented.

Cruttenden (1994) in a review of phonetic and prosodic aspects of **B**aby **T**alk (BTph and BTPr), suggests that the universal existence of prosodic adjustments by adults in talk directed at very young children, such as wide pitch range, use of higher pitch, more frequent use of rising intonation for encouragement, slower articulation rate, longer pauses and whispered speech supports the case for the facilitative effects of infant-directed speech on language acquisition. Although it is reported that infants perceive rhythmic differences in their own language in the first year and by age two can produce novel compounds, the perceptual distinction between compound and phrase stress can take up to and beyond 12;0 years to develop (Vogel and Raimy, 2002). Vogel and Raimy suggest that infant studies explore sensitivity to acoustic patterns (pitch, duration and loudness) but this does not necessarily mean that a specific linguistic meaning is associated with the acoustic pattern. The contrastive use of stress, however, does require higher level processing to associate a specific meaning with an acoustic stress pattern, and is investigated at a later stage of development (p.226). Pitch adjustments by adults such as those listed above may not be accessible to young children using cochlear implants during the early stages of language acquisition because very limited pitch information is delivered via the implant. The aim of the current investigation is to establish whether children using implants can rely on other more accessible cues (i.e. timing and/or loudness) to benefit from prosodic input.

1.3.1.2 Production

McNeilage (1997) suggests that in the babbling stage before a lexicon develops, hearing infants show an ability to reflect the ambient language in their babbling output (p.319). Moreover, the delay in the onset of well-formed syllables, canonical babbling (i.e. strings of alternating vowels and consonants), and reduced babbling repertoires in deaf infants is, according to McNeilage, contrary to Lenneberg's 'innatist perspective' (p. 316) which claims that the onset of babbling is not dependent on auditory experience. This is also contrary to Locke who suggested that sounds produced in normal babbling are independent of the ambient language environment. Subsequently, studies have shown the effects of ambient language on infant productions from 8 months (p. 317). McNeilage suggests that an infant's ability to imitate adults at the beginning of babbling when there is no lexicon provides evidence of a pre-speech relationship between input and output. Juscyzk (1997) also addresses these issues stating that since the 1970's studies have provided evidence that childrens' first words are a continuation of babbling, and that the ambient language influences the production of prosodic patterns. Reports showing that hearing babies begin canonical babbling between 6-10 months while it is delayed in deaf babies to between 11-25 months indicate that babbling does not develop normally in the absence of auditory input (p.172). Although Clement, den Os and Koopmans-van Beinum (1996, p.10) found interpretation of the results of some previous studies difficult due to differences in definitions of babbling and lack of clear information on the degree of hearing loss, they state that no canonical babbling was found in deaf infants by Oller and Eilers (1988) before 11 months.

According to Lieberman (1986) there are similarities between new-born cry and adult speech such as terminal fall in F_0 and amplitude, longer duration of expiration than inspiration phase, and level F_0 in the non-terminal portion of a breath-group. This provides evidence of some innate biological mechanism which controls subglottal pressure during phonation. He also states that physiological limitations in early infancy prevent babies from regulating subglottal pressure for long breath groups, and the steady declination of F_0 described in previous studies is not observed.

McNeliage (p.310) outlines three sub-stages of development identified in the literature. In the stage 1 pre-babbling period 0-7mths: (i) closed mouth phonation giving the impression of a syllabic nasal; (ii) (2-4 months) response to smiling with phonation and velars first as single sounds and later as a series; (iii) vocal play with regular syllable timing, manipulation of pitch (squeals and growls) and loudness (yells and whisper). McNeiliage (p. 310) also cites studies which report that 2-5 month old infants showed approaches to the imitation of the absolute value of adult fundamental frequencies (e.g. Papoušek and Papoušek, 1989), and where 4-5 month infants were observed to imitate formant patterns in /i/ and /a/ vowels with rise-fall pitch contours resembling an adult's. However, it is reported that the infants had higher fundamental frequency because their vocal cords are shorter (Kuhl and Meltzoff, 1982).

In a study of the development of deaf and normally hearing infants, Clement et al. (1996) report that there were no clear differences in mean fundamental frequencies (F_0) between 3 normal hearing and 3 profoundly hearing impaired subjects aged between 5 and 10 months. The authors suggest that the development of mean F_0 at this stage is determined by anatomical and physiological growth rather than hearing status. However, differences were found at the articulatory, durational and syllabic level which Clement et al. conclude was due to the lack of auditory feedback (p.17).

In the Stage 2 babbling period at 7-10 months the normal hearing infant begins to babble, and the opening and closing of the mandible, provides a universal motor basis for rhythmic patterns in speech (McNeilage, p. 311; Juscyzk 1997, p.175). Reduplication of the same syllable occurs from 7-10 months and variegated babbling using various consonants and vowels in multisyllable words occurs from 10-12 months (McNeilage, p. 315).

Cruttenden (1997, p.166) outlines four periods in infant vocal development with some overlap between them: i. *Crying* (birth – 3 months ii. *babbling* (3 months – 1;0 year); iii. 1 word period (1;0 year – 1;9 year); iv. *2 word period* (1;9 years – 2;0 years). During the babbling period around 8 months *imitation of adult intonation patterns* (high level and mid level) in English phrases such as *all gone!* can occur, and Cruttenden suggests the infant uses pitch as if learning a tone language. At the end of babbling and beginning of the 1 word stage *'jargon intonation'* or whole sentence intonation may be produced (p.166-7). During the one and two word periods *rises* are reported during counting, echoing, listing, questioning, attention seeking and *a high fall* is used to express surprise and insistence. A child can vary *nucleus placement* when he has developed two word sentences and by the time he has three or four word sentences he can vary the nucleus to indicate *old information*. However, Cruttenden points out that although some aspects of intonation develop early, children of ten years still have difficulty with intonational meaning (p. 168).

According to Vogel and Raimy (2002), as soon as children acquire word order they can assign phrasal stress at the right edge in SVO (subject + verb + object) languages such as English (p.229). They also state that although in English, compound stress is rule governed and stress is assigned to the first member of a compound, correctly produced compounds by 2 year olds in previous studies might be due to a tendency to stress new items of information (usually the first member of a compound, p. 230).

In a comprehensive review of the development of intonation (Snow and Balog, 2002) the development of intonational meaning is reported to begin at 10 months. Before that (i.e. 4 – 8 months) infants are reported to use gesture and prosody to express pragmatic intention and affective meaning (p. 1046) such as interaction in utterances directed at mother, strength of emotion (pitch height), call cries associated with high anxiety and high F_0 when mother is absent from the room. Vocalizations during shared experience accompanied by rising intonation and eye contact indicate that a response is required, whereas vocalizations without eye contact while the infant is manipulating a toy indicates no response required. During the single word period there seems to be a shift from the universal physiological and emotional associations with F_0 to a linguistic system and grammatical system. A predominance of falling intonation is noted in the first 3 – 9 months of life because of the physiological demands of rising intonation but from about 8 months infants begin to reflect the ambient intonational and rhythmic characteristics and frequency of rises and falls of their native language. However it is suggested that the complexity of different rises i.e. a simple rise in French and more complex fall-rise in English may account for more rises produced by French children.

To summarise, studies discussed above suggest that during the language acquisition process prosodic patterns produced by hearing infants are influenced by their ambient language environment. Onset of canonical babbling occurs between 6 and 10 months, and the first words are a continuation of babbling. By the one to two word stage children can imitate adult intonation patterns and produce rising intonation. At this stage they are also capable of varying nuclear placement and by the three to four word stage children can vary the nucleus to convey new information. Lack of auditory input puts deaf children at a disadvantage in the acquisition process and canonical babbling is delayed with onset occurring between 11 and 25 months. The main consideration in the present study is whether in the absence of adequate pitch information children with cochlear implants can rely on other acoustic and paralinguistic information (e.g. timing, loudness, gesture, facial expression) during prosodic development.

1.3.2 The school years

1.3.2.1 Perception

Limited previous research on the acquisition of compound vs. phrase stress led Atkinson-King (1973) to carry out an investigation of 285 normal hearing children aged 5;0 -13;0 years in the US. The results of this study show that the ability to identify compound or phrase stress is not acquired until late in the language acquisition process, and may develop gradually up to 12;0 years. In contrast with this, Ashby (1992) reports perfect discrimination between compound and phrase stress by two children aged 5;8 and 8;2 years.

Results of a study by Doherty, Fitzsimons, Assenbauer and Staunton (1999) show an overall improvement in the ability to discriminate between phrase and compound pairs, questions, statements and commands across the age range in a group of 37 school-going Irish children (aged between 5;5 and 8;5 years). This study also suggests that ability to discriminate differences in vocal affect or emotional prosody may take longer to develop.

Cutler and Swinney (1987) studied response times in the detection of accented and focused word targets in young children. In the first experiment accented (i.e. prominent) and unaccented versions of target words (e.g. ball, my, mat) were presented in sentences to two groups of children (21 in total) aged 4;0 -7;11 years. Both groups had difficulty with pronouns or function words but the authors state that according to the acquisition literature, word recognition processes for these words do not develop until after age 7;0. The younger group (aged 4;0 - 6;0 years) showed no significant effect of accent. In the second experiment the sentences were scrambled syntactically but the target words occurred in the same position in the list as in the first experiment. Two versions without sentence prosody were presented to ten subjects aged 5;0 -7;1 years with the target words stressed in one and unstressed in the other. Results show a significant effect for word class and stress level and the authors suggest that at this age children rely on lexical semantics whereas in the first experiment lexical semantics were not affected by varying accent or sentence semantics for this age group. In a third experiment higher level processing of sentence semantics was investigated in children aged 3;0– 6;0 years in stories where focus was determined by questions preceding the sentences. Although the focus effect was not significant for the group the results for individuals show that it does appear with age. When divided into three groups the focus effect was significant for the 5 year-old group but not for younger groups. Overall results of these experiments show that a processing advantage for focus words is not fully developed in pre-school children and is acquired before the ability to process accented words between age 4;0 and 6;0 years.

A similar study to Atkinson-King (1973) was carried out by Vogel and Raimy (2002) to investigate the role of prosodic constituents in the acquisition of compound and phrasal stress by 40 children ranging in age from 4;9 and 12;3 years. Their results show a gradual increase in percentage correct scores in the distinction between these contrasts up to 12;0 years and are in general agreement with Atkinson-King (1973). However, Vogel and Raimy’s percentage correct scores for the older group were lower (74%) than for the corresponding group in Atkinson-King’s study (100 %).

Vogel and Raimy suggest that the lower scores in their study might be due to the inclusion of a set of novel compounds and differences in scores for known and unknown items for all ages. It was suggested that better scores in the Atkinson-King study might be due to a training component before the test. Vogel and Raimy also observed a preference for compounds by children aged 4;9 to 7;7 years for known items regardless of stress patterns, but by 7;0 years subjects were beginning to become sensitive to patterns they knew. When the distinctions between compound and phrasal patterns were recognised they were not generalized to novel items because there were no lexical entries for them to be matched with (p.241).

A study of more than 120 British children aged 5;0 -14;0 years was carried out by Wells, Peppé and Goulandris (2004) who investigated perception/comprehension (and production) skills using the test battery PEPS-C i.e. Profiling Elements of Prosodic Systems–Child version (Peppé and McCann, 2003). According to the authors there is limited previous research into prosodic perception over this age range. However, some previous studies cited have conflicting reports on children’s abilities to match pictures to identical phrases with different phrase boundaries (chunking), or to identical sentences with focus on a different lexical item. The results of the study by Wells et al. indicate that in the chunking perception/comprehension tasks there was considerable variation between individual children. Between ages 5;0 and 11;2 years, performance in chunking tasks correlated significantly with subtests of receptive and expressive language measures such as the TROG (Test for Reception of Grammar, Bishop, 1989) and the CELF (Clinical Evaluation of Language Fundamentals-Revised, Semel, Wiig and Secord, 1987). One of the chunking tasks involved matching pictures to a compound (coffee-cake) or two nouns (coffee, cake) and the results show improvements between 5 and 10 year-old groups. In the focus test, understanding the use of accent /focus to highlight a key element in a sentence was found to lag behind the children’s ability to use the appropriate phonetic feature in their own speech. The fact that not all children performed at ceiling in all cases suggested to the authors that some aspects of intonation may be acquired later than the age ranges covered (5;0–14;0 years), or might never be acquired even in adulthood (Peppé, Maxim and Wells, 2000).

1.3.2.2 Production

Atkinson-King (1973) carried out a study of the production of unemphatic stress in compounds and phrases (e.g. *blackboard* versus *black board*) in 300 children aged 5;0-13;0 years. Although the majority of young children were unable to produce compound versus phrase stress and tended to place primary stress on the first syllable, even the youngest children could imitate without difficulty and were able to make a contrast when minimal pairs were produced one after the other. At a later stage they learned to produce each one in isolation and results show that the ability to distinguish between compound and phrase stress is acquired gradually as a function of age. Atkinson-King suggests that younger children are more likely to store learned lexical items first and the rules of stress placement are acquired later. She concludes that stress contrasts were acquired in a particular order i.e. imitation, comprehension and production. Children who were successful with production tasks had no difficulty with comprehension but the reverse was not always the case.

In a comprehensive study of intonation development in 193 children aged between 5;0 and 13;0 years Wells, Peppé and Goulandris (2004) used the PEPS-C (Profiling Elements of Prosodic Systems-Child Version) to investigate production skills. They found that some aspects of intonation such as chunking, affect and focus were established in 5 year-olds and results supported findings in some previous studies. However, they conflicted with Katz, Beach, Jenouri and Verma (1996) who reported that 5 –7 year-olds in their study did not use phrase boundary cues such as pause and duration in an adult way for grouping (chunking) of objects. Wells et al. suggest that differences in the findings may be attributed to the fact that subjects in their own study had to make a lexical (compound versus string of two nouns) rather than a syntactic [(*pink and green) and white*] versus [*pink and (green and white*)] distinction in a study by Katz et al. (1996, p.3181). They also found that some functional prosodic contrasts which were more difficult for some younger children were acquired by most 8 year-olds. For example, some of the younger children had difficulty incorporating two words (*coffee, cake*) into a single intonation phrase in a compound (*coffee-cake*), and they also had diffculty producing a rise pitch on particular syllables for questioning or a fall-rise to indicate ‘not-keen’. They also had a preference for utterance final position in the placement of focus. Wells et al. (2004) also found variation in all the age groups with some 5 year-olds reaching ceiling and some 10 year-olds still performing at chance level. Wells and Local (1993) suggest that other intonational functions such as maintaining or signalling the end of a conversational turn may compete with focus and accent placement in young children as a result of delayed or immature prosodic development (p.71). Unlike Atkinson-King (1973), Wells et al. (2004) found that focus production skills lagged behind focus comprehension skills and their results support some previous studies (e.g. Cutler and Swinney, 1987; Vogel and Raimy, 2002).

Dankovičová, Pigott, Wells and Peppé (2004) investigated temporal boundary markers in a subset of the data in Wells et al. (2004). Acoustic analysis of pause duration and phrase final lengthening in two versus three items (e.g. *coffee-cake and tea* versus *coffee, cake and tea*) produced by ten 8 year-old children using picture prompts was combined with adults' perception of the productions. Overall results show that the children's use of boundary markers was in the right direction and pause was found to be a more salient boundary marker than phrase-final lengthening. However there was considerable individual variation across children, and the authors suggest that further investigation needs to be carried out to establish the relationship between temporal markers and pitch cues. Three groups were identified in the data: a) *accurate and unambiguous* (where the system was considered to be acquired); b) *accurate but ambiguous* (where the contrast was not perceived by listeners); c) *inaccurate and ambiguous* (where children were at a more immature stage of development).

1.3.2.3 Developmental issues relating to the production of stress and intonation by deaf children

For children with severe to profound hearing losses prosodic development is delayed and studies of hearing aid users show different rates of development in production for individuals. For example, Abberton, Fourcin and Hazan (1991) report on fundamental frequency range and intonation development in four severe to profoundly deaf children (aged between 7;0 and 8;0 years) with pure tone average HL ranging from 83 dB to 115 dB). The four hearing impaired children showed different patterns of intonation development over a four year period. Although progress was slow and delayed these children did acquire linguistic pitch control. Two children with 83 dB and 90 dB hearing loss learned to use a range of tones for syntactic or attitudinal purposes as well as rising intonation. Although more delayed the other two children (112 dB HL and 115 dB HL) developed better pitch control and one of them was beginning to produce rising intonation.

Most and Frank (1994) carried out a study of 63 severe to profoundly hearing impaired children (aged between 5;0 and 12;0 years) with average hearing loss ranging from 80 dB to 110 dB, and a group of normal hearing subjects was also included. Spontaneous productions of questions and statements as well as imitations of nonsense syllables and imitations or reading aloud of sentences were recorded and analysed. Results show that in spontaneous speech the older hearing-impaired subjects were different from the normal hearing group in their production of question intonation. The ability to produce appropriate intonation by the hearing impaired subjects seems to develop during between 6;0 and 9;0 years.

More recently Titterington, Henry, Kramer, Toner and Stevenson (2006) investigated weak syllable processing in school age children with cochlear implants. Results suggest that the group of implanted children had a similar prosodic hierarchy to the group of language matched normal hearing children. They showed a preference for footed weak syllables (i.e. in a strong/weak or trochaic template) which influenced the effects of delayed access to audition on the development of linguistic processing and short-term memory. The authors conclude that difficulties associated with perceptual salience cannot fully account for differences in the processing of footed and unfooted weak syllables, and that the influence of prosodic foot structure on the omission of some weak syllables (e.g. in **ba**nana) has not previously been considered for children with cochlear implants (p.263). The normal hearing group (aged 3;0 – 13;0 years) in this study showed increasing ability to process unfooted weak syllables as age increased whereas processing of footed syllables was equivalent across all ages. Despite the fact that English-speaking children are generally reported to use a trochaic template up to age 3;6 years, the language-matched normal hearing subjects in Titterington et al. (aged between 3;6 – 5;8 years) processed footed over unfooted weak syllables when memory load was high (p. 264). Although not central to the current investigation, these results have implications for weak syllable perception and the development of appropriate rhythmic patterns in the speech production of children with cochlear implants.

1.3.2.4 The relationship between perception and production

Cutler and Swinney's experiments (1987) also discussed earlier support other previous investigations by showing that hearing children aged 5;0 or 6;0 years are poor at exploiting prosodic information in language comprehension. Although in general pragmatic and semantic abilities are thought to develop in parallel in 4 – 6 year-old children (p.162) the authors suggest that prosodic development is different. Studies are cited which show that 4 – 6 year-old children cannot process semantic or pragmatic information e.g. *given* versus *new, topic* versus *comment* in production or comprehension, but that they can produce appropriate accentuation to convey new information or focus. According to Cutler and Swinney (p.163) a universal physiological explanation for this 'paradox' is provided by Bolinger (1983) who states that a semantically interesting word generates greater tension and excitement in a speaker which leads to the rise in pitch in accented words. Productions of 3 – 4 year-old children are apparently similar to productions of 5 – 6 year-old children. However, the former are just a physiological reflex and not due to prosodic competence, and the latter are producing accent patterns with a prosodic production system interacting with discourse level factors. Wells et al. (2004) also conclude in their study that children may be able to produce accent and focus in their own speech before they can interpret accent and focus in other speakers and the results support the findings of Cutler and Swinney (1987) above. However, as suggested by Juscyzk (1997, p.183) individual differences in prosodic development might also be influenced by different learning styles in children such as an analytic approach (focus on vowels and consonants in words) rather than attention to stress and intonation in multisyllable utterances.

There seems to be a consensus supporting the gradual acquisition of the stress and intonation contrasts in the studies discussed above for English for normal hearing children and that development is delayed for hearing aid users. The issues discussed above are particularly relevant to the current investigation of the perception of compound versus phrase stress and focus in Experiments II and in the production of focus by children using cochlear implants in Experiment III. As the studies of normal hearing infants and school-going children indicate, pitch seems to be an important cue to the perception and production of stress and intonation. However, in the absence of adequate pitch information through current speech processing strategies, children with cochlear implants will have to rely on other cues such as timing, loudness and paralinguistic cues during prosodic development. This issue is investigated in the current perception and production experiments.

1.4 The perceptual and physical correlates of stress

1.4.1 Acoustic cues to stress and intonation

Limitations of current speech processors in delivering adequate pitch information (section 1.7 below) have implications for how stress and intonation contrasts are perceived by cochlear implant users, and it is possible that other perceptual cues such as timing and loudness are particularly important. The relative importance of the acoustic correlates of stress for normal hearing listeners is discussed in this section. Generally the terms 'pitch' and 'F_0' refer respectively to the perceptual and physical correlates of stress, but they are used interchangeably in some of the studies mentioned in the present discussion. Although the terms 'intensity' and 'amplitude' refer to different physical quantities, these terns are often used interchangeably, and when amplitude and intensity differences are expressed in decibels these difference measures are equivalent. Experiments with normal hearing speakers have shown that the physical parameters of stress (i.e. F_0, duration, and amplitude) contributed to the perception of stress. Some studies have suggested that F_0 provides the most important cue (Fry, 1955, 1958; Lehiste, 1970; Gay, 1978a, 1978b; Ladd, 1996). There is a physiological relationship between increased subglottal pressure from the lungs and both increased vocal amplitude and the frequency of vibration (F_0) of the vocal folds. Although other factors can also change F_0, an increase in F_0 is often accompanied by an increase in amplitude (Gay, 1978; Borden, Raphael and Harris, 1994).

In Fry's 1955 study listeners were presented with noun and verb forms of words such as *subject*, *digest*, *permit* and asked whether they heard the stress on the first or second syllable. Results show that when a syllable was long and of high intensity it was perceived as strongly stressed and when it was short and of low intensity it was perceived as weakly stressed. The results of Fry's 1958 study show that F_0 differed from duration and intensity in that it tended to produce an 'all-or-none effect'. The fact that there was a change in frequency was more important than the magnitude of the change (p. 151). When intensity and duration were studied separately, duration was the overriding cue. These findings have been confirmed by later studies although failure to include intrinsic vowel intensities in one early study by Bolinger (1958) was noted by Lehiste (1970, p.128). Lehiste maintains that because vowels have different intrinsic intensities (Lehiste, 1970; Fry 1979), intensity can only be regarded as a reliable cue to stress where two syllables are intrinsically identical and vowel quality remains constant as in PERvert vs. perVERT. Generally, however, noun/verb pairs like this are not segmentally identical. For example in IMport vs. imPORT the intrinsic intensity of the open vowel /o/ in IMport for speakers in Irish English or /ɔ/ for speakers of British English might obscure increased intensity on the /ɪ/ vowel in the stressed syllable (see the relative intensities of English consonants and vowels in Fry, 1979, p.127). There is a similar connection between vowel quality and fundamental frequency (F_0) associated with it. If other factors are kept constant, high /i/ and /u/ have higher intrinsic F_0, and open vowels such as /a/ are associated with lower intrinsic F_0. F_0 at the peak of the F_0 contour averaged across five speakers was 183 Hz for /i/, 182 Hz for /u/, and 163 Hz for /a/ (Lehiste 1996, p.233). However, the effects of intrinsic F_0 are probably compensated for perceptually by listeners (Silverman, 1984), and are unlikely to affect the importance of pitch as a cue to stress.

Fry's experiments are also reviewed by Gay (1978a, 1978b) in the light of his own investigations. He concludes that production differences in amplitude, fundamental frequency, and first and second formant frequencies between stressed and unstressed syllable pairs were preserved across fast and slow speaking rates. Vowel duration differences, however, were not so great for the faster speaking condition, and for two speakers vowel duration in the faster speaking rate was the same in stressed and unstressed pairs. The possibility that duration might be independent of the other cues was investigated in another experiment by Isenberg and Gay (1978) involving the perception of stress in isolated disyllables OBject vs. obJECT. The results show a trade off between duration and the other cues where F_0, intensity and spectral differences in a comparison syllable of fixed duration were more reliably perceived when duration was manipulated in the other variable syllable.

In a review of the above and other related studies Ladd (1996) suggests that if words in citation form such as perMIT and PERmit become questions then it can no longer be said that the noun/verb contrast is cued by a pitch peak. If these words are put in a longer sentence after the main intonational peak of the utterance, the word is not cued by pitch differences in the contour but yet the stress differences between the two patterns can be heard. He also states that autosegmental metrical (AM) theorists are critical of an approach which regards stress as ‘simply a scalar phonetic property of individual syllables’ (p.47). AM theorists make a distinction between utterance level stress and intonational accent. They claim that there are different degrees of prominence between the elements of the utterance and that in addition, there is an intonation pattern which consists of pitch accents and edge tones i.e. phrasal or boundary tones. Ladd concludes that duration, intensity and spectral properties, if properly measured, could be reliable indicators of stress in English (p.59).

1.4.2 How important is F_0 in the perception of stress and intonation?

A major consideration in the current study is how important F_0 is in signalling stress and intonation contrasts to listeners and whether speakers vary in the use of acoustic cues in order to convey different stress and intonation contrasts. This issue is investigated in Experiment I (Chapter Two) and Experiment II (Chapter Three) in the present study. In Experiment I non-meaningful pairs of synthesised stimuli with syllable 1 and syllable 2 stress (e.g. BAba vs. baBA) are presented to both implanted and normal hearing children with controlled changes in F_0, duration and amplitude.

Compound vs. phrase stress

In Experiment II, however, words with compound vs. phrase stress are presented in a carrier phrase i.e. *give me the BLUEbell* or *give me the blue BELL.* The carrier phrase is identical for all items presented so sentence intonation does not vary and the target item is always in final position to reduce the memory load for implanted children. Lexical stress in compounds vs. noun phrases is signalled by primary stress or accent i.e. in the first element in *BLUEbell* and in second element in *blue BELL.* According to Cruttenden (1997) primary stress/accent refers to the main pitch prominence in an utterance. However, results of a study of prosodic variation in adult speakers of Southern British English (Peppé, Maxim and Wells, 2000) show that differences between compounds and simple nouns may not always be signalled in the same way for different speakers. For example in a chunking production task the majority of speakers were able to make a distinction between the compound (*creambuns*) and simple nouns (*cream, buns,* and *jam*) but pitch movement and pitch reset were not as reliable at signalling differences as lengthening and pause. This would suggest that implanted children might have less difficulty hearing these contrasts produced by some adults if they were differentiated mainly by timing cues and the current study should provide information on perception of compound and phrase stress by normal hearing children up to 17;11 years. Since it is reported in previous studies that normal hearing listeners acquire these lexical contrasts gradually (Atkinson-King, 1973; Wells, Peppé and Goulandris, 2004) it is likely that implanted children might acquire these contrasts later. Performance in the present perception tests by the implanted children is likely to be influenced by level of prosodic development as well as hearing ability.

Focus

In the general intonation literature (see section 1.2) it is suggested that contrastive items have a steeper fall in pitch (Chafe, 1974; Brown et al. 1980; Brazil et al. 1980). Ladd (1996), for example, suggests that words can be in focus for various reasons and are marked by pitch accents, and corrective, narrow or contrastive focus (Gussenhoven, 2006) are signalled by pitch accents in the intonation contour. There seems to be an accepted view that when narrow focus is conveyed to a listener it is signalled by pitch adjustments i.e. increase in F_0 peak, followed by a high fall as well and increases in duration and intensity. Xu and Xu (2005) suggest that in English focus modifies the pitch ranges of F_0 peaks and valleys which are already there and the characteristics of F_0 peaks on stressed syllables are determined by narrow focus with pitch adjustments such as 'expansion under focus, compression after focus, and little or no change before focus' (see section 1.2). Peppé, Maxim and Wells (2000) also report in the study of speakers of Southern British English mentioned above that there can be variation in how individuals signal narrow focus. When focus was conveyed to a listener a falling glide occurred on the focus item for most subjects but there were differences in how other phonetic exponents were used e.g. silence, lengthening, loudness and pitch-reset. The authors concluded that their study indicated that there may be differences in the phonetic realization of intonational contrasts in less controlled social situations compared to laboratory conditions. However, there were some cases where all the accented words sounded prominent, and broad rather than narrow focus was conveyed. Others had 'dual' accents i.e. a pre-final accent for focus and a final accent indicating end of a turn. (See earlier discussion of a single *nucleus* on the last accented syllable in section 1.2.1). The authors conclude that there are variations in how pre-final focus is conveyed to listeners by adults.

This issue is also raised by Kochanski, Grabe, Coleman and Rosner (2005) who carried out quantitative measurements of accented syllables in a large corpus of natural speech in the IViE project (Intonational Variation in English) (including Belfast and Dublin). Contrary to widely held views in the intonational literature (mainly based on laboratory speech) that F_0 is a major cue to prominence, the authors concluded that accent and prominence is marked by loudness and duration cues and that F_0 plays a minor role. They state that none of their subjects used large excursions of F_0 previously associated with prominence in the general literature, and loudness was a better predictor of prominence. However, mean age of the subjects was 16;0 years and they were still in secondary school. In the analysis functional distinctions were not made between lexical stress, focus or other contrasts, so results are difficult to compare with other studies where specific contrasts are elicited. The authors conclude that they do not disagree that F_0 changes can cause speakers to perceive prominence. F_0 (and duration and amplitude) measurements will be carried out for the focus stimuli presented in Experiment II for the normal hearing talkers in the perception tasks as well as the focus production data for the implanted children in Experiment III. The importance of F_0 in signalling focus to normal hearing and implanted listeners will be discussed and general issues for consideration are whether

(i) F_0 adjustments by the talkers in Experiment II are big enough to signal focus to implanted listeners
(ii) F_0 adjustments by CI talkers in Experiment III are big enough to signal focus to a trained listener
(iii) whether normal hearing or implanted talkers use other cues to signal focus such as amplitude and/or duration in combination with F_0 or instead of F_0

1.4.3. Theoretical basis for acoustic analysis of the production data in the current study

There is an extensive literature on different frameworks for representing intonation in normal speech (Cutler and Ladd, 1983; Ladd, 1996; Xu and Xu, 2005) which can be adapted to capture erratic, monotonous or inappropriate F_0 contours in the speech of deaf speakers (O'Halpin, 2001). Some deaf talkers have difficulties co-ordinating respiratory and laryngeal muscles which lead to rhythmic problems (La Bruna Murphy, McGarr, and Bell Berti, 1990), inappropriate pausing and the absence of a gradual decline in F_0 across a sentence (Osberger and McGarr, 1982). This in turn contributes to what listeners perceive as monotony or excessive pitch variation and inappropriate intonation (Monsen, 1979; Allen and Andorfer, 2000). Previous studies with deaf children with hearing aids report some improvements after a training period using visual displays with F_0 and intensity displays but carry-over into spontaneous speech has been limited (Abberton, 1972; Boothroyd, 1973; King and Parker, 1980; McGarr, Head, Friedman, Behrman and Youdelman, 1986; Youdelman, MacEachron and McGarr, 1989; McGarr, Youdelman and Head, 1989; Mahsie, 1995; Spaii, Derkson, Hermes and Kaufholz, 1996). Improvements following cochlear implantation have been reported for different aspects of speech production and perception in children (Waltzman and Cohen, 2000; Svirsky, Teoh and Neuburger, 2004). However, to date there have been no systematic studies involving detailed acoustic analysis of intonation abilities for English speaking implanted children and the present study is the first attempt to do this.

Declination

One aspect of intonation relevant to the present investigation is a universal tendency for F_0 to decline across utterances (Vaissiere, 1983; Cruttenden, 1997; Ladd, 1996); Lieberman, 1986). Different approaches to measuring declination (Cooper and Sorensen, 1981; Thorsen, 1983; Cutler and Ladd, 1983; Ladd, 1993, 1996) involve drawing abstract lines through accent peaks in an overall F_0 contour, and experiments have shown that in shorter sentences rate of declination is often more rapid whereas declination slope is less steep over longer domains (Ladd, 1996). For some speakers, F_0 may increase rapidly at the beginning of a sentence and then either remain flat or decline more slowly at the end. However, in a different approach proposed by Pierrehumbert (1980) and Beckman and Pierrehumbert (1986) accents are scaled above a declining baseline, and they are more concerned with levels and tone sequences rather than the overall F_0 contour. The accent peaks are downstepped so that each one is a constant proportion of the previous peak. Downstepping is also referred to as deaccenting or distressing of old information (Ladd, 1980). More recently Xu and Xu (2005) investigated the phonetic realization of focus for normal hearing talkers and their model simplifies the different approaches described above by taking into account both communicative and articulatory aspects of F_0 variation. They suggest that focus determines the characteristics of F_0 peaks which are already present in an utterance by increasing the size of the F_0 peak and lengthening the duration of the stressed syllable (see also under *Focus* in section 1.4.2).

Representing F_0 contours for NH and CI talkers in the current study

The present study draws on the approaches to measurement referred to above involving drawing abstract lines through F_0 peaks but is remains to be seen whether typical F_0 contours or attempts at conveying focus appropriately can be adequately captured for CI talkers (Experiment III in Chapter Four). Scaling accents and F_0 peaks above a declining baseline might be difficult for deaf talkers if there is frequent pausing, erratic or monotonous F_0, or inappropriate F_0 peaks, but it is a useful way of showing any improvements or change in F_0 control following training or cochlear implantation. For the normal hearing talkers in the current study the first accented word *DOG* may be in focus in the sentence *the DOG is eating a bone* and a step-up to a boosted F_0 peak would be expected on *DOG* followed by a more striking decline in F_0. However, if focus occurs later in the sentence on *EATing* or *BONE* for example, declination can be reset or suspended earlier in the sentence. F_0 can start low, decline gradually, and rise again in anticipation of the boosted F_0 peak later in the sentence. Deaf talkers with breathing problems and difficulty controlling F_0 can also have excessive pausing or excessive duration of syllables which can result in inappropriate pitch reset, a noticeable absence of F_0 decline across utterances, and inappropriate or absence of F_0 peaks normally associated with stressed or accented syllables. For examples and more detailed discussion of these issues and examples of stylized graphs for hearing and deaf subjects pre- and post training see O'Halpin (1993, 1997, 2001). In the present study acoustic measurement of F_0, duration and amplitude for children with cochlear implants and normal hearing talkers are presented in stylized line graphs in Chapter Four. The rationale for analysis of the production data is discussed in section 4.3.

1.4.4 Acoustic cues in the production of stress and intonation in Southern Hiberno English

Very little attention has been paid to Southern Hiberno English intonation but research to date reports that falling nuclear tones (H* + L %) for declaratives were produced by 16-18 year old school-going subjects in Dublin (Grabe and Post, 2002) and are different from the rising tones (L* + H%) reported for Belfast English (Rahilly, 1991, 1997, 1998; Grabe, Post, Nolan and Farrar, 2000; Lowry, 2002). In another preliminary investigation of contrastive stress (O'Halpin, 1994) two adult speakers in Dublin produced falling tones in accented syllables but focus or contrast was not always conveyed to a trained listener possibly due to smaller boosted F_0 peaks on target words especially in final position, and although both speakers had increased duration and intensity of these words it did not always contribute to the perception of focus.

The variation and ambiguity in this study would support Peppé, Maxim and Wells (2000) for SBE speakers. Other varieties of Southern Hiberno English have not yet been investigated but in a study of Irish Dalton and Ní Chasaide (2003, 2005) reported rising tones in Ulster Irish and falling tones similar to the Dublin Hiberno English pattern were reported for Irish in Southern Connaught, Kerry and Mayo. According to the authors it remains to be seen whether there are similar patterns to be found in matching dialects of Southern Hiberno English. Differences in the studies discussed above such as age of the subjects, variety of English and how focus is elicited (spontaneous, semi-spontaneous or in laboratory conditions) may affect results so it is difficult to be conclusive. In the present study only stimuli which are unambiguous and convey focus on the target item to a trained listener (i.e. the author) will be presented to the normal hearing and implanted children. Acoustic measurements of these stimuli and additional data for the same talkers which will be carried out in Chapter Four will confirm the patterns reported above for Dublin English i.e. whether they convey focus in the same way as described for other varieties of English.

1.4.5 Acoustic cues to stress and intonation in the speech of normal hearing and deaf children

Few studies of intonation in normal hearing children are specifically concerned with focus. However, issues raised in studies of other aspects on intonation are relevant to the acoustic analysis of the production data in the current study in Experiment III. For example, Patel and Grigos (2006) found differences between 4, 7 and 11 year-old children in their production of statement-question contrasts. The 4 year-olds used modified duration, the 7 year - olds used F_0, duration and intensity, and the 11 year-olds used more F_0 and less duration and intensity which was similar to adults. Snow (1998, 2001) reported that 4 year-olds in his study differed from adults in that they lengthened the duration of final syllables (i.e. FSL final syllable lengthening) but had a narrower accent range than adults in sentence-final rising tones. The final lengthening produced by the children in Snow's study was accompanied by a narrow pitch excursion due to motor difficulties with rising intonation, whereas for adults a slower speed of pitch change is generally accompanied by wider pitch excursion. Although the current study does not involve question intonation it is possible that the step – up in F_0 or rise – fall associated with a focus item might be difficult to produce in final position especially against terminal fall or declining F_0. Wells et al. (2004) found variability in their study of 5 – 13 year-olds with some 8 year-olds still showing preference for utterance final position in the placement of focus, but they also observed a high incidence of ambiguity. As a final fall in F_0 also signals end of a turn or a sentence, the fall in F_0 may have been insufficient to signal focus to a listener.

Evidence from the experimental studies discussed in 1.4.1 for hearing subjects suggests that F_0 may not always provide an overriding cue to stress, and this may also be the case for deaf speakers. Rubin-Spitz and McGarr (1990), for example, investigated the perception of terminal fall in the speech of eight talkers aged between 8:0 and 18:0 years with pure tone averages HL (hearing loss) ranging from 98 dB to 118 dB. They were asked to read declarative sentences, and *why?* and *yes/no* questions with varying length and contrastive stress. The authors suggest that although listeners may sometimes perceive appropriately stressed syllables and falling terminal pitch contours to be produced, these may not be conveyed by the same acoustic correlates as for hearing speakers. Results show little difference in mean F_0 in declarative and non-declarative sentences, and in terminal falling contours there was also no difference in mean F_0 between these two sentence types. Listeners perceived F_0 contours to be flat in many cases where there was a terminal fall in F_0 and results suggest that contours which fall more quickly regardless of the amount are more likely to be perceived as falling. The authors conclude that there may be conflicting cues (i.e. duration or amplitude) which might affect listeners' perception of F_0.

Murphy, McGarr and Bell-Berti (1990) investigated stress contrasts produced by 13 deaf subjects ranging from 9;0 – 19;0 years with average pure tone hearing loss ranging from 92 dB to 118 dB. Spondaic words such as *cupcake* or *hotdog* were elicited with lexical stress alternating between the first and second syllable. Results show that stressed syllables produced by the deaf subjects tended to have increased F_0 and amplitude, and longer duration. However, if only one or two of these cues were present, the stress patterns were not necessarily judged as 'incorrect' (p. 89) by a panel of listeners. This study highlights individual differences in the use of acoustic cues by hearing impaired talkers.

Most (1999) reports on a study of syllable stress in 15 deaf 10 – 13 year-old Hebrew speakers with average pure tone hearing loss ranging between 82 dB and 125 dB. Results show that syllable duration in bisyllabic meaningful minimal pairs (similar to *`object* versus ob*`ject* in English) did not play an important role in listeners' perception of correct or incorrect stress production. F_0 and amplitude were higher in stressed than unstressed syllables for correctly perceived productions and the reverse was found for patterns which were perceived as incorrect (p.64).

In another study (O'Halpin, 1993, 2001) two 8 year-old deaf subjects (average pure tone hearing loss 96 dB and 100 dB) did not use F_0 or convey contrastive stress in declarative sentences before training and it was anticipated they might have used duration or intensity appropriately. The results, however, show that appropriate lengthening of target syllables was present but was obscured by inappropriate F_0 peaks on normally unstressed syllables. After a period of training only one of the subjects used similar strategies to a hearing subject with appropriate (but exaggerated) boosting of F_0, proportionate durational adjustments, and increased intensity in a structured task only.

Allen and Andorfer (2000) report that all three cues were used in falling and rising intonation patterns by six severe to profoundly deaf and six normal hearing children aged between 7;9 and 14;7 years. Both groups increased F_0 on the second syllable for interrogatives and decreased F_0 for declaratives, but the deaf group had larger mean durational differences between syllables. However, results suggest that the contrastive use of F_0, duration and amplitude cues was less pronounced for the deaf subjects, and statements and questions produced by them were not always correctly categorised by listeners (p. 452).

Other studies of hearing aid users suggest that falling contours are acquired before rising contours (Abberton et al., 1991; Most and Frank, 1994) or that conflicting cues (duration and amplitude) may affect listeners' perception of appropriate F_0 e.g. contours which fall more quickly are likely to be perceived as falling rather than level (Rubin-Spitz and McGarr, 1990). Although it has been reported that all three cues are used in stress and intonation contrasts by English speaking hearing and deaf children using hearing aids by aged 7;0 or 8;0 years it remains to be seen whether children with implants also use these cues in the same way. Some reports of deaf children suggest that even if F_0, duration, and intensity adjustments are appropriate they may not be sufficient to convey focus or contrast. Others suggest rising intonation is difficult for young normal hearing children especially in final position, and for English speaking deaf hearing aid users and Mandarin Chinese speakers falling tones are acquired before rising tones. These issues will be considered for the focus data in the present study and because of time constraints compound and phrase data for the children with cochlear implants will be analysed in a follow up study.

The deaf subjects in the studies cited above were hearing aid users and similar investigations need to be carried out for cochlear implant users to establish which cues are accessible to them in the perception of stress and intonation contrasts. In the absence of adequate pitch information through cochlear implants (section 1.7) they would have to rely more on other perceptual cues to stress such as timing and loudness. The issues raised in this section will be taken into consideration for the implanted children in the present study in the analysis of the speech perception results in Chapters Two and Three, and in the discussion of F_0, duration and amplitude measurements in the production of focus in Chapter Four.

1.5 Representation of the correlates of pitch in the acoustic signal

When the vocal folds vibrate in speech, a complex periodic wave is produced. The length of time a wave takes to repeat is known as its period. The period of repetition is expressed in seconds or milliseconds and the term *frequency* refers to the number of times that a periodic waveform repeats per second (cycles per second). The unit of measurement for frequency is hertz (Hz) and 1Hz, for example, corresponds to one cycle per second. Unlike a pure tone, which has only one frequency of vibration, a complex wave is composed of a number of component frequencies or overtones called harmonics (Denes and Pinson, 1993, pp. 17-45) which are integral multiples of the lowest frequency of pattern repetition or the *fundamental frequency* (F_0). The *pitch* we hear in speech is closely correlated to the fundamental frequency of a complex sound. Generally when the frequency of vibration is increased we hear a rise in pitch and when frequency is lowered we hear a decrease in pitch. However, fundamental frequency and pitch are not identical, as the frequency is a physical property that can be measured instrumentally whereas pitch is a sensation or psychological phenomenon which can only be measured by asking listeners to make judgements (Borden, Harris and Raphael, 1994, p.35-36).

1.6 Coding of pitch and loudness in the inner ear: acoustic stimulation in normal hearing

Decomposition of a complex wave into its component frequencies and amplitudes is referred to as Fourier analysis (Lieberman and Blumstein, 1988, p.26; Denes and Pinson, 1993 p.31; Johnson, 1997, p.13). In normal hearing, the cochlea performs a kind of Fourier analysis of a complex sound into its component frequencies. Frequency information is extracted by a combination of place location along the basilar membrane, and temporal information from the timing of neural impulses (Borden, Harris and Raphael, 1994, p.182). In the cochlea, each point on the basilar membrane (BM) is tuned, responding best to a particular frequency called a characteristic frequency (CF) which decreases from the *base* to the *apex*. The BM behaves like a number of bandpass *filters* which respond best to limited ranges of frequencies around the CFs.

In addition to place coding on the BM, frequency information can be obtained from neural synchrony or phase locking. The nerve spikes, which occur in response to a sinewave, tend to be *phase locked* or synchronised to the stimulating waveform for frequencies up to 4-5 kHz. A nerve fibre may not fire for every cycle but when it does, it occurs at roughly the same phase of the waveform each time. Thus the time interval between the spikes tends to be an integer multiple of the period of the stimulating waveform. Similarly, the resolved lower harmonics of a complex sound also have their own nerve spikes occurring at the same phase of the waveform each time (Moore, 2003, p.246).

Loudness, which is subjective and related to the physical level of sound, appears to be coded according to overall neural firing rate in the nerve. Neurons can have high, medium or low firing rates but above a certain level become saturated and do not respond further increases in sound level. The *dynamic range* (difference between threshold and saturation) is only 10-30 dB for neurons with high firing rates whereas neurons with low and medium firing rates have a wider dynamic range. For neurons with medium and low firing rates, firing rate increases rapidly at first with increasing sound level, and then firing rate continues to increase gradually with increasing sound level over a wider range of levels. For high sound levels, which could be up to 120 dB, neurons with low firing rates and wide dynamic range play an important role (Moore, 2003, p. 246).

1.7 Coding of pitch and loudness in cochlear implants: electrical stimulation

In cochlear implants an array of electrodes is implanted into the cochlea. The electrical signal stimulates the auditory nerve at selected places along the electrode array, and mimics the place coding of the basilar membrane (BM) described above through a filter bank or explicit Fourier analysis. As mentioned in section 1.6, in normal hearing the lower harmonics are resolved and separated on the basilar membrane.

However, in cochlear implants, the frequency range in any one channel generally covers more than one harmonic for fundamental frequencies typical of speech, resulting in unresolved lower harmonics. In cochlear implants, increases in pulse magnitude or duration results in increased neural spike rates in the auditory nerve and in increasing loudness (Moore, 2003, p.246). Because the BM is bypassed in electrical stimulation there is no natural compression and spike rates in single neurons can exceed the maximum rates found in acoustic stimulation resulting in large changes in the sensation of loudness. The dynamic range from threshold to discomfort is only 3-30 dB which is very limited compared to acoustic hearing (up to 120 dB). In cochlear implants the incoming signal for an everyday sound is compressed after it is band-pass filtered into different frequency bands which are then mapped onto electrodes in accordance with place coding in the normal BM.

In speech processors generally, the output of a set of band-pass filters is rectified and smoothed (low-pass filtered) to remove faster fluctuations due to higher frequencies, resulting in an approximation of the amplitude envelope. If the smoothing cut-off frequency is above the F_0 in speech, then F_0 appears as a temporal fluctuation in the speech envelope waveform (Moore, 2003; Guerts and Wouters, 2001; Rosen and Howell, 1991). In a common speech processing strategy such as CIS (continuous interleaved sampling), carrier pulse trains, which are modulated by the extracted speech envelope, are delivered to each electrode at a fixed rate of around 1000 pulses per second (pps). Physiological and psychophysical evidence suggests that to get a good representation of $F_{0,}$ the carrier pulse rate should be 4-5 times the modulation rate). If the speech fundamental frequency range is 80 – 350 Hz, the corresponding carrier pulse rates should be at least 1400 pps if the whole range is to be represented. Higher stimulation rates may provide increased temporal detail and may provide neural firing patterns approximating acoustic stimulation (Wilson, 1997; McKay McDermott and Clark, 1994). However, other widely used speech processing strategies have different carrier pulse rates. For example, ACE (Advanced Encoded Conversion) (Skinner, Arndt, and Staller, 2002) has a high pulse rate of 900 –1800 pps whereas SPEAK (Spectral Peak Coding Strategy) (Skinner, Clark, Whitford, Seligman, Staller, Shipp, Shallop, Everingham, Menapace, Arndt, Antogenelli, Brimacombe, Pijl, Daniels, George, McDermott and Beiter, 1994) has a lower pulse rate of 250 pps. Because of the higher carrier pulse rates, cochlear implant users with ACE strategies might be expected to be provided with better pitch information (up to 300 Hz) than SPEAK users (up to 75 Hz).

1.8 The perception and production of natural tone by children with cochlear implants

1.8.1 Perception

Few studies of pitch perception have been carried out with children and most of what is currently known about the perception of pitch from speech through cochlear implants is from studies of tone languages.

In lexical tone languages such as Mandarin and Cantonese, pitch determines meaning in otherwise identical syllables. Peng, Tomblin, Cheung, Lin and Wang (2004) investigated tone identification skills for 30 CI children (aged between 6;0 and 12;6 years) and presented pairs of Mandarin tones in monosyllables and disyllables in a picture task using a live voice procedure. Overall average score was 72.88 % (chance level 50%), and scores for pairs involving the high falling tone T4 (i.e. T1 versus T4 64.7%; T2 versus T4 78.33%; T3 versus T4 76.25%) were higher than other pairs (T1 versus T2 68.96%; T1 versus T3 70%; T2 versus T3 64.79%). The authors suggest that the shorter duration of T4 may have provided a temporal cue for the implanted children to distinguish it from other tones.

Ciocca, Francis, Aisha and Wong (2002) carried out an investigation of Cantonese tones in a group of 17 prelingually deafened implanted children aged between 4;6 and 8;11 years. They were all using Nucleus 22 or 24 cochlear implants with either ACE or SPEAK speech processing strategies. Natural /ji/ stimuli representing concrete lexical items were recorded by a native Cantonese speaker and presented in a context sentence with six contrastive Hong Kong Cantonese tones (high-level, high-rising, mid-level, low-falling, low-rising, low-level). Stimuli were grouped by Ciocca et al. into eight tonal contrasts (i. HL- ML; ii. HL-LL; iii. ML-LL; iv. HR-LR; v. LR-LL; vi. LF-LR; vii. LF-LL; viii. HL-HR) in order to investigate pitch height and pitch direction. The first three contrasts were used to investigate the separation between three pitch levels (high, mid, and low) on tone perception whereas contrasts iv-vii with a similar initial F_0 were used to test listeners' sensitivity to F_0 at the end point of the second tone in each pair.

As a group, the children performed above chance for three out of the eight contrasts (HL-ML, HL-LL and HL-HR), but only a few individual children performed above chance. None of the children performed above chance for the other contrasts. Although overall performance was poor, results suggest that listeners were more accurate when pairs of stimuli differed by a large F_0 separation and one of the pair was a high tone. Average F_0 separation in the level portion of the tones was about 45 Hz for HL - LL tones, and about 35 Hz for HL - ML tones. Contrasts between ML-LL tones were not perceived above chance and were separated by an average F_0 difference of 10 Hz. Overall, correlations with age at test, post operative duration, age at implant and onset of deafness were not significant. Unlike Mandarin, tone in Cantonese is almost exclusively cued by F_0 contour and height but in high level tones amplitude can be higher for some speakers. According to the authors amplitude in high tones might have been used as a cue by the subjects in this experiment. Because of unresolved lower harmonics in implants, Cantonese implant users have to rely on periodicity cues for pitch perception, but ACE users with fairly high pulse rates (900-1000 pps) and increased periodicity information still had difficulty recognising lexical tones in this study. The authors concluded that further research was needed to establish whether auditory input or cognitive and linguistic factors contribute to lexical tone perception in Cantonese.

As discussed in section 1.4, stress in English is also cued by F_0, but duration and amplitude also play a role. Unlike Cantonese, where tone is cued almost exclusively by F_0, it is possible that duration and amplitude cues might be available to English speaking children with cochlear implants. The results of the study carried out by Ciocca et al. suggest that as a group subjects performed above chance for only three out of eight tonal contrasts where one member of a contrasting pair was a high tone. It is suggested that the reason for this was the relatively large F_0 separation (i.e. 35 Hz-45 Hz) between the high tone and other tones. Other contrasts such as ML-LL with only 10 Hz separation between the tones were not perceived above chance.

In another study of Cantonese tonal contrasts, Barry, Blamey, Martin, Lees, Tang, Ming and van Hasselt (2002a) investigated a group of 16 congenitally deaf children with implants (aged 4;2 - 11;3 years) in an adapted speech feature test (Dawson, Nott, Clark and Cowan, 1998) involving a change/no change test paradigm. The children were using Nucleus 22 and 24 speech processors with either ACE or SPEAK speech processing strategies and had received their implants between the ages of 2 and 6 years. A group of younger normal hearing children (3;9 - 6;0 years) were also included to provide a lower limit of discrimination performance by Cantonese speaking children. Barry et al. suggest that the poor results of Ciocca et al. (2002) might have been influenced by the gradual acquisition of tones and the demands of a lexical labelling task, and they decided to use non-meaningful /wi/ stimuli so that performance depended on hearing ability rather than on age or linguistic ability. Recordings of /wi/ stimuli with the six Cantonese tones were made by a trained native Cantonese speaker and comparisons of acoustic details of all the relevant tones in productions of /ji/ stimuli indicated a standard F_0 range in accordance with reported mean F_0 values for a Cantonese-speaking female (i.e. 250 Hz onset – 272 Hz offset for high level tone and 210 Hz onset – 172 Hz offset for low-falling tone). However, because of difficulty discriminating tones 3 (mid-level) and 6 (low-level) in the non-word /wi/ by both implanted and normal hearing children in the early stages of testing, a decision was taken to use /ji/ stimuli for these tones. A total of 15 tonal contrasts were presented i.e. Tones 1-6 HL, HR ML, LF, LR, LL.

Tone discrimination was significantly better for the normal hearing children although the children with cochlear implants gained sufficient information to perform reasonably well on a number of contrasts. The children using the SPEAK processing strategy obtained group average scores of greater than 0.67 (above chance) in discriminating all except four tonal contrasts whereas the poorest performers were ACE users who achieved a group average of less than 0.67 for seven contrasts (p.90-93). As for Ciocca et al. (2002) above, scores were better for contrasts when one member of a contrast was a high tone than for contrasts involving mid or low tones. A possible reason for this, according to Barry et al., is that the onset frequencies of the mid and low tones were crowded into the lower frequency range. For example, although there were different dynamic contrasts between tone 4 (low-falling with onset 198.6 Hz - offset 155.8 Hz) versus tone 5 (low-rising with onset 188.6 Hz - offset at 224.1 Hz), this contrast was particularly difficult for both ACE and SPEAK users. Barry et al. predicted the ACE users with the higher pulse rate (900-1000 pps) might have performed better but there was no significant difference between strategies. Overall the SPEAK group performed better, and the higher stimulation rate in ACE was not found to be an advantage. Although ACE users were younger than the SPEAK users, years of experience was not found to be statistically significant. Lack of advantage for ACE users could not be attributed to limited experience with the implant. The authors suggest that differences between the strategies and increased individual variation in ACE users in this study might be due to coding strategies not being optimised to individual needs (see section 1.7 above). According to Barry et al., previous studies of adults suggest that pitch height would appear to be of primary perceptual importance to Cantonese speakers generally, whereas subtle pitch direction changes might not be easily perceived. Implanted children in their study had difficulty discriminating contrasts involving mid and low tones with onset frequencies crowded into the lower frequency range. Results support Ciocca et al. (2002) above who also found pitch height to be more perceptually salient than pitch contours.

The variation across normal hearing and implanted children investigated in Barry et al. (2002a) and the possibility of gradual development of tonal perception led to further analysis by Barry, Blamey and Martin (2002b). A multidimensional scaling (MDS) analysis of 9 normal hearing children (aged between 3;9-6;0 years) and 14 implanted children (aged between 7;2-11;3 years) was carried out. The results of the study show that despite differences in linguistic experience and auditory input, all listeners used two dimensions i.e. pitch height (level) and pitch direction (contour) in their perception of tone contrasts. The results confirm previous studies of normally hearing adult listeners using the same technique. The findings of Barry et al. (2002b) suggest that SPEAK users rely more heavily on information about pitch height for making judgements about tone contrast than ACE users. Although there is considerable variability in performance in ACE users, the higher stimulation rates seem to provide more information about pitch direction than pitch height. The authors conclude that further investigations will focus on normal hearing children to establish the effects of linguistic experience and the gradual development of tone discrimination.

More recently in a study of the perception of voice similarity, Cleary, Pisoni and Kirk (2005) investigated how different F_0 and formant frequencies needed to be in English sentences before two different talkers were perceived by normal hearing and children with cochlear implants aged between 5;0 and 12;0 years. Sentences which were originally produced by a female talker (average F_0 175 Hz) were resynthesised and mean F_0 for the tokens at the low end of the continuum averaged at 123.7 Hz corresponding to a difference of six semitones (p.208 – 209). They were presented in half semitone increments in ‘fixed’ or ‘varied’ conditions (i.e. the linguistic content either remained the same or varied). Results show that a group of 30 normal hearing subjects heard two different talkers when F_0 differences were greater than 19.5 Hz (i.e. 2 - 2.5 semitones) with proportionate shifts in formant frequencies. As predicted there was huge variability for individuals across a group of 18 implanted subjects (using SPEAK, ACE or CIS strategies) but performance was significantly greater than chance at 30.5 Hz (i.e. 3.5 semitones) in one condition where the linguistic content varied and no different from chance in all other conditions. Contrary to the authors’ expectations there was a subgroup of 8 implanted subjects who were able to hear two different talkers at F_0 differences which were audible to the normal hearing subjects. According to Cleary et al., some factors which affect speaker recognition such as speaker location, perceived loudness, and speaking rate were controlled in this experiment (p.206, citing Nolan, 1997). However, the authors also suggest that there may be other influencing factors besides insufficient spectral information which may account for variability in implanted children such as neural survival and placement of electrodes.

1.8.2 Production

Peng et al. (2004) carried out a study of the production of Mandarin tone in a group of thirty prelingually-deafened children (aged between 6 and 12 years) in Taiwan. Age at implant ranged from 2;3 to 10;3 years and duration of implant use ranged from 1;7 - 6;5 years, and 19 children used Nucleus (SPEAK) and 11 used MEDEL COMBI 40 (CIS). Four target tones (Tones1-4) in monosyllables and disyllables were elicited spontaneously in most cases and degree of accuracy was rated by a panel of native speakers. Average score for the children’s tone production was 53%. However for individual tones scores were better for T1 (62% level) and T4 (62% high falling) than for T2 (42% mid high-rising) or for T3 (46% low-dipping). The authors conclude that although the acquisition of the Mandarin tone system is delayed for the CI children in their study, results are consistent with reports on the order of tone acquisition in normal hearing (NH) children where level and falling tones (T1and T4) are acquired before contour or rising tones (T2 and T3). English-speaking hearing aid users discussed in section (1.4.5) also produce falling earlier than rising contours.

Mandarin tone production was also investigated by Xu, Li, Hao, Chen, Xue and Han (2004) in seven NH and four prelingually deafened Chinese-speaking children (aged 4;0 – 8;75 years) and using NUCLEUS implants with 2 ACE and 2 SPEAK processing strategies. Acoustic analysis of imitated samples of the four target tones and elicited samples of the subjects counting from 1-10 in Mandarin Chinese showed great individual variation among the CI children. T4 (falling) seemed to be easiest for CI children to produce. Individual errors in tone production included inability to produce rising tones and prolonged duration of T3 due to added effort. The use of glottal stops by one subject instead of low or dipping contours was considered normal (p. 365). The NH group received perfect scores (10) in the subjective intelligibility test whereas the mean scores ranged from 0.25 – 8.5 for the CI group. Differences in intelligibility scores between NH and CI children and differences in scores among CI children were found to be statistically significant. The authors conclude that inadequate pitch information delivered through cochlear implants may hinder tone development in CI children, and other variables such as age at onset of deafness, hearing aid usage, duration of deafness, age at implantation, and speech processing strategy should also be considered (p. 124).

A different approach was taken by Barry and Blamey (2004) in a study of Cantonese tones produced by 16 prelingually deafened children (4;2 – 11;3) using NUCLEUS 22 (6 subjects) and NUCLEUS 24 (10 subjects) implants with either SPEAK or ACE speech processing strategies. Also included were 5 NH adults (23 – 40 years) and 8 NH children (3;8 – 6;0 years). Spontaneous productions of six Cantonese tonemes in words frequently used by children over the age of 3;0 were elicited in a different syllables using picture prompts, and acoustic measurements of F_0 onsets (x axis) and offsets (y axis) were plotted and grouped according to tone types in six ellipses for each speaker. The ellipses were calculated by determining the distribution of points around a mean to provide a visual summary of the location of six tonemes. It was expected that rising tones would cluster close to the y axis and falling tones close to the x axis and level tones would fall midway. The number of correct tones produced by a speaker is reflected in degree of differentiation between the ellipses (p. 1741), and the approach has been found to be appropriate for Cantonese where pitch level is suggested to be more perceptually salient than pitch contour (p. 1746).

Results show significant differences in median tone areas for the three groups of speakers for all tones, with larger ellipse areas for the CI and NH children than for the adult group. Intertonal median differences for the CI group (10.1 Hz-32 Hz) were smaller than for the NH adults (85.5 Hz and 16.6 Hz) and NH children (147.2 Hz – 16.9Hz) and the differences between the three groups were significant. The authors conclude that larger tonal ellipse areas for the NH children suggested more differentiation and greater spread of pitch usage for each tone type than for the CI children (p. 1746), and this is reflected in the auditory transcription where average percentage correct tones for the NH children was 78%. The authors also suggest that smaller tonal ellipses might have been expected given that NH children are reported by some studies to have acquired a tone production system by aged two but the variation found in the results may be due to the fact that a tonal system is still developing in 3-6 year olds. Measurements of the relationship between tonal space and ellipse area show very little differentiation in the production of tone by the CI children and this is born out in the auditory transcription of the data where the average percentage correct tones was below chance at 38%.

1.8.3 The relationship between perception and production

Although a statistically significant correlation was found by Peng et al. (2004) between average overall scores for tone production and identification in a group of 6;0 to 12;0 year old CI children, the correlation was not found to be significant when three high scoring children were removed. No significant correlations were found between tone production and identification and device types. Significant correlations were found between tone production scores and age at implant, and between overall tone identification and duration of implant use for NUCLEUS users only. However, results show that a group of MEDEL users, despite more limited range of experience (18-30 months), performed just as well as NUCLEUS users (31-77 months), and the authors suggest that the faster acquisition rate might be due to a higher stimulation rate (CIS). Peng et al. also suggest that the performance of some very high scoring children must be accounted for by variables other than device type. The children who performed well in tone production in this study also performed well in tone identification but the reverse was not always the case. The authors conclude that tone production and tone identification may not develop in parallel and may be associated with age at implant and duration of implant use.

Barry and Blamey (2004) report that contrary to previous studies of tone production in young Cantonese normal hearing children their findings suggest that the 3-6 year olds have not yet fully acquired a tonal system. Although previous studies of profoundly hearing impaired children report that tone production skills were better than perception skills, Barry and Blamey found that their CI children produced some F_0 contours that could be labelled as correct in the auditory transcription, but these were not produced consistently enough to be considered acquired. The authors suggest that the results support previous studies of tone perception which show that young children are still developing skills for normalisation of pitch level differences between tone. They conclude that longitudinal studies using their methodology would be appropriate for monitoring tone development in individual children.

1.9 Experiments with adult cochlear implant users

Experiments involving a variety of current speech processing strategies with adult cochlear implant users carried out by Richardson, Busby, Blamey and Clarke (1998), Guerts and Wouters (2001) and Green et al. (2004) indicate pitch perception ability of adult CI users.

Richardson, Busby, Blamey and Clark (1998) carried out two experiments in a study of six post-lingually deafened adults using Nucleus 22 cochlear implants. The subjects were all using the MPEAK speech processing strategy where acoustic F_0 is coded is pulse *rate* and acoustic amplitude is coded as pulse *duration* (p. 231).

The first psychophysical experiment investigated the discrimination of pairs of steady state and time-varying stimuli of different pulse rates i.e. F_0 (100 pps, 200 pps, 400 pps) over a series of stimulus durations i.e amplitude (100 ms, 250 ms, 500 ms, 1000 ms) using an adaptive procedure converging around the 50% point. The results of the pulse rate study show that for steady - state stimuli difference limens (i.e. F_0 thresholds) for 100 pps and 400 pps were 6% and 17 % respectively, whereas for the time-varying pulse rates, F_0 thresholds were larger (26% or 32 % at 400 pps) for some subjects or similar (8 - 11% at 100 pps) for others. The authors also noted a large range of performance between subjects.

In the second experiment, performance was measured for five prosodic contrasts with MPEAK strategy and three other strategies which removed pulse rate or pulse duration information. The prosodic contrasts tested involved roving stress (SPAC-1), rise-fall (SPAC-2), and pitch and intonation (SPAC-3), and accent and question and statement (MAC-1 and MAC-2). In general scores were better for the MPEAK strategy than other strategies and a significant difference was found between strategies except in one subtest (SPAC-3) which involved discriminating between gender and intonation. There was a significant difference between strategies for most tests and the results suggest that elimination of pulse duration or pulse rate information results in poor prosody perception performance. However, it was also found that mean performance for the three SPAC tests (91%, 88%, 66% respectively) with the MPEAK strategy in this study was better than earlier versions of Cochlear speech processing strategies (i.e. F_0-F2 and F_0-F1-F2 combined) reported in other studies for the same SPAC tests (74%, 69%, 55% respectively). Richardson et al. also state that for the two MAC tests, mean scores with the MPEAK strategy were 83% and 86% compared with 64% and 87% reported previously for an earlier F_0-F2 strategy. However, the authors conclude that because of the small number of subjects, results should be interpreted with caution. They also suggest that performance with modified strategies might improve with training and experience.

Guerts and Wouters (2001) investigated how different modulation depths (i.e. the difference between maximum and minimum pulse amplitude) might affect the discrimination of modulation rate as a temporal cue to pitch in four post-lingually deafened adults using the LAURA cochlear implant with a CIS processing strategy with a carrier pulse rate of 1250 pps to each electrode.

In the first experiment subjects had to indicate which of two sinusoidally amplitude modulated pulse trains (SAM) had the higher pitch. Modulation frequencies in each pair were either 150 Hz and 180 Hz or 250 Hz and 300 Hz and they were presented at different modulation depths to a single channel. Results varied according to subject, channel, frequency range of the stimuli and modulation depth (20% - 99%) with some who met the criterion of 75% correct and others who did not for any modulation depth. The authors suggest that poor performance in the higher range (250 Hz) may be because relative change in modulation depth (20%) may be below the detection limit for this frequency range.

In the second experiment the smallest discriminable difference was measured between pairs of synthesised /a/ or /i/ vowels with F_0 ranging between 370 Hz and 149 Hz. The standard stimulus (F_0 at either 149 Hz or 250 Hz) and the comparison which varied in F_0 were presented to all available channels in three different speech processing algorithms based on CIS. Good results were obtained for all four subjects for /i/ at an F_0 of 250 Hz only with an envelope cut-off frequency of 50 Hz removing all temporal cues (FLAT CIS). Although the subjects may have been helped by average relative amplitude in each channel for the high frequencies, the authors suggest that amplitude would be unlikely to provide a reliable cue in natural speech as there are other sources of information such as formant frequencies and variation in size of vocal tract for male and female speakers.

In the other two algorithms (i.e. CIS with an envelope cut-off frequency of 400 Hz and fluctuations present, and F_0 CIS with increased modulation depths) all subjects perceived lower F_0 differences ranging from 6-20 Hz when the standard stimulus was at 150 Hz. For two individuals who were sensitive to differences above 250 Hz for /a/, F_0 differences perceived ranged from 12 Hz to 19 Hz. There was no significant difference between the second and third algorithms. The results of these experiments suggest that adult implant users are obtaining some pitch information but the minimum F_0 difference thresholds between the stimuli vary according to subject, processing strategy (algorithm), and F_0 range. The results show that in the absence of temporal information in one algorithm, listeners used average amplitude as a cue to F_0 difference. In the other algorithms which included temporal fluctuations, some individuals only perceived large F_0 differences between vowels.

Green, Faulkner and Rosen (2004) carried out another experiment with eight post-lingually deafened adults using Clarion cochlear implants with CIS and two modified strategies based on CIS. Synthesised diphthong stimuli with dynamically changing spectral structures were presented in a glide labelling task to assess the impact of variations in formant structure on cues to voice pitch. The diphthongs (/au/ /eɪ/ /oɪ/ /aɪ/) had start-to-end frequency ratios which varied in logarithmic steps, in two F_0 ranges, with centre F_0 (mean of start and end F_0) of each glide at 113 Hz and 226 Hz. For each dipththong, start-to-end ratio, and F_0 range there was one ascending and one descending glide and listeners had to identify a glide as rising or falling in pitch. In the standard processing condition, CIS, mean performance for the 113 Hz range, although above chance, was very limited. Pitch direction was only correctly identified in 70 % of trials with an octave change in F_0 over the course of the glide and performance was poorer for smaller glides. It is suggested that temporal pitch cues were less effective in the presence of dynamic slow-rate spectral variation caused by the changing formant structure of the diphthongs (p. 2309).

In the studies discussed above F_0 thresholds varied according to subject, speech processing strategy and F_0 range. The stimuli presented also varied and became increasingly complex and more speech-like ranging from pulse trains to synthesised vowels and diphthongs, and in one early study (Richardson et al. 1998) prosodic contrasts in natural speech such as stress and intonation were presented. Although overall results indicate limited abilities in the experiments discussed above, adults do gain some pitch information from their implants, and this improves slightly with modified speech-processing strategies.

1.10 Cochlear implant simulations with normal hearing adults

The use of vocoders in simulation studies with normal hearing listeners has useful applications in the improvement of cochlear implants as they mimic the limited spectral resolution and unresolved lower harmonics of speech processing strategies. Simulation studies with normal hearing adults such as those discussed below (Green, Faulkner and Rosen, 2002, 2004; Laneau, Moonen and Wouters, 2006) involve the manipulation of spectral and temporal information in the stimuli (i.e. tone glides and synthesized diphthongs or synthesised vowels). The results have implications for young children with cochlear implants at the early stages of prosodic development using standard speech processing strategies.

In the study by Green, Faulkner and Rosen (2002), seven normally hearing listeners were presented with synthesised complex tone glides in three F_0 ranges, with ratios of start to end frequencies varied in six logarithmic steps. The midpoint for each start to end F_0 (centre frequency) in the three F_0 ranges was 146, 208, and 292 Hz. For each ratio and F_0 range, subjects had to identify each glide as falling or rising. They were presented in two four-band and two single-band conditions, with and without spectral information respectively. Cut-off frequencies were at 400 Hz and 32 Hz with temporal F_0 related fluctuations removed from the latter (see discussion in section 1.7).

The results show that in the absence of temporal and spectral cues in the *Single32* condition listeners could not discriminate between falling and rising glides in any of the F_0 ranges, and performance was below 50%. However, in all the other conditions with either limited spectral or temporal information (i.e. *Single400, Four32, Four400*) performance was at or near ceiling for the lower 146 Hz range, but only for the largest start to end F_0 ratios. Performance was also near ceiling for the 208 Hz range in the *Four32* condition only, and as no temporal information was available performance could only be due to spectral information at this centre frequency. The results of the experiment indicate listeners derive some limited pitch information particularly in the lower 146 Hz range but only for large F_0 start-to-end ratios in three of the simulation conditions. These results have implications for the prosodic development of cochlear implant users as F_0 ranges for females and children extend beyond this range and very limited temporal cues to pitch are available through standard processing conditions.

In a second experiment, synthesised diphthongs with time varying formants were presented to six of the adult hearing listeners referred to above. The same F_0 ranges, start-to-end frequency ratios and centre F_0 values, and processing conditions were used except for *Single32*. The stimuli used in the two experiments above produced different results. For example, performance with diphthongs was near ceiling for the lower 146 Hz range in three processing conditions with glides in the first experiment, but in only one (*Four400*) of the three processing conditions used in the second experiment. When temporal F_0 related fluctuations were removed in the *Four32* condition in the first experiment, subjects had good glide labelling performance, but chance performance at 50% in the second experiment indicated that spectral cues were obscured by the spectral dynamics of the diphthongs. The authors conclude that increased numbers of channels or natural rather than synthesised speech stimuli (p. 2163) may provide listeners with additional cues.

More recently, similar results for synthesised diphthongs and an increased number of channels were obtained by Green, Faulkner and Rosen (2004) when spectral cues were available in a speech processing condition simulating the standard CIS (continuous interleaved sampling). In this condition, listeners were unable to discriminate pitch change even for an octave change in F_0 over the course of the glide. However, in other conditions with improved temporal information (*sine and sawsharp*) performance was 90% in the low 141 Hz range for an octave change in F_0. As for Green et al. (2002) performance for these two conditions declined across the F_0 ranges (141 Hz, 199 Hz, and 282 Hz) but was still above chance. Comparisons between the simulations and the experiments with implanted adults are informative and show that the best implant users achieved scores within the range obtained by normal hearing subjects in the simulations (Green et al., 2004, p. 2306).

Effects of different filters and vocoders on temporal and spectral cues

Factors affecting the use of noise-band vocoders as acoustic models for pitch perception in cochlear implants were investigated by Laneau, Moonen and Wouters (2006). The first two experiments concern the effects of spectral smearing on simulated electrode discrimination and F_0 discrimination by NH subjects using a CI simulation (CISIM vocoder) and by CI subjects which were reported in a previous study (Laneau, 2004). Place pitch just noticeable differences (jnd) between a reference and comparison frequency (in the first experiment) and stylized vowel stimuli with temporal cues removed (in the second experiment) were matched for the two groups when the width of the excitation pattern (i.e space constant) was increased to 1 mm. Results of the second experiment show that the NH CISIM group had better place pitch discrimination with smaller space constants than the CI group.

In a third experiment the same synthesised vowels were presented in two conditions (a. with place pitch cues only and b. with temporal and place pitch cues) and results show that different vocoders and filters have important effects on temporal and spectral cues. For example, when only place pitch cues were present there was no significant difference between the performance of the NH subjects using a CISIM vocoder and the CI subjects. When temporal cues were added there was a smaller improvement for the NH CISIM group than for the CI group.

The authors point out that the CI subjects were post-lingually deafened adults and children implanted earlier during the critical period may perform better than later implanted children (p. 504). However, results must be interpreted with caution because vocoder simulation generally does not represent an exact match for the information provided by a cochlear implant. In Experiment I in the present study an acoustic simulation of a cochlear implant is presented to a group of normal hearing children within the same age range as the implanted children for comparison. The purpose of this experiment is to establish whether performance is similar or different for both groups. If performance is similar it is possible that difficulties could be related to device or speech processing strategy whereas if the normal hearing children are better in the simulation condition there could be other factors affecting implanted children such as placement of the electrodes in the cochlea (see section 1.11.6).

1.11 Relevance of the literature to the present investigation

1.11.1 Higher order acquisition issues

Early Acquisition of intonation and stress contrasts in English

The role of pitch in helping infants acquire the rhythmic properties of a stress language such as English and its importance in the development of a lexicon and language generally has been discussed in section 1.3 and 1.4. In English pitch carries important information about stress and intonation for pragmatic, emotional and syntactic purposes, and also for gender identity. As stated in section 1.3, reports show that hearing babies begin canonical babbling (i.e. strings of alternating consonants and vowels) between 6 -10 months while it is delayed in deaf babies to between 11-25 months indicating that babbling does not develop normally in the absence of auditory input (McNeilage, 1997; Clement et al., 1996; Oller and Eilers, 1988). The importance of ambient environment and its influence on babbling and prosodic production in normal hearing infants as young as 8 months has also been documented by Juscyzk (1997). Prosodic adjustments by adults in speech directed at very young children (Baby Talk i.e. BabyPr) such as frequent use of higher pitch, rising intonation for encouragement, slower articulation, whispered speech and longer pauses may facilitate language acquisition (Cruttenden, 1994). However, these adjustments may not be accessible to deaf babies with limited residual hearing and prosodic development may be delayed.

Without available normative data to draw on for very young hearing children it could be expected that implanted children might develop prosodic abilities and particularly intonation more slowly and possibly differently than hearing children as a result of auditory deficits. In addition, device limitations in cochlear implants (see section 1.7) may mean that pitch cues are not accessible to implanted children even when exaggerated so they have to rely on duration and amplitude cues. However, as outlined in Chapter One (see hypotheses in section 1.1.2) it has yet to be established whether the perception and production of intonation is directly linked to implanted children's ability to hear pitch cues (i.e. F_0). The hypotheses are as follows: (i) If F_0 is a necessary cue, intonation contrasts will not be accessible to implanted children and they will not be able to hear F_0 patterns associated with pragmatic contrasts such as given vs. new or focussed words, or grammatical contrasts such as compound vs. noun phrase. If they have no stored representation or prior knowledge of how intonation conveys these contrasts, they will not learn to produce them meaningfully in the same way as hearing children. (ii) If on the other hand F_0 is not a necessary cue to intonation, implanted children will be at less of a disadvantage during the early stages of prosodic development. Eye contact, gestures, actions, jumping up and down, reaching (Crystal, 1986; Snow and Balog, 2002) may draw attention to certain features such as rhythm, response required or not required during interaction with an adult and help develop some prosodic awareness in combination with loudness or duration cues even if pitch cues are not accessible. It may be the case that implanted children perceive stress, intonation and other prosodic contrasts using whatever cues are available to them. In this way they might be able to develop an abstract prosodic and linguistic system which is independent of their ability to hear a particular cue. The intonational contrasts which are of particular interest in the current study of school-going children are *compound vs. phrase stress* and *focus (tonicity)* and they are discussed in more detail below.

Compound vs. phrase stress

As discussed in section 1.3.2 there seems to be a consensus in previous studies of school aged hearing children in the US, Britain and Southern Ireland (Atkinson-King, 1973; Vogel and Raimy, 2002; Wells et al., 2004; Doherty et al., 1999) which suggest that the ability to discriminate between compound vs. phrase stress (e.g. BLUEbell vs. blue BELL) does not seem to be developed until late in the acquisition process. Some of these studies suggest it can continue to develop up to and beyond 12;0 years. Vogel and Raimy found a preference for compounds for known items regardless of stress patterns between 4;4 and 7;7 years and that by 7;0 years children were becoming sensitive to patterns they were familiar with, but compound and phrase patterns were not generalized to novel items. Wells et al. (2004) found that the ability to discriminate between compound (coffee-cake) and two nouns (coffee cake) in a group of children in Southern England showed improvements between 5;0 and 10;0 years. In the present study the issue for consideration is whether implanted and normal hearing children can hear differences in lexical stress by 6;10 years. Although there is only a small number of implanted and normal hearing subjects in the current study the age range extends up to 17;11 years and should provide some insight into the pattern of development that might be expected for both groups of children beyond 13;0 years. This will provide a baseline for future research with other normal hearing and implanted subjects within this age range for Southern Hiberno English and different varieties of English. These contrasts have not been investigated for children with cochlear implants and as discussed above it has yet to be established whether they can ever be acquired in the absence of pitch cues or whether they can draw on other cues to develop an abstract linguistic system with representation of these contrasts. The acoustic cues to compound vs. phrase stress are discussed in section 1.11.2 below.

Focus (Tonicity)

Of particular interest in the general acquisition literature for normal hearing children is nuclear or tonic placement (also referred to as *tonicity* by some authors) which concerns the placement of maximum prominence on a particular syllable for grammatical or pragmatic purposes (Crystal 1969, 1987; Wells and Local, 1993). Evidence from previous studies of normal hearing children (Snow and Balog, 2002) indicates that intentional pragmatic and grammatical intonational functions develop after 10 months whereas before that intonation is associated with physiological and emotional needs. According to Crystal (1986), young children at the two word stage (i.e. 1;6 years) can produce variations in tonicity to distinguish old from new information. Cutler and Swinney (1987), however, report that processing of focus words in their study was significant for a group of 5 year-old subjects but not for a preschool group when focus was determined by questions preceding the sentences which were presented to them. Cruttenden (1997), on the other hand, states that at the two - word stage children can vary nucleus placement, and by the time they produce three or four word utterances they can vary nuclear placement to indicate old information. However, he also reports that some aspects of intonation develop early but some children as old as 10;0 years have difficulty with intonational meaning. Wells et al. found that some aspects of intonation e.g. chunking, affect and focus were established in 5 year-olds whereas other aspects of intonation which were more difficult for younger children were acquired by most 8 year-olds. Most relevant to the current study of focus production is a preference for utterance final focus and Wells et al. suggest that maintaining or ending the end of a conversational turn might compete with focus and accent placement as a result of delayed or immature prosody. Individual variation was also reported by Wells et al. across the age range (5;0 to 13;0 years) but they concluded that children's ability to interpret focus or accent in other speakers lagged behind the ability to realise focus in their own speech. Ambiguity is also found across the age range for contrastive (i.e. narrow) focus which they state is not uncommon amongst adult speakers of English.

The normal hearing subjects in the current study are aged between 6;10 - 17;10 years and the implanted subjects are aged 5;0 – 17;1 years. Although some studies cited above would suggest that normal hearing children aged 6;10 years should be able to process focus words, others report that variation, ambiguity and difficulty with intonational meaning may occur across the age range. The 5 year-old children with cochlear implants might also have difficulty processing focus words, but this could also be compounded by early auditory deprivation and device limitations of the cochlear implant discussed in section 1.7. As we have no available data on implanted children to draw on it needs to be established whether in the absence of pitch (F_0) information they can develop prosodic abilities and particularly intonation more slowly or differently than hearing children.

It also remains to be seen whether implanted children can acquire an abstract representation of focus and tonicity using whatever cues that might be available to them through the implant. As set out earlier, if F_0 is a necessary cue to the perception of stress and intonation, children with implants may not acquire abstract concepts of intonation contrasts or learn to use F_0 to convey or interpret meaningful intonational contrasts. On the other hand if F_0 is not a necessary cue to stress and intonation, a preference for utterance final focus up to or beyond 8;0 years. Difficulty interpreting intonational contrasts produced by others might be due to delayed prosody development or early auditory deprivation rather than pitch limitations of the implant. In the absence of pitch (F_0) information children with implants may be able to rely on duration and/or amplitude cues. In the following section, acoustic cues to compound vs. phrase stress and focus (tonicity) are discussed.

1.11.2 Lower order issues

Development Issues

McNeilage outlines the stages of vocal development reported in the literature on normal hearing infants (section 1.3.1.2) and infants as early as 2-4 months use vocal play with regular syllable timing, manipulation of pitch (squeals and growl) and loudness (yells and whisper). Studies have also shown the effects of ambient language on normal hearing infant prosodic patterns from 8 months (McNeilage, 1997; Snow and Balog, 2002) for example, and more rising intonation is used by French infants than English infants. However, it is suggested that simple rises in French might be easier to produce than complex rises (i.e. rise-fall or fall-rise) typical in English. A study of normal hearing and deaf infants (Clement et al., 1996) suggests that that there are no clear differences in mean fundamental frequencies between 5 and 10 months. The reason given for this is that the development of fundamental frequency at this stage is determined by anatomical and physiological growth rather than hearing status and this accounts for a predominance of falling intonation in the first 3 – 9 months of life (Snow and Balog, 2002). Snow (2001) also reports in another study that normal hearing 4 year-old English speaking subjects had slower rate of pitch change, narrower accent range than adults and lengthened word durations in rising tones. Wells et al. also found that some younger children had difficulty with complex intonation patterns e.g. fall-rise (not keen) and rise-fall (keen) or rising intonation for clarification, and a bias toward utterance final focus placement but these patterns were mastered by 8:0 years.

It remains to be seen in the present study whether children with cochlear implants can interpret or convey focus in the absence of pitch information and if so whether they use the same or different acoustic cues as the hearing subjects As discussed earlier it is not clear in the literature whether F_0 is a necessary cue to stress and intonation, but if implanted children can acquire an abstract concept of focus by relying on other acoustic cues it is possible they may be able produce appropriate F_0 patterns. However, like normal hearing children they might also continue to have a slower rate of pitch change in addition to a narrower accent range, and there may be difficulties with rising intonation for developmental reasons. The acoustic cues to compound vs. phrase stress and focus (tonicity) are discussed below and some of the issues raised above will be considered in detail in Experiment III (Chapter Four) in the analysis of the production of focus on target words by the implanted children in the current study.

Acoustic cues to compound vs. phrase stress

As discussed earlier in section 1.4, early experiments with normal hearing subjects showed that F_0, duration and intensity contributed to the perception of stress and F_0 provided the most important cue in words with first or second syllable stress such as SUBject or subJECT (Fry, 1955, 1958; Lehiste, 1970; Gay, 1978a, 1978b). Ladd (1996), however, suggests that if such words occur after the main intonation peak in a sentence or if question intonation is imposed on the sentence, stress differences can still be heard but are not cued by a pitch peak. Despite the view expressed by Ladd, there is still a widely held view in the literature that lexical stress is signalled by primary stress/accent on the first element in a compound word such as *BLUEbell* and on the second element in a noun phrase such as *blue BELL.* Acccording to Cruttenden (1997) primary stress/accent refers to the main pitch prominence in an utterance. However, a more recent study of prosodic variation in adult speakers of Southern British English by Peppé et al. (2000) shows that differences between compounds and phrases may not be signalled in the same way by different speakers and that pitch movement and pitch reset may not be as reliable at signalling differences between compounds and phrases as lengthening and pause. The traditional view that pitch is a necessary cue to compound vs. phrase stress may be based on laboratory experiments whereas it could be that case that in more natural speech pitch cues are not necessary to cue these contrasts. The possible implications of this view is that listeners with cochlear implants may be able to hear differences between compound vs. phrase stress using duration rather than pitch cues.

In Chapter Two in the current study (see overview of the experiments in sections 1.1 and 1.11.8) pairs of non-meaningful synthesised (e.g. baBA vs. BAba) stimuli are presented with controlled changes in F_0, duration and amplitude signalling first or second syllable stress (Experiment I). The results should inform us how accessible these cues are (and particularly F_0) in signalling lexical stress to both implanted children and normal hearing children in a cochlear implant simulation (section 1.11.5). However, in Experiment II in Chapter Three, natural speech stimuli are presented to the same subjects, but the acoustic cues are not controlled so speakers may vary in their use of F_0, duration and amplitude, and listeners might be able to rely on combinations of these cues to hear differences compound or phrase stress. If, as suggested above, F_0 is *not* a necessary cue to compound vs. phrase stress, poor F_0 discrimination between synthesised /baba/ syllables by implanted listeners in Experiment I may not necessarily mean poor performance in the linguistic task in Experiment II because other timing and amplitude cues should be more accessible to them. On the other hand if F_0 *is* a necessary cue to compound vs. phrase stress then subjects will have difficulty hearing F_0 differences in Experiment I which will lead to difficulty discriminating between compound vs. phrase stress in Experiment II. In addition to pitch limitations of the implants there are also the acquisition issues to be considered which could account for individual differences and difficulties in discriminating between compound vs. phrase stress across the age range.

Acoustic cues to focus (or tonicity)

There seems to be consensus in the literature that narrow focus on a target word is conveyed to a listener by an increase in F_0 peak, followed by a high F_0 fall as well as increases in duration and intensity. Different focus types and oppositions were discussed in section 1.2, and there is a general view that English speakers can make a distinction between new or contrastive information, or broad or narrow focus, or express different focus types by deaccenting or boosting stressed syllables in an utterance (Ladd, 1996; Gussenhoven, 2006). Studies of adult hearing speakers show that this can be achieved by different means such as a change in pitch configuration (contour or direction) or in pitch height, or expansion and compression of F_0 in focus and post-focus words (Xu and Xu, 2005), and by durational and amplitude adjustments. Peppé et al. also report individual variation in how narrow focus is signalled. They report that although a falling glide occurred for most individuals there were differences in how other phonetic exponents were used e.g. silence, lengthening, loudness and pitch reset. However, the authors also suggest that there may be differences in the phonetic realisation of intonational contrasts in less controlled situations compared to laboratory conditions.

This view is supported by the results of a quantitative study (Kochanski et al., 2005) of accented syllables in natural speech in school going subjects (mean age 16;0 years) using different varieties of British English (including Belfast and Dublin). Although Kochanski et al. reported that accented syllables perceived as prominent by listeners were marked by loudness and duration cues and that F_0 played a minor role, these results are not conclusive as specific contrasts were not analysed and results might differ if contrasts such as focus or compound and phrase stress were elicited. The results suggest that F_0 may *not* be a necessary cue to stress and intonation in English (hypothesis (ii) section 1.1.2). If this is the case the absence of F_0 or pitch cues may not be such a disadvantage to cochlear implant users as they may be able to convey and interpret intonational contrasts such as focus using duration and amplitude cues. As stated earlier there may be physiological reasons for appropriate increases in F_0 in the production of focus words by implanted children simply because of tension associated with interest in the target word. Increased interest in a word may lead to an increase in F_0 which is also linked with an increase in amplitude.

So it is possible that durational cues and also F_0 and amplitude might be used appropriately on target focus words by CI children even if they cannot hear pitch differences in the natural speech stimuli in Experiment II or in the controlled /baba/ stimuli in Experiment I. However, if F_0 is a necessary cue to focus (see hypothesis (i) in section 1.1.2) then F_0 changes may be insufficient to be heard by implanted children in the focus stimuli in Experiment II. In the production of focus in

Experiment III implanted talkers might produce F_0 contours which are appropriate for physiological reasons stated earlier but insufficient boosting or deaccenting F_0 might lead to ambiguity or failure to convey focus to a listener. As discussed above for compound vs. phrase stress there may also be developmental issues affecting implanted subjects' ability to interpret or produce focus. The relationship between perception and production of stress and intonation is not straightforward and is discussed again in section 1.11.4 below.

Production of intonation by children using hearing aids

As outlined above for normal hearing children, the development of falling intonation before rising intonation is also reported for English-speaking children with hearing aids aged between 7:0 and 8:0 years (Abberton et al., 1991) and in another study (Most and Frank, 1994) hearing impaired children between 5:0 and 12:0 years were found to be less successful at producing rising than falling intonation. In another study (O'Halpin, 1993; 2001) two 8;0 year old hearing aid users did not convey contrastive stress before training but after training one subject used exaggerated but appropriate F_0 contours (including rise-fall patterns) and increases in duration and intensity similar to a hearing subject of the same age. However, previous studies of the speech of children using hearing aids (Rubin Spitz and McGarr, 1990; Murphy, McGarr and Bell-Berti, 1990; Most, 1999) also report that correctly perceived stress and intonation patterns may not be conveyed by the same acoustic correlates or there may be conflicting cues e.g. duration or amplitude which may affect listeners' perception of F_0. These results would also support hypothesis (ii) in section 1.1.2 that F_0 is not a necessary cue to stress and intonation.

Production of intonation by children using cochlear implants

It remains to be seen whether CI children can make use of appropriate F_0 contours to convey differences in stress and intonation in English. As discussed earlier if F_0 <u>is</u> a necessary cue to stress and intonation, the F_0 changes associated with the grammatical use of intonation in their linguistic environment may not be accessible to these children and they may not learn to use F_0 appropriately. On the other hand if F_0 is <u>not</u> a necessary cue then implanted children can rely on other cues such as duration and amplitude to help develop an abstract prosodic system such as focus and may produce appropriate F_0 without necessarily hearing it. As stated above the relationship between perception and production of stress and intonation is complex and will be discussed again below in section 1.11.4. It may be the case that different cues might be used in perception and production or that some children produce appropriate F_0 contours because of the physiological tension associated with a focus word. In the present study the appropriate use of F_0, duration and amplitude is investigated in sentences with target focus words produced by CI talkers and a small group of NH talkers in Experiment III. Although the methodology differs from the various studies mentioned above, changes in F_0 (and duration and amplitude) on the target focus words and the ability to convey focus to a listener will be considered.

The developmental studies discussed earlier mostly involved American and British subjects so the current investigation will provide additional new data from an Irish population. A few experimental studies of intonation in Dublin English (Dalton and Ní Chasaide, 2005; Grabe and Post, 2002) suggest that falling tones are associated with declarative sentences which is similar to Southern British English whereas rising tones are more typical in Belfast English. One preliminary study of adult speakers of Dublin English, however, suggests that focus or contrast might not always be conveyed to a listener in initial or final position (O'Halpin, 1994), despite appropriate increases in F_0, duration and intensity. According to Wells et al. focus in final position may compete with end of a conversational turn, and they also report that ambiguity in narrow focus is not uncommon in children and adults.

1.11.3 Acoustic cues to lexical stress in tone languages: what can we predict for English speaking implanted children from the results of experimental studies of pitch perception and production of Chinese tones?

In tone languages such as Cantonese, pitch plays an important role in determining lexical meaning and intelligibility in otherwise identical syllables and is a necessary cue to tone discrimination. Most of what is currently known to date about the perception of pitch in speech through cochlear implants is from tone languages but there may be a closer link between perception and production than for English where listeners can also rely on temporal and amplitude cues. Although Ciocca et al. report that overall performance was poor in their study, they found that children performed best in three out of eight contrasts where the average separation of tones was either 35 Hz or 45 Hz and also when one of a pair of tones was a high tone. In other words the implanted children needed almost half an octave difference between pairs of tones before they could identify them. Barry et al. suggest that poor discrimination of contrasts involving low to mid tones regardless of direction might be due to onset frequencies being crowded into lower frequency range, and these onset differences may not be perceptible to cochlear implants users in the absence of other cues. It would appear that F_0 *is* a necessary cue to tone discrimination particularly in Cantonese and has important implications for the acquisition of tones by young implanted children. Although performance seems to be better when there is almost half an octave separation between tones it is also possible that the CI listeners could be perceiving higher amplitude often associated with the high tones. As reported in the acquisition literature generally, adults may use exaggerated pitch contours in speech addressed to children (Cruttenden, 1994, p. 150) but the pitch changes in natural speech in English may be less than half an octave and might not be perceptually salient to implanted children. The natural speech stimuli presented in Experiment II in the current study were not specifically addressed to children so pitch differences may be less than half an octave and so might be less perceptible to the implant subjects.

Similarly, Mandarin tones, although mainly cued by F_0, have some limited temporal cues which might account for better tone identification reported by Peng et al. (2004), and it is reported that pitch height seems to be more perceptually salient than pitch direction (contour). The results of the experiments with tone languages suggest that implanted listeners might be able to hear pitch changes of almost half an octave but this issue needs to be investigated systematically for English. One study of voice similarity (Cleary et al., 2005) investigated how different F_0 and formant differences in English sentences needed to be before two different talkers were perceived by NH and CI children. Results show that performance by CI children was significantly greater than chance in only one condition where linguistic content varied and F_0 differences of 3.5 semitones were audible. However, there was a subgroup of CI children who could hear two different talkers with a difference of 2.7 semitones in one condition, and a difference of 2.17 semitones in another suggesting variability within the group of cochlear implant subjects. There was less variability for the NH group who could hear different talkers when F_0 differences were greater than 19.5 Hz (i.e. 2 – 2.5 semitones). Although the study by Cleary et al. was concerned with voice similarity and not stress and intonation it does give some indication of how big the F_0 differences need to be before two different talkers were perceived by the normal hearing and implanted listeners.

To date there are no other available data for implanted children in English so in the current investigation in Experiment I synthesised pairs of non-meaningful /baba/ stimuli were also presented to the implanted and hearing children in order to establish how big the controlled differences in F_0, duration and amplitude needed to be before they were audible to individual listeners. As discussed above in section 1.11.2 it might be possible to shed some light on whether perception of linguistic contrasts in natural speech stimuli in Experiment II (i.e. *focus* and *compound vs. phrase stress*) is linked up with the ability to hear controlled changes in F_0 (hypothesis (i) in sections 1.1.2 and 1.11.4). On the other hand the results may indicate whether implant users can rely on other cues to stress and intonation such as duration and/ or amplitude in the absence of pitch information (see hypothesis (ii) in sections 1.1.2 and 1.11.4).

Results of studies of the development of tone production in Mandarin speaking 6 to 12 year-old children with cochlear implants (Peng et al., 2004; Xu et al., 2004) report that falling and level tones are acquired before rising tones which was also reported for studies cited earlier of English speaking normal hearing and hearing aid users. In a study of tone production in Cantonese, Barry and Blamey (2004) report smaller inter tonal differences for young CI children (4;2 to 11;3 years) than NH children (aged 3;8 to 6;0 years) and adults. A greater spread of pitch usage for each tone type used by the NH group is reflected in the percentage correct scores rated by listeners (i.e. 78% for the NH group and 38% for the CI group). In Experiment III in the current study measurements of F_0, duration and amplitude in target English words produced by implanted children will indicate the extent to which appropriate changes in F_0 and/or duration and amplitude in the focus words are sufficient to convey focus to a listener.

1.11.4 Perception vs. production of tone, stress and intonation

Perception vs. production of stress and intonation contrasts

An important issue for consideration in the current study is whether implanted children's perception of stress and intonation contrasts is a prerequisite for production. In other words does the appropriate production of intonational contrasts depend on how well implanted children can hear and interpret these contrasts. It is widely accepted that perception precedes production in language development generally but this may not be the case for prosodic development. Although Stackhouse and Wells (1997) suggest that the ability to draw attention to new information is well established by the fourth year, it is possible that children may be able to produce accent and focus in their own speech before they can interpret it in the speech of others (Wells et al., 2004). This supports a previous study by Cutler and Swinney (1987), who suggest that the productions of 3 to 4 year-olds may be apparently similar to productions of 5-6 year-olds because a semantically interesting word generates excitement and tension. They also suggest that a rise in pitch on accented words might be due to a physiological reflex rather than prosodic competence. This may be because the younger group cannot yet process given vs. new, or topic vs. comment but can produce appropriate accentuation to convey focus or new information.

Perception vs. production of tone

Evidence of a similar mismatch between perception and production is also reported in tonal development in Cantonese speaking children (Barry and Blamey, 2004) and although most subjects produced appropriate F_0 contours that could be labelled correct, only a few were judged to be able to produce meaningful tonal differentiation (p. 1747). Studies of perception and production of pitch contours in Cantonese and Mandarin tones can give us some indication of what kind of difficulties might be expected for English implanted children, although it must be borne in mind that Cantonese and Mandarin tones are mainly cued by pitch except for some durational cues in Mandarin tones or increased amplitude in the high tones in Cantonese. Peng et al. (2004) found that a correlation between tone perception and tone production in 6 – 12 year-old children was not found to be significant when high scoring children were removed. The children who performed well in tone production also performed well in tone identification but not the reverse, and the authors conclude that tone identification and production do not develop in parallel and may be associated with duration of implant use and age at implant discussed below in section 1.11.5. Contrary to previous reports which suggest that tone production was better than tone perception, Barry et al. (2002, p. 1747) found that for some of the subjects (age 3;0 – 6;0) tone production and tone perception skills were still developing, and they recommended longitudinal monitoring of tonal development.

Relevance of previous studies of perception vs. production to current study

The children in the experiments on Chinese tones were younger than the children in the current experiment. However, the issues mentioned above will be considered for English speaking implanted children and in the analysis of performance in the perception and production of linguistic focus. Unlike Chinese tones which are cued mainly by F_0, stress and intonation contrasts in English are cued by a combination of F_0, duration and/ or amplitude cues. There are no corresponding studies of focus in English speaking implanted children but it is possible that the developmental issues relating to perception and production normal hearing children in section 1.11.1 might also apply. For example, the physiological reflex referred to earlier (Bolinger, 1983) generating a rise in pitch with excitement and tension associated with an interesting word might occur in implanted children even without being able to hear pitch contrasts and possibly before they can interpret focus in the speech of others.

As set out in the hypotheses in section 1.1.2 and again in section 1.11.4 it is not yet certain whether F_0 really is a necessary cue for the perception of stress and intonation in English. However, like Cantonese speaking implanted children it may be the case that English speaking children with implants are able to produce F_0 contours that sound appropriate but are not produced consistently enough for focus to be considered acquired. As outlined earlier in the discussion of acquisition issues there may be variation and ambiguity across subjects. In Chapter Five the relationship between perception and production of focus in English by CI children will be explored further. For example, if CI talkers can produce appropriate F_0 contours but can only perceive amplitude and/or duration differences through their implants we might expect a correlation between the production of appropriate F_0 in focus words in Experiment III and the perception of duration and/or amplitude in the /baba/ stimuli in Experiment I. Since increased F_0 is generally associated with an increase in amplitude we might also expect a correlation between the production of appropriate amplitude in target focus words in Experiment III and the perception of duration and/ or amplitude in Experiment I. Correlations between the acoustic cues (i.e. F_0, duration and amplitude) which may or may not be used in the perception and production of focus by CI talkers will be analysed and discussed in detail in Chapter Five.

Summary of the hypotheses

It remains to be seen whether F_0 is a necessary cue to stress and intonation, particularly to the intonational contrasts investigated in the present study (i.e. compound vs. phrase stress, and focus). The importance of F_0 as a necessary cue to stress and intonation in English is not clear and straightforward in the literature and the two main hypotheses considered in this present investigation (sections 1.1.2 and 1.11.2) are summarized again below:

hypothesis (i)

> If F_0 *is* a necessary cue to stress and intonation in English, implanted children will need good access to pitch cues (or F_0) in order to hear them if they do not have access to pitch cues, the intonation contrasts will not be accessible to them and so they will not develop abstract phonological representations of compound vs. phrase stress or focus like normal hearing children. Without stored representation of these contrasts they will not learn to produce them appropriately to convey meaning.

hypothesis (ii)

> If on the other hand if F_0 is *not* a necessary cue and plays a less important role in the perception of intonation, implanted children will be able to rely on other cues such as duration and amplitude, which puts them at much less of a disadvantage during early stages of prosodic development. As stated above implanted children will use whatever cues are available to them to develop an abstract prosodic system independent of their ability to hear a particular cue. It is possible that having acquired representation of prominence, they may try to convey focus by producing appropriate increases in F_0 (see physiological reflex above) without necessarily hearing F_0 changes when produced by others. This would support the hypothesis that the intonation contrasts develop as abstract phonological systems which may or may not be perceived or produced by the same cues.

1.11.5 Variables which might affect perception (Experiments I and II) and production (Experiment III) performance: stimulation rate, age at implant, duration of implant use

Variability in results of previous studies: an overview

The effects of variables such as aetiology, communication mode, duration of implant use, age at implant, speech processing strategy, and age on individual performances have been documented in some general outcome studies of speech perception and production skills for English for children (Nikolopoulos, Archbold and O'Donoghue, 1999; Tait and Lutman, 1997; Walzman and Cohen, 2000; Blamey, Sarant, Praatch, Barry, Bow, Wales, Wright, Psarros, Rattigan and Tooher, 2001). Some of these variables also affect outcomes for adult implant users and they are discussed below.

Experiments with adult implant users

Experimental studies of pitch discrimination in adult implant speakers of English (Richardson et al., 1998; Green et al., 2004) and Flemish (Geurts and Wouters, 2001) found that F_0 thresholds varied according to subject, speech processing strategy, and F_0 range. The stimuli presented varied and became more complex and speech-like (i.e. pulse trains, vowels, diphthongs and stress and intonation in natural speech). In Green et al. (2004) discrimination between synthesised vowels varied according to subject, speech processing strategy (i.e. standard CIS and modified strategies), and F_0 range. Poor glide discrimination (i.e. diphthongs) was obtained by some adult implant users even with an octave change in F_0 over the course of the diphthongs. It is suggested that temporal pitch cues were less effective in the presence of dynamically changing spectral structures (i.e. formants) in the diphthongs. Although the results of all these studies indicate limited abilities, adults gain some pitch information from their implants. Given the poor performance of adults above, similar and perhaps increased difficulties might be expected for implanted children using standard speech processing strategies (i.e. SPEAK and ACE). However, many of the adult implant uses above were post-lingually deafened or had progressive hearing losses so received their implants as adults. Many of the children in the current study had pre-lingual deafness and received their implants at an earlier age before plasticity of the central auditory system diminished (Sharma, Dorman and Spahr, 2002; Sharma and Dorman, 2006), so perception performance might be better for younger implanted children.

Experiments with implanted children

Age and duration of implant use

Variability in performance has also been reported in perception and production in the studies of Chinese tones by CI children (see sections 1.8.1 and 1.8.2). For example, in a study of *Mandarin* Chinese tones Peng et al. report that tone *identification* correlated with duration of implant use and tone production correlated negatively with age at implant i.e. there was better tone production by children who received their implants at a younger age. They concluded that factors other than device limitations e.g. plasticity of the central auditory system, need to be considered to explain high level of performance in perception and production of Mandarin tones by some individual CI children. However, studies of *Cantonese* tones Ciocca et al., 2002) report that correlations between tone perception and age at test, duration of implant use, age at implantation, and onset of deafness were not significant. Ciocca et al. concluded that further research was needed to establish whether auditory input or cognitive and linguistic factors contribute to lexical tone discrimination. Barry et al. (2002a, 2000b) also concluded in a study of tonal development in NH and CI subjects that the effects of linguistic development and gradual development of tone needed to be established. A study by Cleary et al. (2005) found a non-significant tendency for later implanted English speaking children to perform more poorly in a talker discrimination task. The authors suggest that variability in the results might be due to other influencing factors such as neural survival or placement of electrodes which are beyond the scope of the present study.

Barry and Blamey (2004) in their study of tone *production* suggest that a tonal system was still developing in the normal hearing 3 - 6 year old children investigated. They also report that F_0 contours were not produced by their 4 – 11 year CI subjects with sufficient frequency to be considered acquired. Xu et al. (2004) in a study of Mandarin tone production conclude that inadequate pitch information delivered through cochlear implants may hinder tone development. They also suggest that other variables such as age at onset of deafness, duration of deafness, age at implantation, and hearing aid usage should be considered.

Results of the studies cited above are not conclusive regarding a correlation between variables such age at implant or duration of implant use. The age range of the normal hearing and implanted subjects in the current investigation extends beyond the subjects in the studies cited above and variables which might affect perception and production skills such as age at implant and duration of implant use will be considered in the analysis of the perception and production performance in Experiments I, II, and III in the current investigation.

Linguistic ability and the use of meaningful vs. non-meaningful stimuli

Barry et al. (2002a) used non-meaningful /wi/ stimuli in their own study, because they suggested that poor performances by the subjects in the study by Ciocca et al. might have been due to the lexical demands of meaningful /ji/ stimuli. Given the wide age range of the subjects in the present study and the inevitable range of linguistic ability this issue is also taken into account in the experiments. Non-meaningful /baba/ stimuli are presented in Experiment I and meaningful natural linguistic stimuli are presented in Experiment II. As mentioned above by Barry et al. the use of non-meaningful stimuli might ensure that subjects were relying on hearing rather than linguistic ability. The advantage of using the non-meaningful synthesised stimuli in the present study is that the smallest discriminable differences in F_0, duration and amplitude between stressed versus unstressed syllables can be investigated in a controlled experiment with groups of NH and CI children within the same age range without any linguistic demands. The natural speech stimuli presented to both groups in Experiment II are produced by speakers varying in gender and age and the F_0, duration and amplitude correlates of stress and intonation are not controlled for each speaker. Experiment II is concerned with the ability of implanted children to use these intonational cues to stress in a linguistic context. A group of age matched normal hearing subjects are also included in the present experiments for comparison with the implanted children.

Stimulation rate

Experiments with implanted children with commonly used speech processing strategies SPEAK (250 pps) and ACE (900 – 1000 pps) in a study of Cantonese tones (Barry et al. 2002a, 2002b) are of particular relevance to the current study as both these strategies are used by the subjects. Barry et al. report that overall tone discrimination for implanted subjects (aged between 4;2 and 11;4 years) was better for SPEAK users whereas the higher stimulation rate of ACE was not found to be an advantage. However, there was more individual variation among ACE users, and Barry et al. (2002b) concluded that more information about pitch direction (i.e. contour) might be available to ACE users whereas SPEAK users might rely more on information about pitch height (i.e. level). Although the ACE users were younger than SPEAK users years of experience was not statistically significant. Peng et al. (2004) found similar tone identification performances by their subjects (aged between 6;0 – 12;6 years) using two device types (MED-EL and Nucleus) despite a shorter duration of implant use. They suggest that this could be due to faster acquisition by the MED-EL group or higher stimulation rate of CIS speech processing strategy than SPEAK in the Nucleus device. Cleary et al. conclude that good performances by some of the children (aged between 5;0 and 12;0 years) using SPEAK, ACE and CIS in their talker identification study suggests that other factors such as neural survival or placement of the electrode array may determine how electrically coded spectral detail is accessed by individuals. Although Cleary et al. found that one CI subgroup performed better, variability across the group was not correlated with speech processing strategy or device.

In the present experiments, only two speech processing strategies are used (i.e. SPEAK and ACE) and comparisons will also be drawn between the performances of children using different stimulation rates in these speech processing strategies. As discussed in section 1.7 carrier pulse trains modulated by the extracted speech envelope are delivered to each electrode at a fixed rate of 250 pulses per second (pps) for SPEAK and between 900 pps and 1000 pps for ACE. There is physiological and psychological evidence that to get a good representation of F_0 range the carrier rate should be at least 4-5 times the modulation rate. For example, if the F_0 range is 80 – 350 Hz the corresponding carrier pulse rate will need to be 1400 pps to get a good representation of F_0 so it might be expected that the faster pulse rate of ACE will provide implant users with better access to F_0 than the slower pulse rate of SPEAK. Reports vary in the studies cited above for example in a study of Cantonese tones Barry et al. report better performance for SPEAK users whereas in a study of talker similarity in English (Cleary et al.) good performances were reported for both ACE and SPEAK users. As the age range of the subjects in the present study is greater than for the studies of Chinese tones, performance in the perception experiments may improve with implant experience for one or both of these strategies and stimulation rates.

1.11.6 CI simulation studies

A vocoder simulation of cochlear implant processing is used in this research to compare the performance of implanted children to normal hearing controls in the discrimination of F_0, intensity and duration differences in synthetic bisyllables. As noted above (section 1.10) details of different vocoders and filters have important effects on access to temporal and spectral cues to pitch and a simulation cannot be considered to represent an exact match to the information provided by a cochlear implant (Laneau, 2004). However, such simulation can nevertheless approximate the reduced spectral and temporal detail that is delivered through a cochlear implant and hence give some basis for age-matched comparisons between implanted and normal hearing children. The NH simulation and the speech processing strategies in the cochlear implants are not identical but there are individual differences anyway between CI subjects such as number of electrodes inserted, frequencies of the channels and the pulse rates. In any case previous simulations show that results with 8 channel and 22 channel simulations are not much different. However, if performance is similar for both groups, difficulties could be related to device or speech processing strategy, but if the normal hearing children in a cochlear implant simulation perform better than implanted children it may suggest that there are other factors affecting implanted children such as neural survival, placement of electrodes, duration of deafness or duration of implant use.

1.11.7 Methodological considerations

The methodologies used in previous studies of children with cochlear implants vary and listener rating scales have been used for tone production (Peng et al., 2004; Xu et al., 2004; Barry and Blamey, 2004), with additional acoustic analysis of the data by some investigators (Barry and Blamey, 2004; Xu et al., 2004). Tone perception studies also use various methods such as live voice procedure (Peng et al.), recorded natural speech stimuli (Ciocca et al., 2002), an adaptive speech feature test in a change no change paradigm with non-meaningful stimuli (Barry et al., 2002a), and resynthesised English sentences presented in a continuum using a variation of an adaptive staircase procedure (Cleary et al., 2005). Some of these procedures are used in the current study which will make it possible to draw comparisons between the results.

1.11.8 The current study

The present investigation includes both early and later implanted children aged between 5;7 years and 16;11 years using two commonly used speech processing strategies (i.e. SPEAK and ACE) in multi-channel implants. Synthesised /baba/ stimuli with different stress positions are presented in two F_0 ranges corresponding to the male and female ranges (Experiment I). The stimuli are also presented to a group of normal hearing children (NH) within the same age range as the CI children in unprocessed and simulated cochlear implant conditions. Prosodic contrasts (compound vs. phrase stress and focus) in natural speech stimuli are also presented in Experiment II to NH and CI children within the same age range. Production of focus on different target words is elicited from the CI subjects in Experiment III and detailed measurements of F_0, duration and amplitude are analysed.

Age at switch-on, age at time of testing, duration of implant use and stimulation rate for the CI subjects will be considered in the analysis of the results. These variables are likely to contribute to differences in performance. For example, some of the children in the current study were implanted during *the sensitive period* of maximal plasticity of the central auditory system of up to 3.5 years (Sharma, Dorman and Spahr, 2002; Sharma and Dorman, 2006) whereas others were implanted at a later stage. None of the implanted children in the current study received their implants under 2;4 years and some were deaf as a result of meningitis ranging from age 2 weeks to 3;0 years. Others had progressive hearing losses and were implanted at different ages up to 15;9 years. The implanted subjects in the current study were the only available children within the age range in the clinical population at the time of testing who could understand the tasks.

It would appear that results are inconclusive in previous studies of pitch and the analysis of the data in the current experiments will take into account developmental and linguistic factors and other variables listed above which might affect perception and production performance for both groups of children across the age range.

Comparison of the perception performances in the linguistic tasks by the normal hearing and implanted groups of children will indicate an expected trajectory of intonational development for implanted children compared to normal hearing children within a similar linguistic environment. Although there is a small number of subjects, they will provide valuable preliminary data for comparison with normative data for other varieties of English, and issues discussed above such as prosodic and intonational development will be taken into account. The relationship between perception and production of stress and intonation contrasts (i.e. compound vs. phrase stress and focus) as well as variables such as age and speech processing strategy will be considered throughout the discussion of the results.

CHAPTER TWO

EXPERIMENT I: SENSITIVITY TO VARIATIONS IN F_0, DURATION AND AMPLITUDE IN SYNTHESISED SPEECH SOUNDS

2.1 Introduction

The relative importance of the physical correlates of stress (F_0, duration and amplitude) has been discussed in sections 1.4 and 1.11.2 and recent experiments have shown that in less controlled conditions F_0 may not necessarily be the most important cue to stress and intonation for normal hearing listeners (Peppé et al., 2000; Kochanski et al., 2005). The aim of Experiment I is to establish minimum F_0, duration and amplitude differences perceived by implant users in pairs of synthesised /baba/ bisyllables. The use of non-meaningful bisyllables avoids potential difficulties relating to age and linguistic ability so that listeners rely on auditory input only and not on linguistic context. As outlined in Chapter One low scores obtained by implanted children in a study of lexical tones in Cantonese could be attributed to the demands of a lexical labelling task (Barry et al, 2002a; Ciocca et al, 2002). The effects of variables such as mode of communication, duration of deafness, aetiology, speech processing strategy, and age, on individual performances are well documented for other general outcome studies of implanted children (Nikolopoulos, Archbold, and O'Donoghue, 1999; Tait and Lutman, 1997; Walzman and Cohen, 2000; Blamey, Sarant, Praatch, Barry, Bow, Wales, Wright, Psarros, Rattigan and Tooher, 2001). Some of these variables will be taken into account in the discussion of the results.

2.2 Methods

2.2.1 Subjects

A total of seventeen implanted children (CI) aged between 5;7 and 16;11 participated in this experiment. All of them were using Nucleus 24 speech processors (8 Sprint, 8 Esprit 3G and 1 Esprit). Ten were using the SPEAK (250 pps) speech processing strategy and 7 were using ACE (600-1800pps). All of the children were in mainstream school except for one who was in a school for the deaf. At the time of testing, duration of implant use ranged from 1;6 to 6;10 years. (See Table 2.1 for individual subject details). Ethical Approval was obtained by the Beaumont Hospital Ethics Committee 2002, and a sample copy of the consent letter to parents of children with implants is in Appendix 2.3. Sixteen normal hearing (NH) children of friends and neighbours in the Dublin area were also included in Experiment I and ages ranged between 6;10 and 17;10 years.

							EXPERIMENT I		EXPERIMENT II		EXPERIMENT III	
subjects	age at switch-on	processor	strategy	stimulation rate (pps)	educational setting	communication mode	age	duration of CI use	age	duration of CI use	age	duration of CI use
C1	7;0	Esprit 3G	Speak	250	Mainstream	Oral/Aural	11;10	4;9	11;11	4;10	12;3	5;2
C2	3;4	Sprint	ACE	720	Mainstream	Oral/Aural	8;0	4;7	8;1	4;8	8;4	4;11
C3	2;5	Sprint	Speak	250	Mainstream	Oral/Aural	6;1	3;8	5;7	3;1	5;9	3;4
C4	3;7	Sprint	ACE	600	Mainstream	Oral/Aural	7;11	4;4	7;11	4;4	7;11	4;5
C5	3;0	Sprint	ACE	1800	Mainstream	Oral/Aural	8;3	5;3				
C6	2;11	Esprit 3G	Speak	250	Mainstream	Oral/Aural	9;0	6;0	8;10	5;10	9;2	6;2
C7	15;9	Esprit 3G	ACE	900	Mainstream	Oral/Aural	17;4	1;6	16;11	1;1	17;1	1;3
C8	7;8	Esprit	Speak	250	Mainstream	Oral/Aural	14;4	6;8	14;1	6;4	14;4	6;7
C9	2;11	Sprint	Speak	250	Mainstream	Oral/Aural	8;3	5;3	8;3	5;4	8;0	5;8
C10	12;6	Esprit 3G	ACE	900	Mainstream	Oral/Aural	13;8	1:3	13;10	1;4	13;10	1;4
C11	3;3	Sprint	ACE	900	Mainstream	Oral/Aural	8;7	5;4	8;1	4;10	8;3	5;0
C12	10;8	Esprit 3G	Speak	250	Mainstream	Oral/Aural	12;8	2;0	12;8	2;0	13;1	2;4
C13	5;3	Sprint	ACE	900	Mainstream	Oral/Aural	7;6	2;3	7;3	2;0	7;5	2;2
C14	4;0	Esprit 3G	Speak	250	Mainstream	Oral/Aural	10;11	6;10	11;0	6;11	11;5	7;4
C15	3;4	Esprit 3G	Speak	250	Mainstream	Oral/Aural	8;9	5;4	8;10	5;5	9;3	5;10
C16	2;5	Sprint	Speak	250	Mainstream	Oral/Aural	6;11	4;5	6;11	4;6	6;11	4;6
C17	12;7	Esprit 3G	Speak	250	School for the Deaf	Oral/TC	14;7	1;11	14;9	2;1	15;2	2;6

Table 2.1 *Details for CI subjects in Experiments I, II and III. Subject 5 was unable to attend for Experiment II and III. Not all subjects completed the experiments in the same order.*

CI subjects	Gender	Onset	Aetiology	500 Hz	1000 Hz	2000 Hz	4000 Hz
				dB HL	**dB HL**	**dB HL**	**dB HL**
C1	male	3 years	Meningitis	>70	>80	>80	>80
C2	female	10 months	Meningitis	>80	>80	>80	>80
C3	female	Congenital	Unknown	55	60	>80	>80
C4	male	3 years	Meningitis	>80	>80	>80	>80
C5	male	Unknown	Unknown	65	55	70	>80
C6	female	2 weeks	Meningitis	75	>80	>80	>80
C7	male	Congenital	Unknown	55	50	50	>80
C8	male	Congenital	Unknown	45	65	>80	>80
C9	female	Congenital	Unknown	50	60	80	>80
C10	male	Congenital	Unknown	45	46	55	50
C11	female	Congenital	Unknown	40	45	60	80
C12	female	Congenital	Unknown	30	40	75	80
C13	male	Congenital	Unknown	45	50	50	50
C14	female	Congenital	CMV	80	>80	>80	>80
C15	male	Congenital	Unknown	55	65	>80	>80
C16	female	2 years	Meningitis	60	65	>80	>80
C17	male	Congenital	Waardenb.	45	50	55	60

Table 2.2 *Onset of deafness, aetiology, and aided pre-operative hearing loss (expressed as dB HL) between 500 and 4000 Hz for individual CI subjects.*

2.2.2 Stimuli

Laryngograph recordings (adult female) were carried out at UCL to provide a reference set of F_0, duration and amplitude measurements. Repetitions of bisyllables, ***BA**ba* with syllable 1 stress (trochaic) and *ba**BA*** with syllable 2 stress (iambic) were recorded on a TEAC DA-P20 DAT recorder. F_0 contours and narrowband spectrograms were generated for different stress and intonation patterns using SFS/WASP (Speech Filing System, Huckvale, 2004) and provided a reference set for setting parameters for the synthesised stimuli. F_0 measurements for each syllable were taken at onset, peak/mid, and offset of voicing. Peak amplitude and duration for each stressed and unstressed syllable were also measured.

2.2.2.1 Syntheses

The KLATTSYN-88 software synthesiser (Klatt and Klatt, 1990) and Speech Filing System (SFS) software (Huckvale, 2004) were used to generate a set of synthesised / /baba/ stimuli with syllable 1 (*BAba)* and syllable 2 *(baBA)* stress. Acoustic cues to syllable stress, i.e. fundamental frequency (F_0) contour, syllable duration, and vowel amplitude, were manipulated in the synthesised bisyllables. In one series all three cues co-varied, and in the others each cue varied in isolation.

F_0 contour series

To generate a rising and falling F_0 contour in the stressed syllable, F_0 was set to rise (linearly) from onset to the temporal mid-point, and fall (linearly) from the mid-point to syllable offset. At this stage onset and offset F_0 values for both syllables were identical and the unstressed syllable had a flat F_0 contour. The onset F_0 value of syllable 1 was either 100 Hz (low male F_0 range) or 200 Hz (high female F_0 range), and the peak F_0 at the mid-point was higher than at onset according to 48 equally spaced multiplicative factors from 1.013 to 1.84 (maximum difference 84%). The F_0 contours for syllable 1 or syllable 2 stress were identical for any given peak F_0 value. To replicate the decline of F_0 in natural speech a declination component with a linear fall in F_0 was added so that F_0 at syllable offset was 0.94 x F_0 at syllable onset. As a result peak F_0 values in stressed syllables depended on stress position (see Figure 2.1). For the F_0 contour series, amplitude for both syllables was fixed by setting the Klatt AV parameter to 50 dB, and duration for both syllables was fixed at 300 ms (see Figure 2.2. (b).

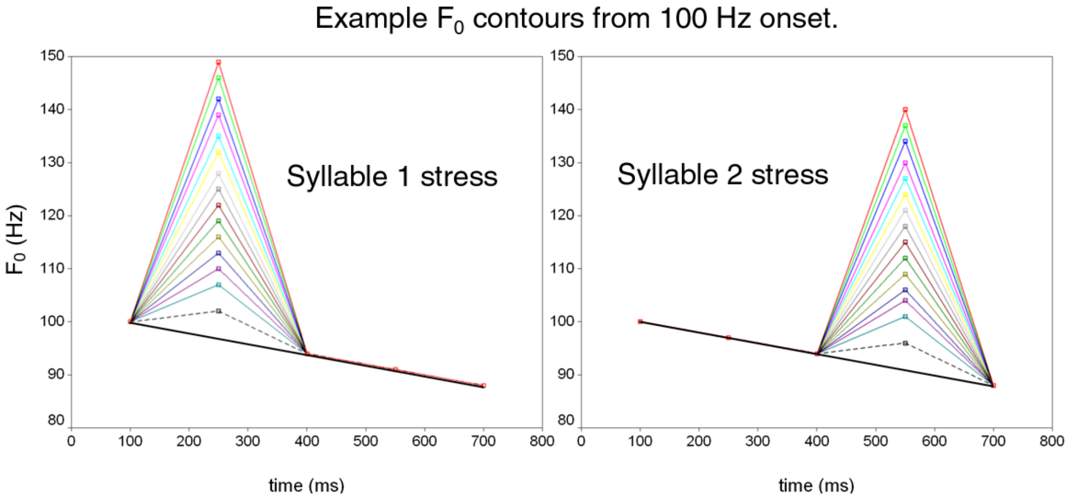
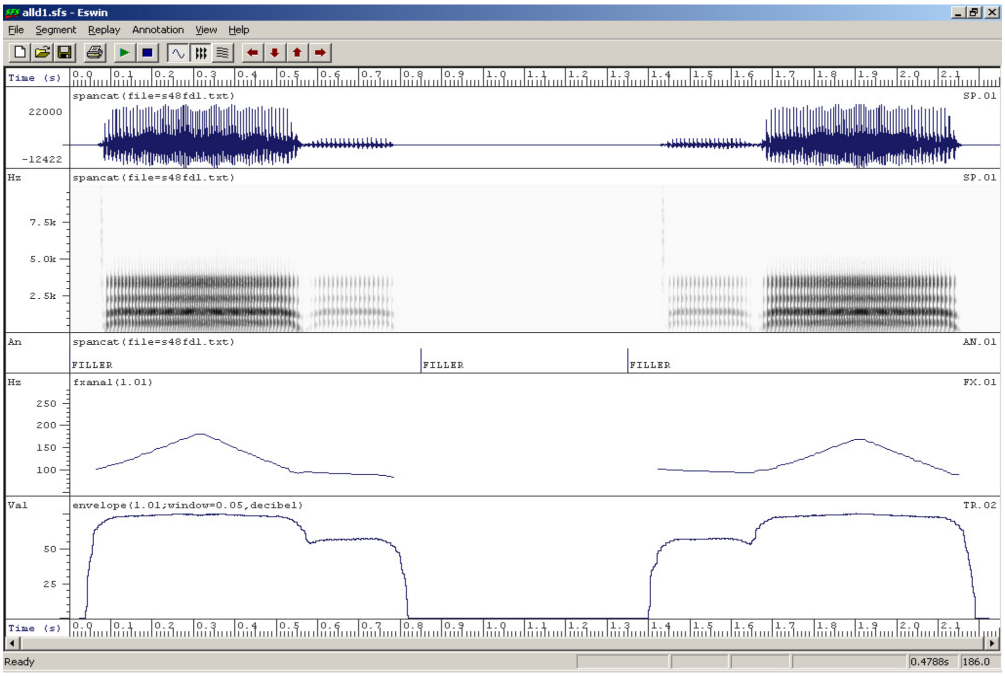


Figure 2.1 *Examples of F_0 contours for syllable 1 stress and syllable 2 stress for two synthesised syllables superimposed on a declination line. Peak F_0 is varied and duration is fixed at 300 ms for both syllables.*

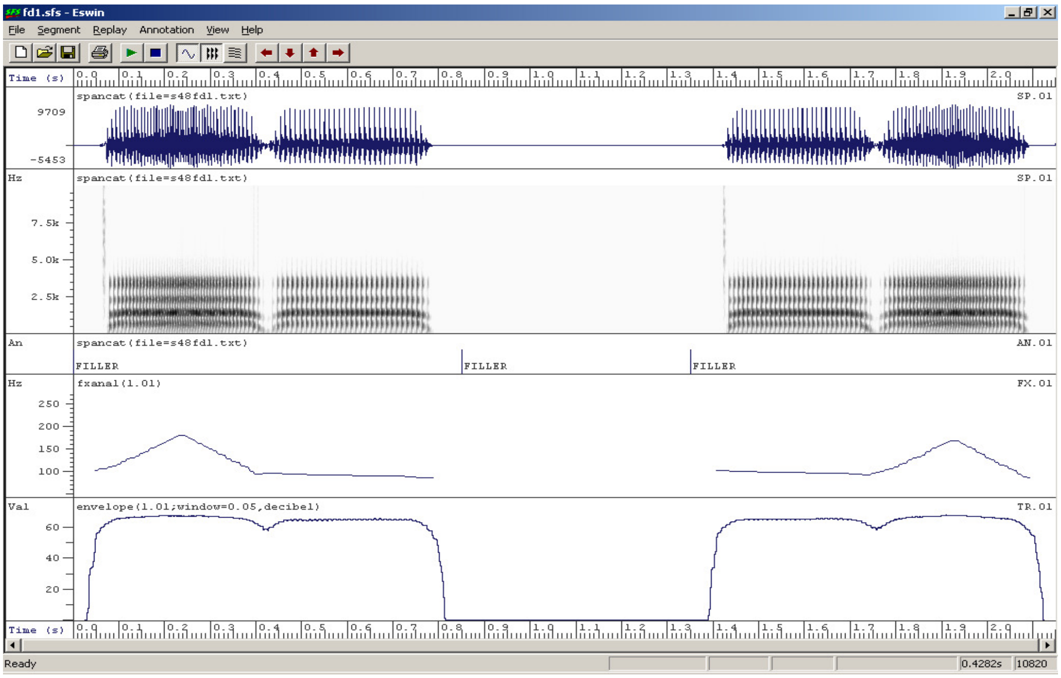
Amplitude series

The Klatt AV parameter was used to vary overall amplitude of the two syllables, and average AV value over two syllables was always 49.5 dB. Difference values for the amplitude series were 1, 3, 5, 7, 9, 11, 13, and 15 dB. The only variation in F_0 was the steady declination with the value at syllable offset always 0.94 of the value of syllable onset. Syllable duration for each syllable was fixed at 300 ms. See Figure 2.2. (c) for an example at the maximum amplitude difference level.

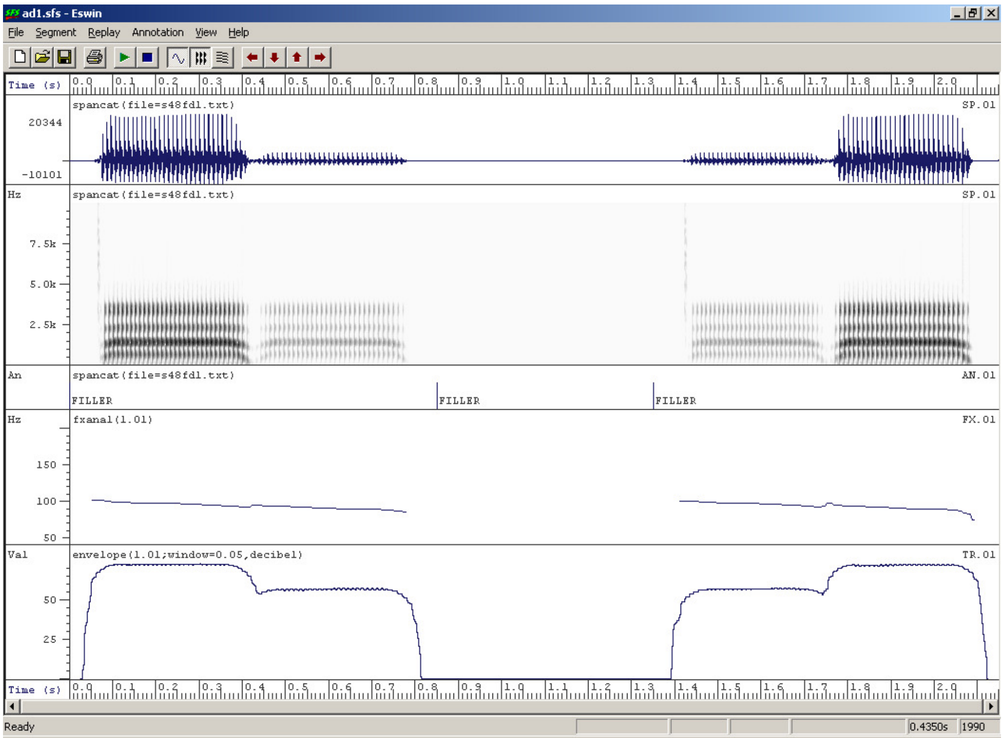
(a) all cues



(b) F_0 only



(c) amplitude only



(d) duration only

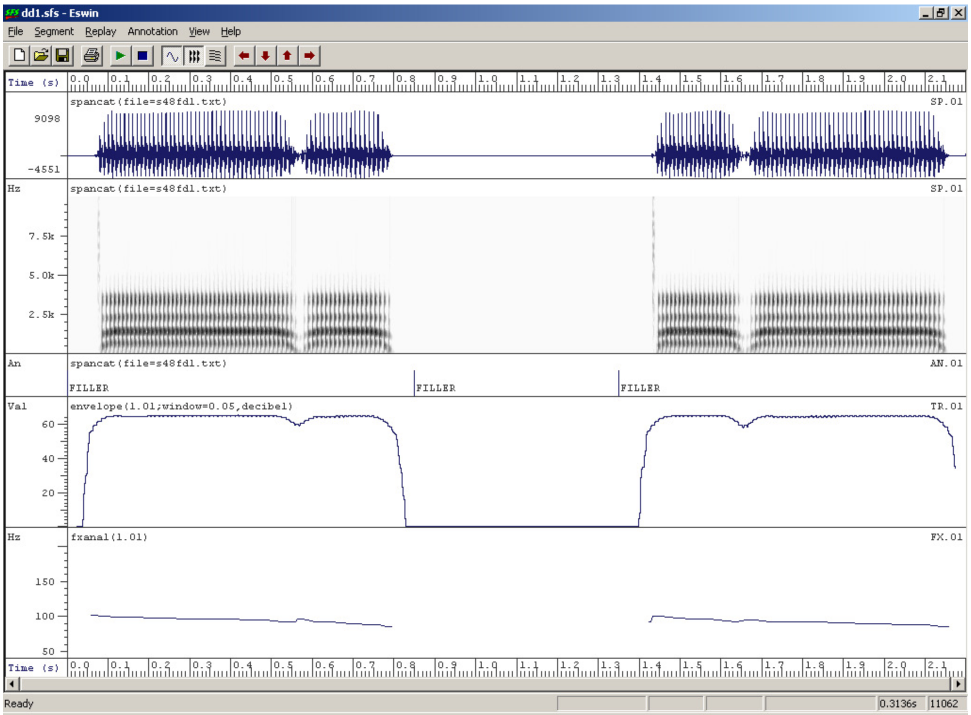


Figure 2.2 *Examples of waveforms, spectrograms, F_0 and amplitude contours for synthesised pairs of bisyllables with the syllable 1 and syllable 2 stress at the maximum difference level for all cues (a), F_0 (Hz) only (b), amplitude (dB) only (c), and duration (secs) only (d).*

Syllable duration series

Overall duration of the two syllables was varied, but average duration was always 300 ms. The duration ratio between stressed and unstressed syllables ranged from 1.02 to 2.38 (maximum difference 138%). The amplitude AV parameter was fixed at 50 dB for both syllables, and the only variation in F_0 was the steady declination with syllable offset always 0.94 of the value of syllable onset. See Figure 2.2. (d) for an example of the maximum duration difference level.

Multiple cue variation series

F_0 contour, amplitude, and duration all co-varied in this series and Appendix 2.1 shows the combinations of F_0 peak height, amplitude difference and duration difference used in the syntheses. The measurements used in these combinations are loosely based on speech recordings described above but were not intended to match the covariation of these cues in natural speech. The multiple cue series was included to provide the listeners with experience with the task and with a more natural stimulus in addition to the series where only one cue varied. See example of all cues varying in Figure 2.2. (a).

Other synthesis parameters

The same vowel formants were used for both F_0 ranges in the syntheses, and Table 2.3 shows the frequency of the first three formants for the vowel steady state drawn from acoustic measurements taken from a male speaker of southern British English. Parameters for the synthesis are shown in Appendix 2.2 where the burst for the first syllable is at time t = 200 ms and the closure between the two syllables is at t = 530 ms.

Talker	Formant frequency
F1 (Hz)	790
F2 (Hz)	1536
F3 (Hz)	2430

Table 2.3 *Measurements for the first three formants of a steady state /a/ vowel drawn from a male speaker of Southern British English.*

Cochlear Implant Simulation

As discussed in section 1.11.6 testing the NH subjects in a CI simulation (CISIM) is useful because we can observe how they perform when certain information is removed or controlled (i.e. F_0, duration and amplitude). If results are similar then difficulties could be related to the device or processing strategy, but if the NH children perform better than CI children there may be other influencing factors such as neural survival, placement of electrodes, duration of implant or duration of implant use.

An acoustic simulation of a cochlear implant was presented to a group of normal hearing children to provide an age-matched comparison for the data from the implanted children. A noise-excited vocoder (Shannon, Zeng, Kamath, Wygonski and Ekelid, 1995; Faulkner, Rosen and Stanton, 2003) was used to generate acoustic stimuli that approximate the spectral and temporal information from a cochlear implant. The simulation used 8 bands covering a frequency range from 100 to 5000 Hz. The band cut-off frequencies for a –3 dB attenuation are shown in Table 2.4.

Band	Lower cutoff (Hz)	Upper cutoff (Hz)
1	100	219
2	219	392
3	392	643
4	643	1006
5	1006	1532
6	1532	2294
7	2294	3399
8	3399	5000

Table 2.4 *The cut-off frequencies (-3 dB attenuation) for 8 bands in a cochlear implant simulation using a noise-excited vocoder (Faulkner et al. 2003)*

Band-pass filters were all sixth-order Butterworth designs, and envelope extraction in each band used half-wave rectification followed by a 400 Hz low-pass smoothing filter (second-order Butterworth). The output for each band was derived from white noise that was first amplitude modulated by the envelope extracted from that band, and subsequently filtered by an identical band-pass filter to the input filter for the band.

2.2.3 Details of testing

2.2.3.1 Adaptive threshold measurement

A two-alternative forced-choice ‘same/different’ discrimination task was used to measure just detectable threshold differences in F_0, duration and amplitude in the four synthetic series discussed above. On any given trial, subjects were presented with two /baba/ bisyllables, with 600 ms silence between the two. For 50% of trials, selected at random, the two bisyllables were identical. Stress position varied between the 2 bisyllables on the remaining 50% of trials, and within each trial the cue representing stress position had a constant value. The order of stress positions within the pair was selected randomly. Subjects indicated their perception of the two bisyllables by clicking on one of two pictures representing the ‘same’ or ‘different’ on a computer screen.

A 2-down 1-up staircase (Levitt, 1971) was used to increase the difference between the pair of bisyllables after each incorrect response and to decrease the difference after two successive correct responses, thus converging on 70.7% correct. After 10 reversals the staircase procedure ended. However, if subjects obtained 8 successive incorrect responses at the maximum or 8 successive correct responses at the minimum stimulus difference that was possible, or if 100 trials were completed before 10 reversals occurred, the procedure also ended. The threshold was estimated from the mean of the stimulus differences at the last 6 reversal points at the end of each staircase.

2.2.3.2 Procedure

All implanted children (CI group) were tested in purpose-built audiology booths and the normal hearing children (NH group) were tested in a quiet room at home. Ambient noise level was monitored with a hand held Monacor SM-4 sound level meter. Stimuli were delivered via a Dell C640 laptop computer connected to a Fostex 6301B Powered Speaker. Laptop and speaker volume controls were preset at 70-75 (SPL) and the speaker was placed one metre from the child’s ear or microphone.

The different series (conditions) for the CI and NH groups are summarized in Table 2.5. All four series were presented in the low F_0 range, and in the high F_0 range, only the multiple cue and F_0 series were presented.

	Summary synthesised `/baba/` series	**Cues**
1	Multiple cue variation series	all cues varying (F_0, duration, amplitude)
2	F_0 contour series	F_0 varying (duration and amplitude fixed)
3	Syllable duration series	duration varying (F_0 and amplitude fixed)
4	Amplitude series	amplitude varying (F_0 and duration fixed)
	F_0 ranges	
1	low (male) F_0 range with initial onset value at 100 Hz	
2	high (female) F_0 range with initial onset value at 200 Hz	

Table 2.5 *Summary of the synthesised /baba/ series presented to the cochlear implant (CI) and normal hearing (NH) subjects in Experiment I. The multiple cue and F_0 contour series were presented in the low and high F_0 ranges. An additional set of the same series was presented to the NH group in a cochlear implant simulation.*

As described above the stimuli were delivered in an adaptive 2-down 1-up procedure. Each child worked individually and at the start of each series, a pair of pictures representing *same/different* appeared on the computer screen. The child responded to the stimulus by clicking on the appropriate picture with a mouse. At the beginning of each series the task was explained and each child was given an opportunity to listen to examples of the stimuli in each series at 8 different difficulty levels covering the range of 48 levels presented in the test. Once the test started each child worked independently without prompting and each subtest lasted 5-10 minutes. There was no time limit and each child worked at his own pace, but younger children required more supervision and breaks between each series than older children. The series in the low F_0 range were presented first followed by the series in the high F_0 range. The order of presentation for each series varied randomly within each range for each subject. This procedure was repeated for the CI group and where possible two sets of each series were completed. However, the total number of series and repetitions completed varied according to the age and concentration of the subject.

The NH children were presented with one set of each the above series in the low and high F_0 ranges. In addition, they were presented with a cochlear implant simulation of each series as described above. Twelve different series were presented to the NH group in total (see Table 2.5). The series in the low F_0 range were presented first and then the high F_0 range. Each unprocessed series was followed by the same series in a cochlear implant simulation condition. The order of presentation for each unprocessed and simulation pair varied randomly within each range for each subject.

2.3 Results

Individual and group results are presented below, and difference thresholds for the F_0, duration and amplitude conditions are discussed separately for the NH and CI subjects. The vertical axes, upon which thresholds are plotted, are expressed in percentage change for peak F_0 and duration. Amplitude differences are expressed in decibels (dB*).* Where two sets of each series were completed by the CI children, minimum and maximum difference thresholds are presented with the mean thresholds in the individual graphs.

2.3.1 F_0 difference thresholds

2.3.1.1 Cochlear implant

Figure 2.3 shows minimum, maximum and mean difference thresholds for individual implanted (CI) children for two sets of the F_0 series in the low and high F_0 ranges. In the low F_0 range mean scores show that all but subject 1 failed to hear F_0 peak differences of less than 40% (0.5 octave) and ten subjects performed at or close to the maximum difference at 84%. Although difference thresholds were generally not much different for the high (female) and low (male) F_0 ranges, the group results in Figure 2.4 show that variability in the high F_0 range (5% -84%) was nearly twice that of the low range (40% -84 %). Eight subjects could hear peak F_0 differences of 40% or less (i.e.15%, 20%, and 25%) in the high F_0 range.

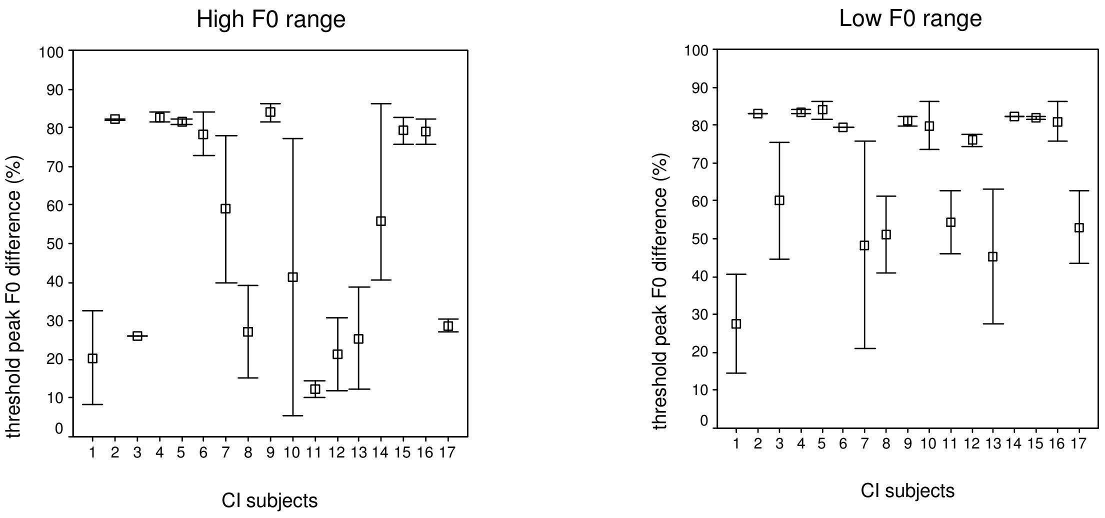


Figure 2.3 *Mean peak F_0 difference thresholds for individual CI subjects in low and high F_0 ranges. Minimum and maximum thresholds are presented as whiskers where two sets of each series were completed.*

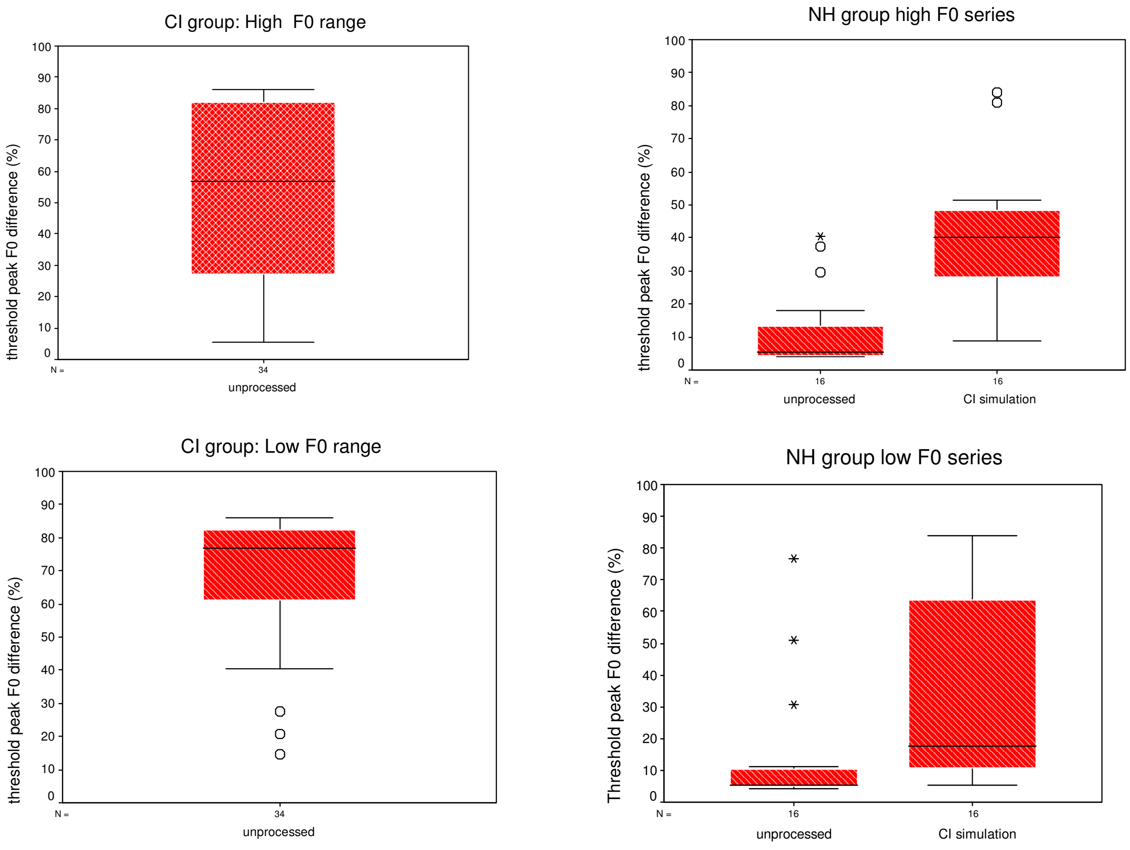


Figure 2.4 *F_0 difference thresholds for low and high F_0 ranges for the CI group on the left and for the NH group in the unprocessed and simulation conditions on the right.*

2.3.1.2 Normal hearing simulation condition

Group performance for the NH group for the simulation condition to the right of Figure 2.4 was more variable in the low F_0 range (5% - 84%) whereas most were hearing differences less than 52% in the high F_0 range.

2.3.1.3 Normal hearing unprocessed condition

The NH group results to the right of Figure 2.4 were similar for both unprocessed F_0 ranges. In the low F_0 range difference thresholds for most were below 10% and for the high F_0 range below 15%.

2.3.1.4 Summary

Although difference thresholds for the CI subjects were not much different for the high and low F_0 ranges, variability in the high F_0 range (5%-84%) was greater than that of the low F_0 range (40%-84%).

Performance for most NH subjects was similar for the low (5%-10%) and high (5%-15%) F_0 ranges in the unprocessed conditions, and performance in the unprocessed condition was better than in the CI simulation condition. In the CI simulation condition peak F_0 thresholds were much more variable (i.e. 5%-84 % in the low F_0 range and 10-52 % in the high F_0 range) but most NH subjects were hearing F_0 differences of 52 % or less in the high F_0 range.

In the low F_0 range, most CI talkers could only hear F_0 differences above 60% whereas most of the NH group could hear F_0 differences of less than 60% in the simulation condition. In an independent samples t test the difference between the CI (unprocessed condition) and NH (CI simulation condition) was found to be significant (equal variances not assumed $p<.001$). In the high F_0 range thresholds were more variable for the CI subjects in the (5%-84%) than the NH subjects in a simulation condition (10–52%). However in an independent samples t test the difference between the CI group and NH group in the simulation condition was not found to be significant ($p=.198$).

A test of analysis of variance (ANOVA) of within-subject effects over two groups (i.e. CI and NH in the simulation condition) showed that F_0 range had no significant effect on thresholds [F(1,31) = 1.418, p=0.243)]. However the interaction of F_0 range and the CI/NH simulation groups showed that the effect of F_0 range was very different for the two groups [F(1, 31) = 9.68, p =0.004]. Tests of between-subjects effects with high and low F_0 ranges averaged together showed a significant difference between the groups [F(1,31) = 8.27, p =0.007)]. Pairwise comparisons for the two groups using a Bonferroni adjustment for multiple comparisons within each F_0 range showed that the two groups are significantly different (p=0.001) in the low F_0 range but not in the high F_0 range (p=0.208). Pairwise comparisons (also using a Bonferroni adjustment) for the two F_0 ranges within each group showed a significant difference for the CI group (p=0.004) at the $p<0.05$ level but not for the NH group (p=0.191).

2.3.2 Duration difference thresholds: CI group vs. simulation vs. unprocessed conditions for the NH group

In this section duration difference thresholds for the low F_0 range are presented below for individual and group CI and NH subjects. Durational differences are expressed in percentages in the vertical axes in the graphs.

2.3.2.1 Cochlear implant

Figure 2.5 shows individual minimum, maximum and mean duration difference thresholds in two sets of the duration series for individual CI children in the low F_0 range only. There was some variability in the mean duration difference thresholds for individual CI children with 8 subjects showing thresholds below 30%, and 4 subjects in excess of 80% up to maximum difference at 138%. This is also reflected in Figure 2.6 for the CI group with duration thresholds ranging from 5% up to maximum level at 138%.

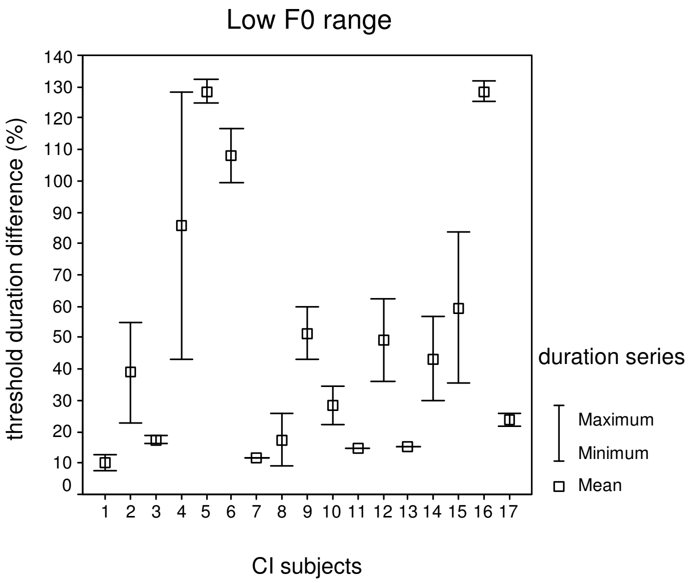


Figure 2.5 *Minimum, maximum and mean threshold duration differences between syllable 1 and syllable 2 stress for individual CI subjects in two sets of each series.*

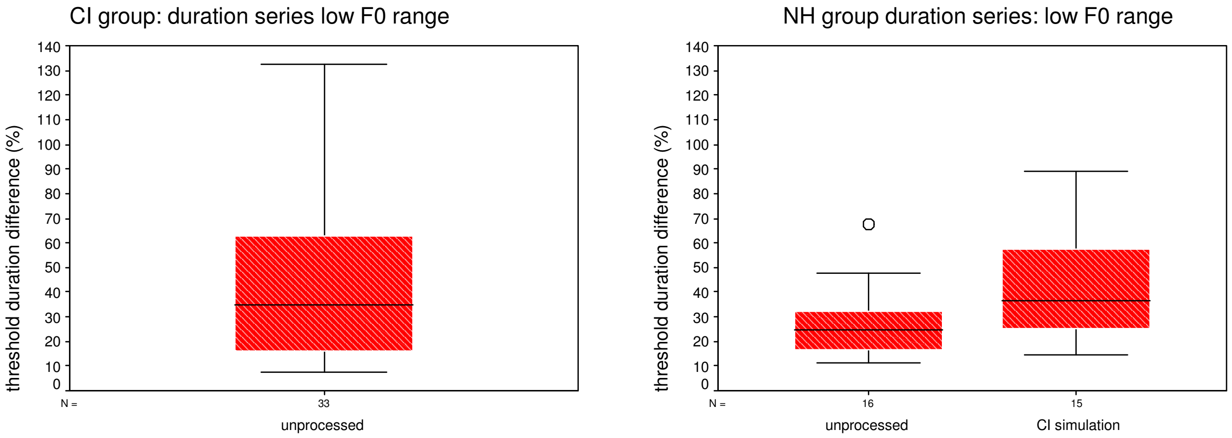


Figure 2.6 *Duration difference thresholds in the lower F_0 range for the CI group and for the NH group in the unprocessed and CI simulation conditions.*

2.3.2.2 Normal hearing simulation condition

In Figure 2.6, duration thresholds in the CI simulation condition only for NH subjects varied from 15%-90% in the low F_0 range. There was more variation for the CI group (5%-138%) with some individuals hearing slightly smaller differences than the NH group. However, Figure 2.6 shows that most subjects in these two groups could hear duration differences less than 60%.

2.3.2.3 Normal hearing unprocessed condition

Figure 2.6 shows that most of the NH group in the unprocessed condition could hear duration differences less than 48% (with one exception at 70%), and some could hear slightly smaller differences (10%) than in the simulation condition (15%).

2.3.2.4 Summary

Overall duration difference thresholds varied in the low F_0 range for the CI group from 5% up to maximum difference at 138%. There was variation for the NH subjects in the unprocessed condition (10% - 48%) and in the simulation condition (15%-90%) with some doing slightly better in the unprocessed condition. When the CI and NH in a CI simulation are compared most subjects in each group could hear differences less than 60% with a few CI subjects hearing slightly smaller differences, an independent samples t test showed that the difference between the two groups was not significant (p=.514).

2.3.3 Amplitude Difference Thresholds: CI group vs. simulated and unprocessed conditions for the NH group

In this section individual and group amplitude difference thresholds for CI and NH subjects in the low F_0 range are presented below, and in the vertical axes in the graphs amplitude differences thresholds are expressed in decibels (dB).

2.3.3.1 Cochlear implant group

Individual minimum, maximum and mean amplitude difference thresholds for CI children are presented in Figure 2.7 for the low F_0 range only. The results show variability across subjects with three subjects (subjects 1, 15, 17) showing mean difference thresholds at and below 5 dB, and seven subjects at or close to the maximum difference at 12-15 dB. The majority of CI subjects, however, could hear differences of less than 12 dB. Group results for the CI subjects in Figure 2.8 show the range of variability for the CI group with difference thresholds ranging from 3 dB up to maximum level at 15 dB.

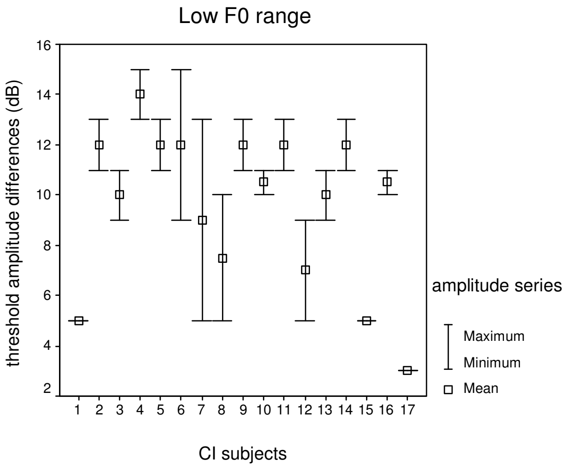


Figure 2.7 *Minimum, maximum and mean threshold amplitude differences for syllable 1 vs. syllable 2 stress for individual CI subjects in pairs of /baba/ stimuli.*

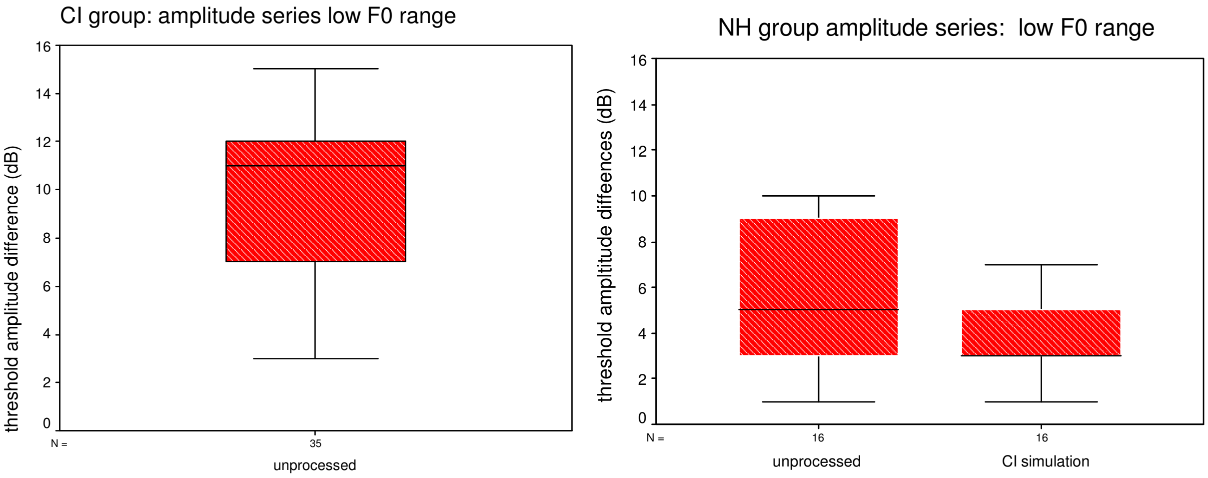


Figure 2.8 *Amplitude difference thresholds in the lower F_0 range for the CI subjects and for the NH subjects in the unprocessed and simulation conditions.*

2.3.3.2 Normal hearing simulation condition

In the simulation condition to the right of Figure 2.8 the NH subjects could hear differences ranging from 1 dB – 7 dB in the low F_0 range.

2.3.3.3 Normal hearing unprocessed condition

Thresholds for the NH group in the unprocessed condition in the low F_0 range presented at the bottom of Figure 2.8 show variability in performance with some subjects performing worse than in the simulation condition (1 dB - 10 dB).

2.3.3.4 Summary

Surprisingly, performance for the NH group was a somewhat better in the simulation (1-7 dB) than in the unprocessed amplitude condition (1 dB-10 dB) and it was considered it might be due to a practice effect because the simulation condition was always presented after the unprocessed condition (see section 2.3.3.5 below). There was more variability for the CI group generally (3 dB -15 dB) and performance for the NH group in a CI simulation was better (1 dB – 7 dB). In an independent samples t test comparing the CI group and NH group in the simulation condition, the difference between the two groups was significant ($p<.001$).

2.3.3.5 Learning effect

The better amplitude thresholds for the NH group in a simulation condition suggested a possible practice effect as a result of order of presentation i.e. unprocessed followed by the simulation condition. However, the duration series were presented to the NH group in a similar order and there was no evidence of a practice effect. There was also no evidence of a practice effect for the CI group who completed two of each series but not immediately following each other. Thresholds in the second run were slightly better or worse for some subjects and similar for others, and only one subject (CI) performed better in the second run of the duration and F_0 series in the high and low ranges.

2.3.4 Correlations between F_0, duration and amplitude thresholds

2.3.4.1 CI subjects

In a Pearson correlation test for the CI group (Table 2.6), correlations were significant for the CI group with Bonferroni correction ($p< 0.05$) between F_0 thresholds in the high and low F_0 ranges and between duration thresholds and F_0 thresholds in the both F_0 ranges. When age was controlled for the correlation between duration and F_0 thresholds remained in the high F_0 range but was only approaching significance ($p = 0.005$) in the low F_0 range which suggests some developmental effect. However, Table 2.6 shows that there was no evidence of any correlation between age, duration of CI use, or stimulation rate (in the speech processing strategies SPEAK or ACE) and minimum difference thresholds in the F_0, and duration and amplitude series for the CI children in Experiment I.

CI Subjects: Pearson Correlations for Experiment I: Bonferroni corrected significance level = 0.0023								
		High F_0	**Duration**	**Amplitude**	**Age**	**Age at switch-on**	**Duration of Implant use**	**Stimulation rate**
Low F_0	Pearson Correlation	0.722	0.684	0.471	-0.391	-0.400	0.242	0.070
	Sig. (1-tailed)	**0.001**	**0.001**	0.028	0.060	0.056	0.174	0.394
	N	17	17	17	17	17	17	17
High F_0	Pearson Correlation		0.721	0.420	-0.360	-0.417	0.330	0.124
	Sig. (1-tailed)		**0.001**	0.047	0.078	0.048	0.098	0.318
	N		17	17	17	17	17	17
Duration	Pearson Correlation			0.476	-0.447	-0.474	0.318	0.181
	Sig. (1-tailed)			0.027	0.036	0.027	0.107	0.243
	N			17	17	17	17	17
Amplitude	Pearson Correlation				-0.465	-0.489	0.328	0.390
	Sig. (1-tailed)				0.030	0.023	0.099	0.061
	N				17	17	17	17

CI subjects: Partial Correlation Coefficients controlling for age in Experiment I: Bonferroni corrected significance level = p=0.036						
		High F_0	**Duration**	**Amplitude**	**Duration of Implant use**	**Stimulation rate**
Low F_0	Coefficient	0.677	0.619	0.355	0.106	0.056
	df	-14	-14	-14	-14	-14
	P (1-tailed)	**P= .002**	P= .005	P= .089	P= .348	P= .419
High F_0	Coefficient		0.671	0.306	0.220	0.114
	df		-14	-14	-14	-14
	P (1-tailed)		**P= .002**	P= .125	P= .206	P= .337
Duration	Coefficient			0.339	0.175	0.179
	df			-14	-14	-14
	P (1-tailed)			P= .100	P= .259	P= .254
Amplitude	Coefficient				0.180	0.416
	df				-14	-14
	P (1-tailed)				P= .252	P= .055

Table 2.6 *Pearson correlations with partial correlations controlling for age at Experiment I are presented in two separate tables above for the CI subjects.*

.

2.3.4.2 NH subjects

CI simulation condition

In a Pearson correlation test (see Table 2.7) for the NH subjects in the CI simulation condition correlations with Bonferroni correction were significant when age was controlled (p= 0.001) between F_0 thresholds in the low and high F_0 ranges. The correlation between duration thresholds and F_0 thresholds with Bonferroni correction was approaching significance (p = 0.002) for the high F_0 range only.

Unprocessed Condition

In the unprocessed conditions for the NH talkers the correlation between F_0 thresholds in the high and low F_0 ranges with Bonferroni correction (p= 0.001) disappeared when age was partialled out (p= 0.006).

Comparisons between CI and NH subjects

Similar correlations between F_0 thresholds in the high and low F_0 ranges were found for both the CI group and NH group in the simulation condition when age was factored out whereas the correlation disappeared for the NH subjects in the unprocessed condition indicating age effects. These results indicate that ability to hear smaller differences in F_0 may have been affected by device limitations for both the CI and the NH subjects in the simulation condition. Although duration thresholds correlated with F_0 thresholds in the high F_0 range for both of these groups there was a weaker correlation for the NH in the simulation condition which remained when age was partialled out. No correlation was found between duration thresholds and F_0 thresholds in the low F_0 range for the NH subjects in the simulation condition whereas for the CI group a correlation between duration thresholds and F_0 thresholds in the low F_0 range with Bonferroni correction was weaker (p = 0.005) when age was partialled out.

NH Subjects: Pearson Correlations for Experiment I

		High F_0	**Low F_0 CISIM**	**High F_0 CISIM**	**Duration**	**Duration CISIM**	**Amplitude**	**Amplitude CISIM**	**Age**
Low F_0	Pearson Correlation	0.692	0.724	0.774	0.534	0.497	-0.060	-0.101	-0.327
	Sig. (1-tailed)	**0.001**	**0.001**	**0.001**	0.017	0.030	0.412	0.355	0.108
	N	16	16	16	16	15	16	16	16
High F_0	Pearson Correlation		0.329	0.632	0.358	0.508	0.149	0.164	-0.394
	Sig. (1-tailed)		0.107	0.004	0.087	0.027	0.290	0.272	0.066
	N		16	16	16	15	16	16	16
Low F_0 CISIM	Pearson Correlation			0.662	0.236	0.588	0.290	0.103	-0.043
	Sig. (1-tailed)			0.003	0.189	0.011	0.138	0.352	0.438
	N			16	16	15	16	16	16
High F_0 CISIM	Pearson Correlation				0.427	0.697	0.107	0.090	-0.554
	Sig. (1-tailed)				0.050	0.002	0.346	0.370	0.013
	N				16	15	16	16	16
Duration	Pearson Correlation					0.460	-0.393	-0.332	-0.422
	Sig. (1-tailed)					0.042	0.066	0.104	0.052
	N					15	16	16	16
Duration CISIM	Pearson Correlation						-0.019	**0.001**	-0.135
	Sig. (1-tailed)						0.474	0.500	0.315
	N						15	15	15
Amplitude	Pearson Correlation							0.693	-0.144
	Sig. (1-tailed)							**0.001**	0.298
	N							16	16
Amplitude CISIM	Pearson Correlation								0.144
	Sig. (1-tailed)								0.297
	N								16

CISIM = Cochlear Implant Simulation **Correlation is significant at p = 0.0014 using a Bonferroni significance level p<0.05**

NH subjects: Partial Correlations controlling for age at Experiment I

		High F_0	**Low F_0 CISIM**	**High F_0 CISIM**	**Duration**	**Duration CISIM**	**Amplitude**	**Amplitude CISIM**
Low F_0	Coefficient	0.648	0.755	0.781	0.479	0.483	-0.115	-0.061
	df	12	12	12	12	12	12	12
	P (1 - tailed)	P= .006	**P= .001**	**P= .001**	P= .042	P= .040	P= .347	P= .418
High F_0	Coefficient		0.338	0.552	0.246	0.497	0.101	0.240
	df		12	12	12	12	12	12
	P (1 - tailed)		P= .119	P= .020	P= .198	P= .035	P= .366	P= .204
Low F_0 CISIM	Coefficient			0.773	0.275	0.593	0.285	0.093
	df			12	12	12	12	12
	P (1 - tailed)			**P= .001**	P= .171	P= .013	P= .162	P= .376
High F_0 CISIM	Coefficient				0.343	0.730	0.023	0.165
	df				12	12	12	12
	P (1 - tailed)				P= .115	P= .002	P= .469	P= .287
Duration	Coefficient					0.453	-0.510	-0.272
	df					12	12	12
	P (1 - tailed)					P= .052	P= .031	P= .174
Duration CISIM	Coefficient						-0.036	0.028
	df						12	12
	P (1 - tailed)						P= .452	P= .462
Amplitude	Coefficient							0.734
	df							12
	P (1 - tailed)							**P= .001**
CISIM = Cochlear Implant Simulation				**Correlation is significant at p=0.0018 using a Bonferroni significance level p<0.05**				

Table 2.7 *Pearson correlations with partial correlations controlling for age at Experiment I are presented in two separate tables above for the NH subjects.*

2.4 Summary and Discussion of the Results

In this section the findings of Experiment I are summarized and the implications are discussed. Comparisons are drawn between the current results and those of other previous relevant studies.

2.4.1 Fundamental Frequency (F_0)

2.4.1.1 Comparisons between F_0 discrimination by CI group and by the NH group in the unprocessed condition

In the F_0 series in Experiment I, peak difference thresholds were not much different for the two F_0 ranges for the CI group but as shown in Figures 2.3 and 2.4 there was greater variability in the high F_0 range (5%-84%) compared to the low F_0 range (40% - 84 %). Most CI children seem to have difficulty hearing F_0 differences of less than half an octave and some of them may not be hearing differences even at the maximum difference level (84%). However, in the high F_0 range some were hearing smaller F_0 differences. In contrast with this there was less variability for the NH subjects in the unprocessed F_0 series, and most were hearing differences of 10% or less in the low F_0 range and less than 15% in the high F_0 range.

2.4.1.2 Implications of the results for the perception of prosodic contrasts?

If F_0 is a necessary cue to stress and intonation in English (see hypothesis (i) in section 1.1.2 and also 1.11.4) these results have serious implications for most of the CI subjects and their ability to hear or even acquire linguistic contrasts such as focus or compound stress if F_0 changes are greater than half an octave. However, the alternative view supported by some recent studies of natural speech discussed in section 1.11.2 suggests that F_0 is not a necessary cue to stress and intonation (see hypothesis (ii) in section 1.1.2 and 1.11.4). If this is the case children with cochlear implants will be at less of a disadvantage during the acquisition process despite the pitch limitations, and they might be able to rely on other cues (e.g. duration and amplitude discussed below) to help then acquire and hear prosodic contrasts such as compound vs. phrase stress and focus. It remains to be seen whether the perception of linguistic stimuli in Experiment II are linked with their ability to hear smaller F_0, duration or amplitude differences in Experiment I.

2.4.1.3 Are results different from previous findings in studies of implanted adults and children and why might this be?

In a previous study of Cantonese tones by Barry et al. (2002a) tone discrimination was also found to be significantly better for the NH group than the CI group in the discrimination of tone contrast but unlike the present study variability was reported across both groups. The results of F_0 series for the CI subjects in Experiment I are similar to results of a study of Cantonese tones by Ciocca et al. (2002) in that a large average F_0 separation of tones was also required by implanted children. However, overall performance was poor and was above chance for only three out of eight tonal contrasts when there was an F_0 separation of 35 Hz or 45 Hz which in this study was just above or below half an octave when one of a pair of tones was a high tone. In other words implanted children needed almost half an octave difference before they could discriminate between pairs of tones, but it has also been suggested that listeners could be responding to higher amplitude associated with higher tones. Tone discrimination by implanted children in Mandarin (Peng et al., 2004) was also better for pairs of tones when one was a high tone but it is suggested that shorter duration of one tone (T4) may have provided an additional duration cue.

Better F_0 discrimination was reported in a study of resynthesised English sentences by Cleary et al. (2005). In that study CI subjects could hear two different talkers when there was an F_0 difference of 30 Hz (3.5 semitones) whereas NH subjects only needed 19.5 Hz (2-2.5 semitones). However there was also a sub-group of CI children who could hear F_0 differences which were audible to the NH listeners. Although this study was concerned with voice similarity and not stress and intonation, it does give us some indication that smaller F_0 differences than the current Experiment I thresholds were needed by their CI subjects to be able to hear two different talkers. In experiments with post-lingually deafened adults Geurts and Wouters (2001) reported smaller F_0 threshold differences than the present study with subjects perceiving F_0 differences between pairs of synthetic /a/ or /i/ vowels i.e. between 6 and 20 Hz in the lower F_0 range and between 12 and 19 Hz in the higher F_0. Individual thresholds in that study varied according to subjects, processing strategy and F_0 range. Both the Cleary et al. and the Geurts and Wouters study differ from the present one in that the

F_0 difference was present through the stimuli rather than at a momentary peak as here, and this may be a factor in the differences seen.

2.4.1.4 Comparisons with the typical acoustic changes in natural speech: F_0

As the F_0 changes in natural speech are unlikely to be more than half an octave, most CI listeners will have difficulty hearing F_0 cues to stress and intonation. This is borne out by the F_0 measurements for the natural speech stimuli in the present study (see Section 3.5.4.1 and Appendix 3.2) which show that in general the F_0 differences between the target focus words and the neighbouring words were less than or just above half an octave, and rarely approached or exceeded an octave (see Talker 2 for ***MAN:*** *paint* 11.88 semit. and Talker 3 for ***EAT***: bone 16.37 semit., and in an extreme case *paint*: ***BOAT*** 26.04 semit.). The boxplots in Appendix 3.3 also indicate that the spread of F_0 differences between focus and neighbouring words rarely exceeded half an octave in focus in focus position 1 (initial position) except for one sentence (i.e. *the man is driving a car*), and were always less than half an octave in focus position 3 (i.e. final position). Experiment I results suggest that CI listeners will have difficulty hearing F_0 differences in the natural speech stimuli in Experiment II.

2.4.1.5 F_0 *discrimination by the NH in a CI Simulation*

As discussed in section 1.11.6 one of the advantages of a cochlear implant simulation is that we can observe how these children perform when certain information is removed (i.e. F_0, duration or amplitude). As indicated in Figures 2.3 and 2.4 in Experiment I in the current study, some NH children in a CI simulation were hearing smaller F_0 differences than some of the CI group in the low F_0 range, and an independent samples t test (Section 2.3.1.4) found a significant difference ($p<0.001$) between these two groups. Most NH subjects in the simulation could hear differences less than 60% whereas most CI subjects could *not* hear differences less than 60%. In the high F_0 range there was greater variability for the CI subjects than the NH subjects in the simulation condition, but the difference between the two groups in an independent samples t test was not found to be significant. In a test of analysis of variance (ANOVA) pairwise comparisons within each F_0 range show that the two groups were significantly different in the low F_0 range only. The slightly better performance in the high F_0 range for a few CI subjects might be because these subjects were responding to spectral information in the different formant structure of the vowels in the stressed and unstressed syllables in the pairs of synthetic /baba/ stimuli. This is in contrast with Green et al. (2002, 2004) who report poorer glide labelling performance by both implanted adults and by normal hearing adults in simulation studies for the higher F_0 ranges in synthetic diphthongs with dynamically changing formant structures.

However, as suggested by Laneau et al. (2004) results of simulation studies should be interpreted with caution as different vocoders and filters in a cochlear implant simulation may have important effects on temporal and spectral cues and may not represent an exact match for information provided by a cochlear implant. In general simulation studies are useful in that they mimic the limited spectral resolution and unresolved harmonics of speech processing strategies. As stated in section 1.11.5 some of the CI subjects in the current study received their implants at an early age during the period of maximum plasticity, and there are individual differences between CI subjects such as number of electrodes inserted, frequencies of the channels and pulse rates. In the current study the poorer performance by the CI group compared to the NH group in a CI simulation in the low frequency range might be accounted for by factors other than device limitations such as duration of deafness or implant use (discussed below) or other factors beyond the scope of this investigation such as placement of electrodes or neural survival.

2.4.2 Discrimination of duration and amplitude cues by NH and CI subjects

As discussed earlier in 1.1.2 and in 1.11 it is unclear whether F_0 is a necessary cue to stress and intonation or whether implant users rely on duration and amplitude cues to hear prosodic contrasts such as focus. The purpose of the amplitude and duration /baba/ series in Experiment I was to establish minimum duration and amplitude difference thresholds in the lower F_0 range for the CI group as well as the NH group in the unprocessed and simulation conditions. The results might indicate whether duration or amplitude might provide reliable cues to stress and intonation in the absence of F_0 cues through the implant.

2.4.2.1 Duration

Variability occurred across CI subjects (5%- 138%) in the duration series in the low F_0 range and across the NH subjects in the unprocessed condition (10%-48%) and the simulated condition (15%-90%). However, the boxplots in Figure 2.6 show that performance for the NH group in the simulation condition was similar for most of the CI group who could hear duration differences less than 60%. When the NH group in the simulation condition was compared with the CI group in an independent samples t test (Section 2.3.2.4) the difference between the two groups was not found to be significant ($p = 0.514$). These results suggest that duration may be a more reliable cue to listeners in the absence of F_0 information via a cochlear implant or a simulation of a cochlear implant.

Comparisons with typical acoustic changes in natural speech: duration

In natural speech it may be the case that some CI subjects use duration as a cue to stress and intonation in the absence of F_0 information through the implant. The duration measurements in Appendix 3.5 and the boxplots in Appendix 3.6 for the NH focus stimuli (presented in Section 3.5.4 in Experiment II) give us some idea of changes in duration that might be expected in focus words in natural speech. The median duration measurements for three of the four sentences (i.e. all except *the girl is baking a cake*) were consistently longer in the target focus words/syllables than when they were not in focus. As discussed earlier in Section 2.4.1.4 most F_0 differences between the focus words and neighbouring words were less than half an octave (especially in final position) and so would not be accessible to most CI listeners according to Experiment I results. Since the range of duration thresholds in Experiment I was 5% -138% and most CI listeners could hear duration differences of 60% in Figure 2.6, some of the median duration differences in the NH stimuli in the boxplots in Appendix 3.6 would be accessible to them e.g. ***BOY*** (75%), ***DOG*** (75%) ***MAN*** (120%) ***BONE*** (150%) ***DRIVE*** (80%) ***CAR*** (140%). There were eight CI subjects who could hear duration differences of 30% or less and so smaller median duration differences between the focus and unfocussed target words would be accessible to these listeners e.g. ***PAINT*** (20%), ***BOAT*** (25%). In one sentence (i.e. *the girl is baking a cake*) however there were only minimal changes in the median duration differences for ***BAKE*** and ***CAKE*** which might not be accessible to most CI listeners.

2.4.2.2 Amplitude

In the amplitude series in the low F_0 range (see Figure 2.8), mean threshold differences varied across the CI subjects from 3 dB up to the maximum difference level of 15 dB but the majority could hear differences of less than 12 dB, and so some CI subjects might be able to rely on amplitude changes in target focus words in natural speech. In the simulation condition the NH group performed better with threshold differences ranging from 1 dB to 7 dB, whereas in the unprocessed condition thresholds ranged from 1 dB to 10 dB. In an independent samples t test the difference between the CI group (3 dB –15 dB) and the NH in a simulation condition (1 dB – 7 dB) was found to be significant (see Section 2.3.3.4).

Comparisons with typical acoustic changes in natural speech: amplitude

As stated earlier Appendix 3.2 and boxplots in Appendix 3.3 show that in final focus position and in other positions, F_0 differences between the target focus word and the neighbouring words were less than half an octave and probably inaccessible to most implanted subjects. The boxplots in Appendix 3.8 show a step up in the median amplitude differences for each of the stimulus sentences ranging between 4 dB and 9 dB to the final focus position and might be a more reliable cue to focus than F_0 for some CI listeners (see Section 3.5.4.3)..

2.4.3 Were there any correlations between F_0, duration and amplitude thresholds for CI and NH subjects in a simulation condition?

The NH group in the simulation condition (CISIM) resembled the CI group (see Tables 2.6 and 2.7) when age was controlled and correlations were found between F_0 thresholds in the high and low F_0 ranges. However, there were some differences between these groups. For example there was no correlation between duration thresholds and F_0 thresholds in the low F_0 range for the NH subjects in the simulation condition even when age was partialled out and a weak correlation with Bonferroni correction ($p = 0.002$) remained between duration and F_0 thresholds in the high F_0 range. For the CI subjects when age was partialled out a significant correlation between duration and F_0 thresholds in the high F_0 range remained but the correlation between duration thresholds and F_0 thresholds in the low F_0 range with Bonferroni correction was only approaching significance ($p = 0.005$). For both groups correlations between F_0 thresholds and duration thresholds in the high F_0 range remained when age was partialled out. In other words ability to discriminate differences in F_0 in the high F_0 range correlated with ability to hear differences in duration. For the CI subjects only the correlation between F_0 discrimination in the low F_0 range and ability to hear duration differences was approaching significance when age was controlled.

2.4.4 Did factors such as age, duration of implant use, practice and stimulation rate affect performance in Experiment I?

2.4.4.1 Age and duration of implant use

As indicated in Tables 2.6 and 2.7 no correlations were found for the NH subjects in a simulation condition between age at time of testing and F_0, duration and amplitude thresholds. For the CI subjects also there were no correlations between F_0, duration or amplitude thresholds and age at testing, age at switch-on, duration of implant. Ciocca et al. (2002) also found in their study of Cantonese tones that correlations with age at test, age at implant and use of implant were not significant (section 1.11.5). In contrast with this Peng et al. (2004) found that identification of Mandarin tones correlated with duration of implant use although this could be ascribed to age effects in the use of duration cues which are not found in Cantonese tones.

2.4.4.2 Stimulation Rate

In the present study there was no correlation between stimulation rates of SPEAK and ACE speech processing strategies and F_0, duration and amplitude thresholds in Experiment I. Similarly, Ciocca et al. also reported that ACE users even with higher pulse rates (900 –1000 pps) still had difficulty recognising lexical tones and Barry et al. (2002a) anticipated that ACE users in their study might have performed better but there was no significant difference between strategies (section 1.8). Overall in these studies the SPEAK group performed better and the higher stimulation rate was not found to be an advantage for ACE group. Although the ACE users were younger than the SPEAK group the duration of implant use was not found to be statistically significant.

2.4.4.3 Other contributing factors

As the boxplot in Figure 2.6 indicates, the CI group and the NH in the simulation condition in the duration series were similar in that most could hear duration differences less than 60%. However, in the boxplots in Figure 2.8 the NH group performed significantly better in the simulation condition in the amplitude series in the low F_0 range than the CI group and this suggests that there could be other contributing factors besides device limitations beyond the scope of the current study such as position of the electrodes, neural survival, as well as the normal hearing ability of the NH subjects which provided stimulation of the auditory pathway.

2.4.5 Questions arising from Experiment I results

Questions arising from the results of Experiment I to be considered in Chapter Three are whether

a. CI children can hear prosodic contrasts in natural speech stimuli in Experiment II given that they cannot hear F_0 differences of less than half an octave between pairs of /baba/ syllables in Experiment I

b. the ability to hear differences in stress and intonation in natural speech stimuli is correlated with the ability to hear smaller F_0 and/or duration and amplitude differences

c. the results of Experiments I and II indicate differences between NH and CI groups such as

    (i) differences in the acoustic cues (F_0, duration, amplitude) used to hear prosodic contrasts such as focus or compound vs. phrase stress

    (ii) whether the ability to hear any of these acoustic cues determines the perception of prosodic contrasts in Experiment II

2.5 Appendices

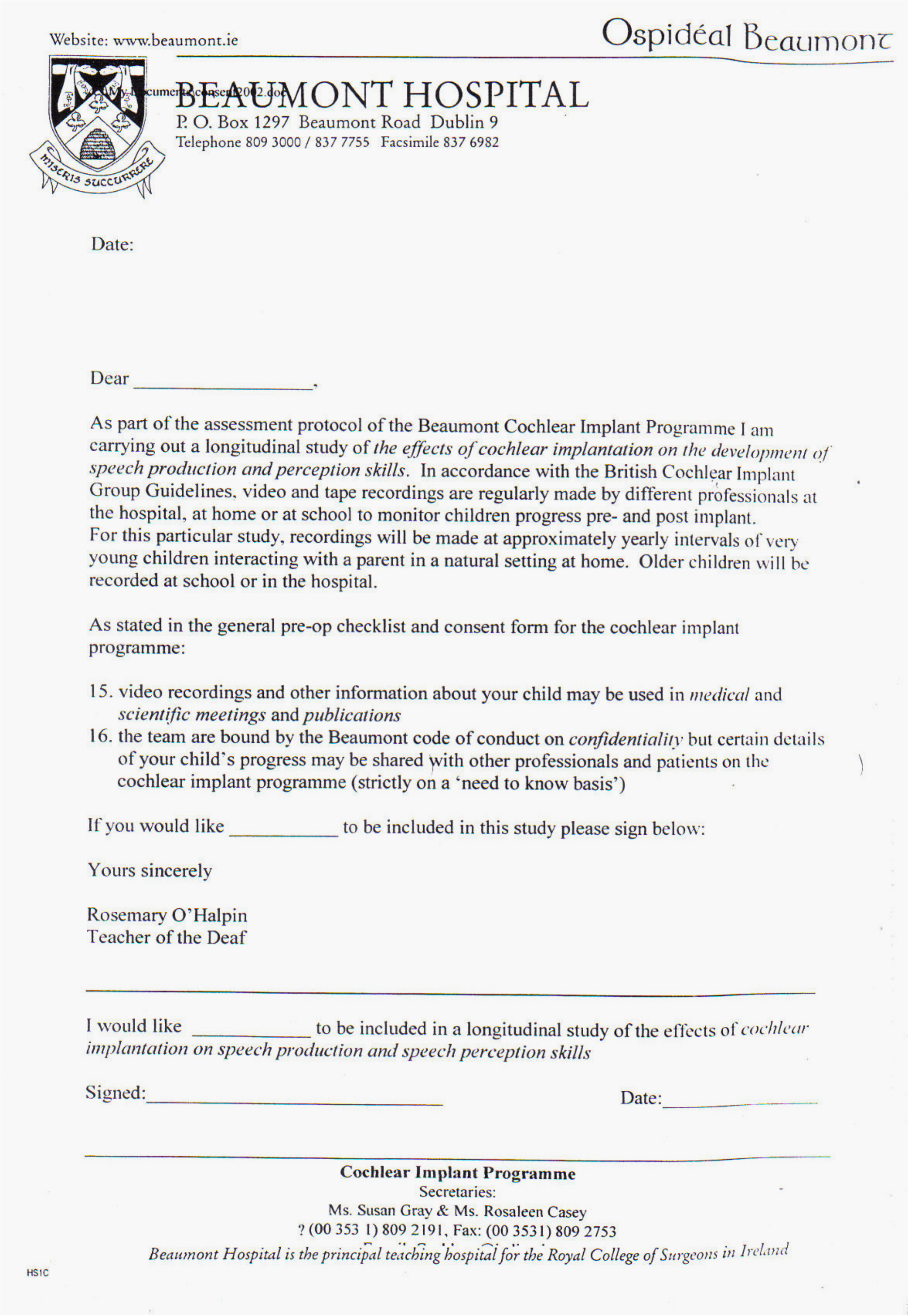
Continuum level	Peak F_0/onset F_0	amplitude difference (dB)	long/short duration
1	1.013	1	1.017
2	1.026	1	1.037
3	1.039	1	1.055
4	1.052	1	1.073
5	1.065	1	1.094
6	1.079	1	1.113
7	1.093	3	1.135
8	1.107	3	1.154
9	1.121	3	1.178
10	1.135	3	1.197
11	1.150	3	1.221
12	1.164	3	1.242
13	1.179	5	1.267
14	1.194	5	1.288
15	1.209	5	1.309
16	1.225	5	1.336
17	1.240	5	1.358
18	1.256	5	1.380
19	1.272	5	1.409
20	1.288	7	1.436
21	1.305	7	1.460
22	1.321	7	1.484
23	1.338	7	1.514
24	1.355	7	1.544
25	1.373	7	1.569
26	1.390	9	1.595
27	1.408	9	1.626
28	1.426	9	1.652
29	1.444	9	1.684
30	1.462	9	1.712
31	1.481	9	1.744
32	1.500	10	1.773
33	1.519	11	1.815
34	1.538	11	1.850
35	1.558	11	1.872
36	1.578	11	1.908
37	1.598	11	1.944
38	1.618	11	1.981
39	1.639	13	2.014
40	1.660	13	2.053
41	1.681	13	2.092
42	1.703	13	2.132
43	1.724	13	2.172
44	1.746	13	2.214
45	1.769	15	2.245
46	1.791	15	2.288
47	1.814	15	2.332
48	1.837	15	2.376

Appendix 2. 1 *Multiple cue variation series showing combinations of F_0 peak height, amplitude difference, and duration difference that were used in the syntheses.*

Time (ms)	AV (dB)	AF (dB)	F1 (Hz)	F2 (Hz)	F3 (Hz)	AB (dB)
190	0	0	200	1100	2080	63
195	17	25	322	1187	2171	63
200	33	50	443	1273	2263	63
205	50	0	565	1360	2354	63
210	50	0	610	1385	2362	63
215	50	0	655	1410	2371	63
220	50	0	700	1435	2379	63
225	50	0	745	1461	2388	63
230	50	0	790	1486	2396	0
235	50	0	790	1511	2405	0
240	50	0	790	1536	2413	0
245	50	0	790	1536	2422	0
250	50	0	790	1536	2430	0
Values constant for steady start part of syllable 1 from 250 to 455 ms						
455	50	0	790	1536	2430	0
460	50	0	790	1536	2428	0
465	50	0	790	1536	2421	0
470	50	0	790	1532	2413	0
475	50	0	790	1510	2406	0
480	50	0	790	1488	2398	0
485	50	0	790	1466	2390	0
490	50	0	790	1444	2383	0
495	50	0	775	1422	2375	0
500	50	0	700	1400	2368	0
505	50	0	625	1378	2360	0
510	50	0	547	1347	2340	0
515	47	0	456	1282	2272	0
520	45	0	364	1217	2203	0
525	42	0	273	1152	2135	0
530	41	0	218	1113	2094	0
535	43	0	310	1178	2162	0
540	46	0	401	1243	2231	0
545	48	0	492	1308	2299	0
550	50	0	576	1366	2356	0
555	50	0	633	1395	2365	0
560	50	0	689	1425	2375	0
565	50	0	745	1454	2384	0
570	50	0	790	1483	2394	0
575	50	0	790	1513	2403	0
580	50	0	790	1536	2413	0
585	50	0	790	1536	2422	0
590	50	0	790	1536	2430	0
Values constant for steady start part of syllable 2 from 590 to 795 ms						
795	50	0	790	1536	2430	0
800	50	0	790	1536	2427	0
805	50	0	790	1536	2419	0
810	50	0	790	1527	2412	0
815	50	0	790	1505	2404	0
820	50	0	790	1483	2397	0
825	50	0	790	1461	2389	0
830	50	0	790	1439	2381	0
835	50	0	760	1417	2374	0

840	50	0	685	1395	2366	0
845	50	0	610	1373	2359	0
850	45	0	529	1334	2327	0
855	33	0	437	1269	2258	0
860	20	0	346	1204	2190	0
865	8	0	255	1139	2121	0
870	0	0	200	1100	2080	0

Appendix 2.2 *Variation of the first three formants for /a/ vowel steady state, with a burst located at time t= 200ms for the first syllable and the closure between the two syllables at t= 530 ms.*



Website: www.beaumont.ie

Ospidéal Beaumont

My Documents\consent2002.doc

BEAUMONT HOSPITAL
P. O. Box 1297 Beaumont Road Dublin 9
Telephone 809 3000 / 837 7755 Facsimile 837 6982

Date:

Dear ____________________,

As part of the assessment protocol of the Beaumont Cochlear Implant Programme I am carrying out a longitudinal study of *the effects of cochlear implantation on the development of speech production and perception skills*. In accordance with the British Cochlear Implant Group Guidelines, video and tape recordings are regularly made by different professionals at the hospital, at home or at school to monitor children progress pre- and post implant.
For this particular study, recordings will be made at approximately yearly intervals of very young children interacting with a parent in a natural setting at home. Older children will be recorded at school or in the hospital.

As stated in the general pre-op checklist and consent form for the cochlear implant programme:

15. video recordings and other information about your child may be used in *medical* and *scientific meetings* and *publications*
16. the team are bound by the Beaumont code of conduct on *confidentiality* but certain details of your child's progress may be shared with other professionals and patients on the cochlear implant programme (strictly on a 'need to know basis')

If you would like ____________ to be included in this study please sign below:

Yours sincerely

Rosemary O'Halpin
Teacher of the Deaf

I would like _____________ to be included in a longitudinal study of the effects of *cochlear implantation on speech production and speech perception skills*

Signed:_________________________________ Date:_______________

Cochlear Implant Programme
Secretaries:
Ms. Susan Gray & Ms. Rosaleen Casey
? (00 353 1) 809 2191, Fax: (00 3531) 809 2753

Beaumont Hospital is the principal teaching hospital for the Royal College of Surgeons in Ireland

HS1C

Appendix 2.3 *Ethical approval was granted by Beaumont Hospital Ethics Committee 2002 and consent was obtained from parent(s) to carry out the experiments (see sample letter above).*

CHAPTER THREE

EXPERIMENT II: SENSITIVITY TO VARIATIONS IN STRESS AND INTONATION IN NATURAL SPEECH STIMULI

3.1 Introduction

The gradual acquisition of stress and intonation in English has already been discussed in Chapter One. There is a general agreement in the literature (e.g. Atkinson-King, 1973; Vogel and Raimy, 2002; Wells et al., 2004) that the perception of stress contrasts such as focus, and compound vs. phrase stress may continue to develop beyond 12;0 years, and it is also suggested that some stress contrasts might never be acquired even in adulthood (Peppé et al., 2000). Because of weak pitch cues available through current speech processing strategies it is possible that implant users rely more on timing and loudness cues.

In Experiment I, listeners had to rely on listening ability only when discriminating between pairs of non-meaningful /baba/ stimuli whereas in Experiment II, the subjects have to identify lexical items with different stress and intonation patterns in a linguistic context.

The aims of Experiment II are to

a. investigate the speech perception abilities of implanted (CI) and normal hearing (NH) children in picture identification tasks involving focus, and compound vs. phrase stress in natural speech stimuli.
b. compare the performances of the CI children with the NH children taking into account factors such as age at time of testing, age at switch-on, duration of CI use, speech processing strategy, and other acquisition issues raised in the review of the literature in Chapter One.
c. establish whether the CI and NH groups of children are responding to the same or different perceptual cues (pitch, timing and loudness) to lexical stress and focus using acoustic measurements of the perception stimuli in Chapter Three.

3.2 Methods

3.2.1 Subjects

A total of sixteen implanted (CI) children from different parts of the Irish Republic participated in Experiment II. The details are the same as for Experiment I (see Table 2.1) except for one subject (subject 5) who was unable to attend for Experiment II tests. Twenty two normal hearing subjects (NH) aged between 5;9 and 16;11 years also participated, and five of them were also included in Experiment I. Eight of the normal hearing children were siblings of the implanted children, and were not involved in Experiment I.

3.2.2 Stimuli

Talkers

Two male (age 16 and 20 years) and 2 female (age 12 and 27 years) speakers of Southern Irish English from Dublin were recorded individually in an anechoic room with a low noise floor at UCL using a Bruel & Kjaer 2231 sound level meter fitted with a 4165 microphone cartridge. A Laryngograph processor was used to record an Lx signal fed to the line input of a Sony DTC-60ES DAT recorder with a sampling rate set to 44.1 kHz. Picture prompts appeared on a screen in front of individual talkers in the anechoic room and each task was explained, and they were instructed to give particular types of responses as described below. There was no time limit and each talker worked at his/her own pace. For the three sub-tests in Experiment II, three different types of stimuli were recorded as shown in Table 3.1, and they are referred to as Phrase Test (compound vs. phrase stress), Focus 2 (focus in two element phrases), and Focus 3 (focus in three element phrases).

Design of the Stimuli

Focus 2 Test

Two element (***Focus 2***) and three element sentences (***Focus 3***) were included in the focus tests in Experiment II. The shorter two element sentences (*Focus 2*) have only two target focus items which reduces the memory load for CI listeners, whose task is to decide whether they hear first or second position focus (e.g. *BLUE* book vs. blue *BOOK*). This is not unlike the task in Experiment 1 which also involves first or second position stress in pairs of /baba/ syllables. However, in Experiment I non-meaningful syllables are used with controlled changes in F_0, duration and amplitude whereas in Experiment II, meaningful two word phrases with shifting focus are presented where F_0, duration and intensity are not controlled. Other factors come into play especially in final position such as boundary markers or turn delimitation which may compete with focus on the final item.

Phrase Test

Although the ***Phrase test*** involves two elements the task for listeners is quite different from Focus 2 as they have to decide whether they hear a phrase with two separate elements (blue *BELL*) or a compound (*BLUE*bell). As discussed earlier in section 1.11.2 differences between compound vs. phrase stress may not be signalled in the same way by different speakers and pitch movement and pitch reset may not be as reliable cues as lengthening and pause.

Focus 3 Test

The advantage of ***Focus 3*** test is that there are three target words with two pre-final focus items which do not compete with boundary markers and/or turn delimitation on the final focus item. Unlike Focus 2 there are unstressed syllables in between the target focus words or syllables which may help the focus words stand out to listeners as a result of a step up or change in F_0, or duration, or amplitude. However, the changes in F_0 on the target words against the natural decline of F_0 will be accessible to normal hearing listeners but it remains to be seen whether implanted subjects can perceive these changes on the focus words or whether they can make use of duration or amplitude cues.

Elicitation of the data

A structured approach was taken to elicit full SVO (i.e. subject +verb+ object) sentences for the Focus 3 rather than elliptical sentences from the four NH talkers for consistency and to facilitate statistical analysis. The use of a schwa /ə/ in unstressed syllables, and the realization of /t/ as a fricative /s̠/ in Hiberno English (e.g. in boat) by the NH talkers adds to the naturalness of the SVO stimuli. The use of picture prompts is commonly reported in the literature (e.g. Peng et al., 2004; Ciocca et al., 2002) and a question and answer sequence (Xu and Xu, 2005; O'Halpin, 2001; Parker, 1999; King and Parker, 1980; Atkinson-King, 1973) or mini dialogue rather than reading aloud or imitation task (Snow, 1998). In this way the responses might be as close to spontaneous speech as possible while maintaining control over experimental variables such as the vocabulary, sentence type or target focus item. Other methods used with older hearing subjects and reported in the wider literature such as retelling a story or a map task or spontaneous conversation (Kochanski et al., 2005; Dalton and Ní Chasaide, 2007) would be too challenging for the younger implanted children who might be delayed in prosodic, pragmatic and semantic development.

The advantage of using simple declarative svo sentences is that the stimuli should not present additional linguistic difficulties to the younger children and could be used right across the age range of the subjects (O'Halpin, 1993, 2001). Ellipsis can sometimes occur in natural speech (e.g. Q: *Is the DOG painting the boat? A: No the BOY is….)* but complete sentences with focus on one word for emphasis or contrast in response to a question are not unusual. For consistency and ease of analysis, full sentences were elicited from the NH talkers in the perception stimuli for Experiment II as well as production data from the CI talkers in Experiment III (see Chapter Four). To make responses as spontaneous as possible, picture prompts were also used in the Phrase test to elicit a compound or noun phrase (i.e. bluebell vs. blue bell) and in the Focus 2 test to elicit focus or contrastive stress in adjective+ noun phrases (*e.g. it's a BLUE door*) in response to questions in mini dialogues (*e.g. Is it a GREEN door?).* Both elliptical (*e.g. No, it's BLUE*) and full responses occur in natural speech but for consistency full adjective + noun phrases were elicited from the NH talkers for the perceptual stimuli in Experiment II. For consistency and measurement in the future the first item from each set of repetitions was selected where possible for the Experiment II subtests unless it was poor quality, ambiguous, or unmeasureable.

PHRASE TEST		
Compound	**Phrase**	
give me the bluebell	give me the blue bell	
give me the blackboard	give m the black board	
give me the greenhouse	give me the green house	
give me the redhead	give me the red head	
give me the bluebottle	give me the blue bottle	
give me the hotdog	give me the hot dog	
FOCUS 2 TEST		
it's a **BLUE** book	it's a blue **BOOK**	
it's a **GREEN** door	it's a green **DOOR**	
FOCUS 3 TEST		
the **BOY** is painting a boat	the boy is **PAINTING** a boat	the boy is painting a **BOAT**
the **GIRL** is baking a cake	the girl is **BAKING** a cake	the girl is baking a **CAKE**
the **MAN** is driving a car	the man is **DRIVING** a car	the man is driving a **CAR**
the **DOG** is eating a bone	the dog is **EATING** a bone	the dog is eating a **BONE**

Table 3.1 *Summary of the natural speech stimuli recorded by four talkers for Phrase, Focus 2, and Focus 3 speech perception tests in Experiment II.*

Phrase Test (48 items)

Six compound versus phrase pairs (e.g. bluebell vs. blue bell) were recorded in a carrier sentence *give me the* _____. Two pictures appeared side by side on a screen in front of the talker for each compound vs. phrase. It was considered less confusing if the test stimulus was recorded in sentence-final position for cochlear implant listeners. Three repetitions of each stimulus were recorded together and a total of 144 items were recorded for the four talkers. The talkers were given time to practice and were instructed to avoid listing intonation in their responses (i.e. a rise in pitch at the end of each elicited item indicating the speaker is not yet finished or there is more to come as in days of the week or counting or a list of names). Instead talkers were encouraged to produce each item as an independent entity and unrelated to the next picture prompt with neutral intonation with a natural decline in F_0. A total of 48 items were selected for the perception test.

Focus 2 Test (16 items)

Two pictures (i.e. *a green door* and *a blue book)* were presented separately on the screen in front of the talkers and they were asked questions *(e.g. is it a GREEN book?*) designed to shift focus (contrast) and elicit a specific pattern (i.e. no*, it's a BLUE book).* Each talker was asked the same set of four questions six times in random order. A total of 94 phrases were recorded for the four talkers and 16 items were selected for the perception test.

Focus 3 Test (48 items)

Four pictures corresponding to the three element phrases were presented separately to the talkers. Each talker was asked three types of question for each picture (e.g. *is the GIRL painting the boat?*) designed to shift focus (contrast) in three element declarative sentences and produce specific patterns (i.e. *no, the BOY is painting the boat*). There were four pictures in total and the talkers were asked the same sets of questions six times in random order. A total of 288 sentences were recorded from all the talkers and 48 items were selected for the perception test.

Stimuli

The prosodic contrasts in the present study (i.e. compound vs. phrase stress and focus discussed above) are of particular interest as they have been investigated in a few studies of normal hearing subjects but not yet for children with cochlear implants. However, studies of other prosodic contrasts in English (Titterington et al., 2006) and in Mandarin Chinese (Peng et al., 2004) suggest that implanted children follow the same order of acquisition as normal hearing children but are delayed. As discussed in section 1.3.2 for normal hearing children compound vs. phrase stress is acquired gradually up to 12;0 or 13;0 years (Wells et al., 2004; Atkinson-King, 1973; Vogel and Raimy, 2002) but there are differences in reports regarding the age at which focus is acquired. For example, Cutler and Swinney (1987), report that the ability to process focus on target words in response to questions develops between 4;0 and 6;0 years. However, Cruttenden (1997) suggests a child can vary nuclear (i.e. tonic) placement when he has developed two word sentences and by the time he has three or four word sentences he can vary the nucleus to indicate old information. Cruttenden also points out that children of ten years can have difficulty with intonational meaning generally, and Wells et al. also suggest that the understanding of focus to highlight a key element lags behind children's ability to use it in their own speech.

3.2.3 Procedure

The test stimuli were saved individually as wav files presented using custom software on a Dell Latitude C640 laptop computer. In the Focus 2 and Focus 3 tests the initial "*No*" was not always produced by the talkers in the recordings so it was removed from all the phrases and sentences selected for the perception test. Implanted children (CI group) were tested individually in purpose-built audiology booths and the normal hearing (NH group) were tested in quiet conditions at home as described in Experiment I in Chapter Two. Laptop and speaker volume controls were set to produce a sound level that peaked at 70 -75 dB SPL and the speaker was placed one metre from the child's ear or microphone.

Before each sub-test the children were familiarised with the vocabulary, pictures, and voices, and they were allowed to practice in a trial run while the task was explained by the investigator. The stimuli were presented randomly to each child on a laptop computer as described above and there was no time limit. Response alternatives were represented by two or three picture alternatives (see Table 3.1 and examples of pictures in Appendix 3.1). In the Phrase test pairs of corresponding pictures (e.g. *bluebell* and *blue bell)* appeared for each stimulus and the subject was required to click on the appropriate picture. In the Focus 2 test two pictures (e.g. *BLUE* and *BOOK*) appeared for each stimulus, and in the Focus 3 test three pictures (e.g. *BOY, PAINTing, BOAT*) appeared with each stimulus. Subjects were asked to decide which word in the stimulus sounded the most important and then click on the appropriate picture. Once the test started the subject was allowed one repetition of each stimulus before responding. Each child worked independently at his/her own pace without prompting, using a mouse to select a picture to match each stimulus.

3.3 Results

The results of the tests in Experiment II are presented for the Phrase, Focus 2 and Focus 3 tests for the CI and NH below. A Pearson correlation test was carried out for age at test, duration of CI use and pulse rate in the speech processing strategies, and a significance level with Bonferroni correction $p<0.05$ (1-tailed) was applied. In addition to the individual test outcomes, an overall focus perception measure (MFocus) was introduced, this being the average of the Focus 2 and Focus 3 scores. Assuming that performance in Focus 2 and Focus 3 was the result of the same set of acoustic cues, this overall measure could be expected to be more reliable than the individual Focus 2 and Focus 3 scores. Similarly, an overall measure of F_0 discrimination threshold (MF_0) was computed, this being the average of the low and high range F_0 thresholds.

3.3.1 Overall CI and NH performance

Figure 3.1 shows variability for both groups in the spread of individual scores in the boxplots for each sub-test. In the Phrase test group scores ranged from 48% to 90% for the CI group and there was greater variability for the NH group with scores ranging from 47% to 96%. Assuming a binomial distribution (48 items, chance level 0.5) individual subjects would need to get 62.5% correct if we are to be 95% confident that they were not responding randomly. In both groups there were some individuals performing significantly above chance at 62.5% and some performing below this level in both groups (i.e. 10 CI subjects and 5 NH subjects),

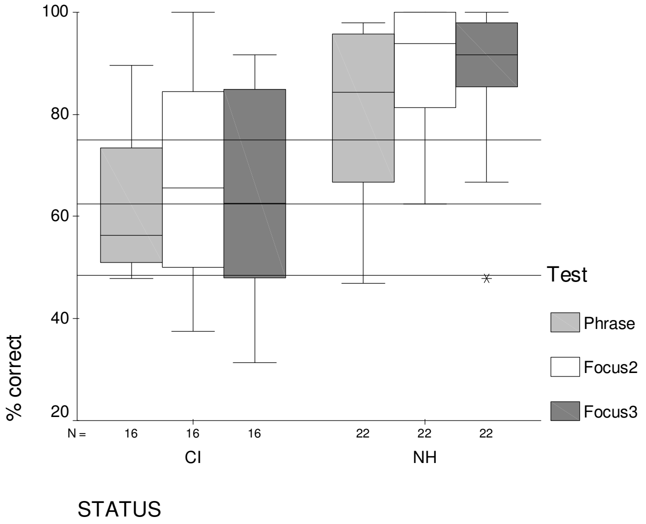


Figure 3.1 *Percentage correct scores (%) for NH and CI subjects in the Phase, Focus 2 and Focus 3 tests in Experiment II. Reference lines for each test at 62.5% (Phrase), 75% (Focus 2) and 48.5% (Focus 3) indicate where we can be 95% confident that subjects were not responding randomly to the stimuli.*

In the Focus 2 test all the NH subjects scored above 63% with some at or close to ceiling at 100%, and the CI group had some lower scores ranging from 38% to 100%. Assuming a binomial distribution for this test (16 items, chance level 0.5) subjects would need to get 75% correct if we are to be 95% confident they were not responding randomly. Ten individual subjects in the CI group performed below the 75% level whereas all except five of the NH group were above this level.

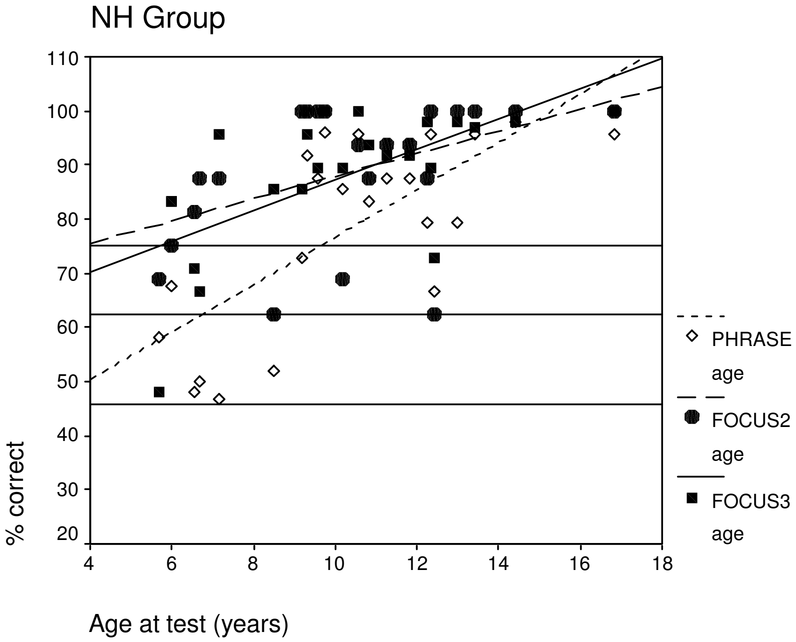
In Focus 3 test, scores for the NH subjects ranged from 65% up to ceiling at 100% with one exception at 47%. There was more variability across CI individuals for Focus 3 ranging from 31% to 93%. Assuming a binomial distribution (48 items, chance level 0.33) subjects would need to get 45.8% correct in this test if we are to be 95% confident they were not responding randomly. All of the NH subjects performed above 45.8% whereas four individual CI subjects were below this level. Overall, these results would suggest that in all three tests more individual subjects in the CI group were responding more randomly than the NH subjects.

3.3.2 Age at test

NH subjects

As discussed in section 1.3, there seems to be a consensus in the literature supporting the gradual acquisition of stress and intonation contrasts for normal hearing children up to and beyond 12;0 years. Figure 3.2 shows that by 8;6 years most of the NH group in the current investigation scored above 80% in all three tests. There was individual variation with some scores at or just above 60% for individual subjects even at 12;6 years, although scores for the Phrase and Focus 3 tests were significantly above chance levels (62.5% and 45.8% respectively). By 13.6 years, all test scores for the NH group were at or close to 100%.

A Pearson correlation test (see Table 3.2) shows that the relationship between age and percentage correct scores is statistically significant for the Phrase test (p= 0.001) and for the Focus 2 and Focus 3 tests averaged together (MFocus: p= 0.002). When Focus 2 and Focus 3 are analysed separately the correlation with age is significant for Focus 3 but only approaching significance with Bonferroni correction (p=0.017) for Focus 2.



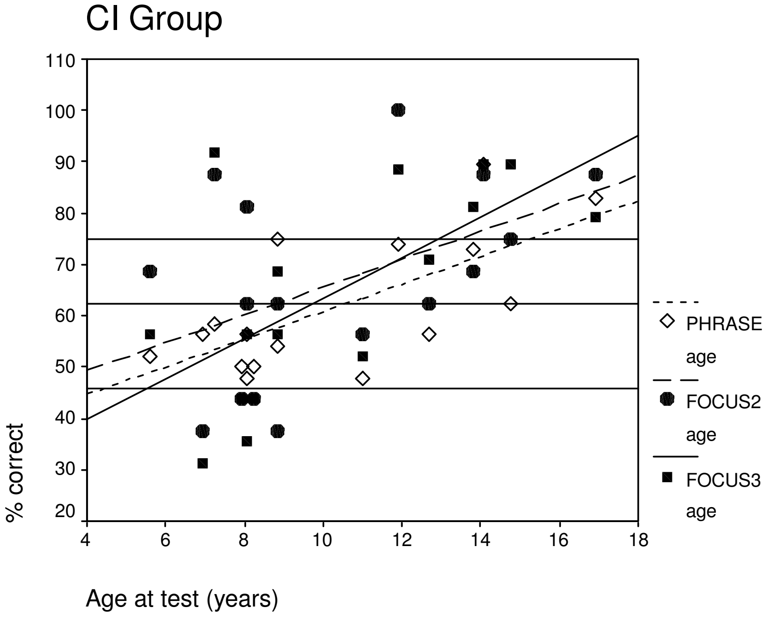


Figure 3.2 *Individual percentage correct scores for Phrase, Focus 2 and Focus 3 tests vs. age at time of testing for the NH group at the top of the figure and the CI group at the bottom. Reference lines at 62.5% (Phrase), 75% (Focus 2) and 45.8% (Focus 3) indicate where we can be 95% confident that subjects were not responding randomly to the stimuli in the three tests.*

NH	**Age at Experiment II**	
PHRASE	Pearson Correlation	0.721
	Sig. (1-tailed)	**0.001**
	N	22
FOCUS3	Pearson Correlation	0.621
	Sig. (1-tailed)	**0.001**
	N	22
FOCUS2	Pearson Correlation	0.454
	Sig. (1-tailed)	0.017
	N	22
BOLD type indicates correlations significant at p=0.0112 using Bonferroni corrected significance level		

NH subjects	**Age at Experiment II**	
PHRASE	Pearson Correlation	0.721
	Sig. (1-tailed)	**0.001**
	N	22
MFOCUS	Pearson Correlation	0.599
	Sig. (1-tailed)	**0.002**
	N	22
Bold type indicates correlation significant at p=0.025 Bonferroni corrected significance level		

Table 3.2 *Pearson correlations for age at test and percentage correct scores for Phrase test, Focus 2 and Focus 3 tests for the NH group in Experiment II. In the bottom table Focus 2 and Focus 3 tests have been averaged together (MFocus).*

CI subjects

Figure 3.2 shows that there was a gradual improvement in performance for the CI group across the age range up to 16;11 years but they were more delayed than the NH group. After age 12;6 the NH subjects scores were at or close to 100% in all three sub-tests whereas the majority of the CI subjects were significantly better than chance and in general did not obtain perfect scores beyond this age. A Pearson correlation test in Table 3.3 shows that there was a correlation between age and performance in the Phrase test (0.002) and a correlation was approaching significance with Bonferroni correction (p = 0.008) between age and performance when Focus 2 and Focus 3 tests were averaged together (MFocus). When these tests were analysed separately the correlation was significant with Bonferroni correction for Focus 3 only (p = 0.004). Similarly, there was a correlation between age at switch-on and MFocus (p = 0.005) and when Focus 2 and Focus 3 were analysed separately the correlation was significant for Focus 3 only (p = 0.002). These results suggest that although the correlations were not significant for all the tests, performance seems to improve with age for both CI and NH groups as indicated in the scattergraphs in Figure 3.2.

CI Subjects		**Duration of implant use**	**Age at switch-on**	**Age at Experiment II**	**Stimulation rate**
PHRASE	Pearson Correlation	-0.172	0.594	0.681	0.086
	Sig. (1-tailed)	0.261	0.008	**0.002**	0.375
	N	16	16	16	16
FOCUS3	Pearson Correlation	-0.421	0.671	0.642	0.125
	Sig. (1-tailed)	0.052	**0.002**	**0.004**	0.323
	N	16	16	16	16
FOCUS2	Pearson Correlation	-0.324	0.494	0.466	0.337
	Sig. (1-tailed)	0.110	0.026	0.034	0.101
	N	16	16	16	16
Bold type indicates correlation significant at p = 0.0042 Bonferroni corrected significance level					

CI Subjects		**Duration of implant use**	**Age at switch-on**	**Age at Experiment II**	**Stimulation rate**
PHRASE	Pearson Correlation	-0.172	0.594	0.681	0.086
	Sig. (1-tailed)	0.261	**0.008**	**0.002**	0.375
	N	16	16	16	16
MFOCUS	Pearson Correlation	-0.396	0.619	0.589	0.241
	Sig. (1-tailed)	0.065	**0.005**	**0.008**	0.184
	N	16	16	16	16
Bold type indicates correlation significant at p=0.0062 Bonferroni corrected sigificance level					

Table 3.3 *Pearson correlations for the CI group in Experiment II are presented above for age at test, duration of CI use, and pulse rate for each speech processing strategy. In the bottom table Focus 2 and Focus 3 tests are averaged together (MFocus).*

3.3.3 Duration of CI use

Performance in the three sub-tests in the present study varied and there is no evidence of children with longer implant experience performing any better than children with less experience. Figure 3.3 shows the variability in individual scores for each test, and in a Pearson correlation test in Table 3.3 there was no evidence of a correlation between duration of implant use and percentage correct scores in Phrase, Focus 2, or Focus 3 tests.

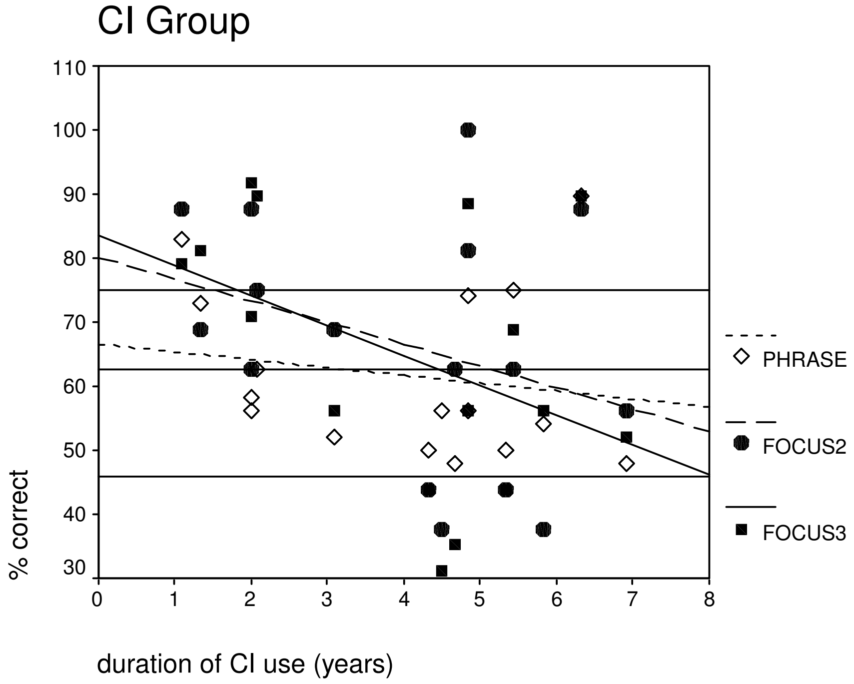


Figure 3.3 *Percentage correct scores (%) for individual CI subjects in the Phrase, Focus 2 and Focus 3 tests and duration of implant use (years). Reference lines at 62.5% (Phrase), 75% (Focus 2), and 45.8% (Focus 3) indicate where we can be 95% confident that subjects were not responding randomly to the stimuli in the three tests.*

3.3.4 Speech processing strategy

Figure 3.4 shows performances of CI children using ACE (stimulation/pulse rate 600-1800 pps) or SPEAK (stimulation/pulse rate 250 pps) speech processing strategies. In the Phrase Test some SPEAK users performed significantly above chance (62.5%) whereas most ACE users performed below this level. In the Focus 2 test, some individual ACE and SPEAK users performed significantly above the 75% chance level and others performed below this level. In the Focus 3 test, most ACE and

SPEAK users performed significantly above chance level (45.8%), although there were also some individual scores below this level. Table 3.3 shows there was no evidence of a correlation between stimulation/pulse rate and percentage correct scores for the Phrase test, Focus 2 test, or for Focus 3 test.

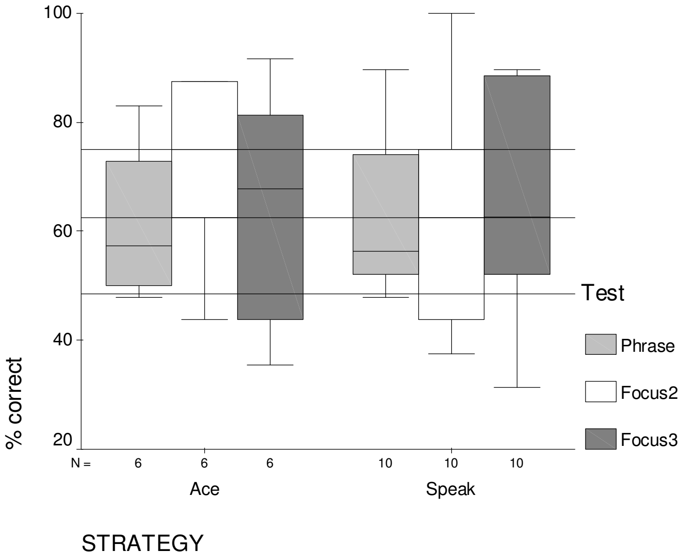


Figure 3.4 *Percentage correct scores (%) in the Phrase, Focus 2 and Focus 3 tests for the CI subjects using ACE and SPEAK speech processing strategies. Reference lines at 62.5% (Phrase), 75% (Focus 2) and 45.8% (Focus 3) indicate where we can be 95% confident that subjects were not responding randomly to the stimuli in the three tests.*

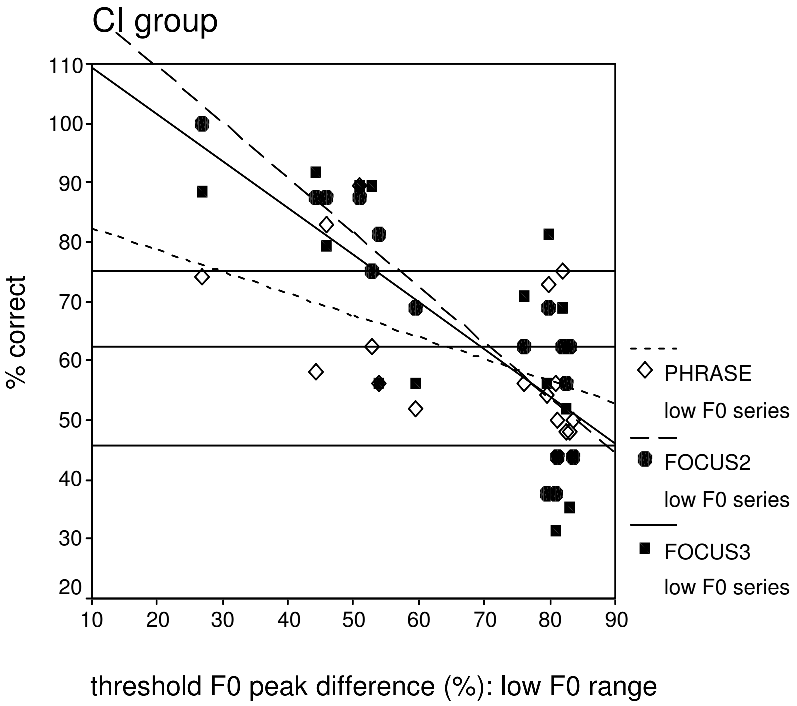
3.4 Experiment I and Experiment II results for the CI group

One of the questions to be addressed in Experiment II (Section 2.4.5) is whether ability to hear differences in compound vs. phrase stress and focus in natural speech stimuli is correlated with ability to hear smaller F_0 and/or duration and amplitude differences. To determine this a Pearson correlation test (Table 3.4) was carried out for F_0, duration and amplitude thresholds in Experiment I and percentage correct scores in the Phrase, Focus 2 and Focus 3 tests in Experiment II. A significance level of $p<0.05$ was applied with Bonferroni correction and individual results are presented below.

3.4.1 Correlations between F_0 discrimination (Experiment I) and Phrase, Focus 2 and Focus 3 scores (Experiment II)

Table 3.4 shows that an average of high and low F_0 thresholds (MF_0) correlated significantly with an average of Focus 2 and Focus 3 scores (MFocus) and the negative correlations with Bonferroni correction remained ($p = 0.001$) when age was controlled in Table 3.5. Correlations were also found when high and low F_0 thresholds and Focus 2 and Focus 3 were analysed separately (Table 3.4) and the correlations remained significant with Bonferroni correction ($p = 0.001$) when age was partialled out in Table 3.5. Results indicate the ability to hear linguistic focus correlated with ability to hear smaller F_0 differences whereas no correlations were found between F_0 thresholds and performance in the Phrase test.

In the scattergraphs in Figure 3.5, F_0 thresholds are presented for the low and high F_0 ranges in Experiment I with percentage scores in all three tests in Experiment II. Some talkers who were significantly above chance levels in Phrase and Focus 3 tests could only hear peak F_0 differences in the low F_0 range at the maximum difference level (see reference lines in the scattergraph in Figure 3.5 showing significance levels at 62.5%, 75% and 45.8% for Phrase, Focus 2 and Focus 3 tests respectively). This would suggest that these talkers were responding either to duration or amplitude cues. In the high F_0 range some of the CI subjects who were significantly above chance in the three Experiment II tests had better F_0 discrimination, except for one or two subjects significantly greater than chance in the Phrase and Focus 3 tests who were only hearing F_0 differences close to the maximum level.



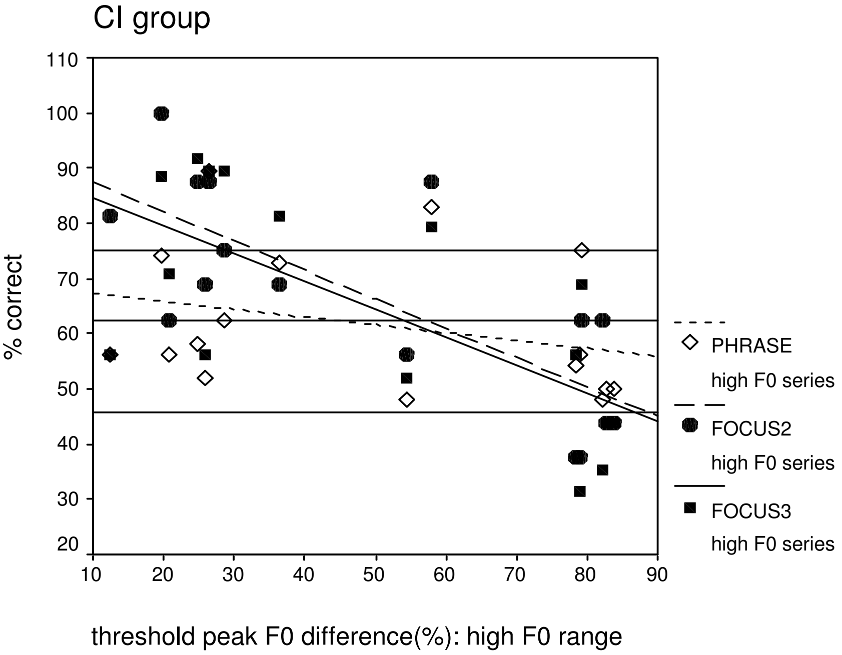


Figure 3.5 *F_0 thresholds in Experiment I and Phrase, Focus 2 and Focus 3 scores in Experiment II for the CI group in the low F_0 range at the top of the figure and in the high F_0 range on the bottom. Reference lines at 62.5% (phase), 75% (focus 2) and 45.8% (focus 3) for the three tests respectively indicate where we can be 95% confident that subjects were not responding randomly to the stimuli in the three tests.*

CI subjects		**Experiment I vs. Experiment II**		
		PHRASE	**FOCUS 3**	**FOCUS 2**
Low F_0	Pearson Correlation	-0.513	-0.711	-0.880
	Sig. (1-tailed)	0.021	**0.001**	**0.001**
	N	16	16	16
High F_0	Pearson Correlation	-0.297	-0.681	-0.756
	Sig. (1-tailed)	0.132	**0.002**	**0.001**
	N	16	16	16
Duration	Pearson Correlation	-0.392	-0.644	-0.878
	Sig. (1-tailed)	0.067	**0.004**	**0.001**
	N	16	16	16
Amplitude	Pearson Correlation	-0.467	-0.597	-0.523
	Sig. (1-tailed)	0.034	0.007	0.019
	N	16	16	16
Bold type indicates correlation significant at p=0.0042 Bonferroni corrected significance level				

CI subjects		**PHRASE**	**MFOCUS**
MF_0	Pearson Correlation	-0.414	-0.854
	Sig. (1-tailed)	0.055	**0.001**
	N	16	16
Duration	Pearson Correlation	-0.392	-0.802
	Sig. (1-tailed)	0.067	**0.001**
	N	16	16
Amplitude	Pearson Correlation	-0.467	-0.594
	Sig. (1-tailed)	0.034	**0.008**
	N	16	16
Bold type indicates correlation significant at p=0.0083 Bonferroni correct significance level			

Table 3.4 *Pearson correlations between F_0, duration and amplitude thresholds in Experiment I vs. percentage correct scores for Phrase, Focus 2 and Focus 3 tests in Experiment II for the CI subjects. In the bottom table Focus 2 and Focus 3 tests are averaged together (MFocus) and the high and low F_0 ranges (MF_0) are also averaged together.*

CI subjects		**Experiment II**		
		PHRASE	**FOCUS 3**	**FOCUS 2**
Low F_0	Coefficient	-0.407	-0.681	-0.870
	df	13	13	13
	P (1_tailed)	P= .066	**P= .003**	**P= .001**
High F_0	Coefficient	-0.110	-0.646	-0.721
	df	-13.000	-13.000	
	P (1_tailed)	P= .348	**P= .005**	**P= .001**
Bold type indicates correlations significant at p=0.0083 Bonferroni corrected significance level				

CI Subjects		**PHRASE**	**MFOCUS**
MF_0	Coefficient	-0.249	-0.853
	df	13	13
	P (1_tailed)	P= .185	**P= .001**
Bold type indicates correlation significant at p=0.025 Bonferroni corrected significance level			

Table 3.5 *Partial correlations controlling for age for the CI subjects between* F_0 *thresholds in the low and high* F_0 *ranges in Experiment I and percentage correct scores in Phrase, Focus 2 and Focus 3 tests in Experiment II. In the bottom table the high and low* F_0 *ranges have been averaged (*MF_0*) and also Focus 2 and Focus 3 tests have been averaged (MFocus).*

3.4.2 Correlations between duration discrimination (Experiment I) and Phrase, Focus 2 and Focus 3 scores (Experiment II)

When Focus 2 and Focus 3 scores were averaged together (MFocus) the correlation with duration thresholds was significant with Bonferroni correction (see Table 3.4) and the correlation remained ($p = 0.001$) when the focus tests were analysed separately. When age was partialled out (see Table 3.6 below) the correlation between Focus 2 and Focus 3 averaged together (MFocus) and duration thresholds was significant with Bonferroni correction. However, the correlation disappeared for Focus 3 ($p = 0.024$) when these two tests and duration thresholds were analysed separately indicating that any association is likely to be due to age. Table 3.3 also indicates a developmental effect where a correlation between age and Focus 3 scores was significant with Bonferroni correction ($p = 0.004$). The correlation between duration thresholds and Focus 2 tests remained significant when age was controlled which suggests that performance in this test depended on ability to hear differences in duration. No correlations were found between duration thresholds and the Phrase test.

The scattergraph in Figure 3.6 shows duration thresholds in Experiment I and all three test scores in Experiment II in the low F_0 range only. Most of the subjects whose performance was significantly greater than chance in all three tests could hear duration differences less than 60%, although there were some who were only able to hear bigger duration differences (e.g. 110% for one talker in Focus 3). These results suggest duration might be a more reliable cue than F_0 for some subjects.

CI subjects		**PHRASE**	**FOCUS3**	**FOCUS2**
Duration	Coefficient	-0.137	-0.518	-0.844
	df	13	13	13
	P (1-tailed)	P= .313	P= .024	**P= .001**
Amplitude	Coefficient	-0.252	-0.451	-0.389
	df	13	13	13
	P (1-tailed)	P= .182	P= .046	P= .076
Bold type indicates correlations significant at p=0.0083 at Bonferroni corrected significance level				

CI subjects		**PHRASE**	**MFOCUS**
Duration	Coefficient	-0.137	-0.743
	df	13	13
	P (1-tailed)	P= .313	**P= .001**
Amplitude	Coefficient	-0.252	-0.454
	df	13	13
	P (1-tailed)	P= .182	P= .045
Bold type indicates correlation significant at p=0.0125 Bonferroni corrected significance level			

Table 3.6 *Partial correlations for the CI subjects controlling for age between duration and amplitude thresholds in the low F_0 range in Experiment I and percentage scores in Phrase, Focus 2 and Focus 3 tests in Experiment II. In the bottom table Focus 2 and Focus 3 have been averaged together (MFocus).*

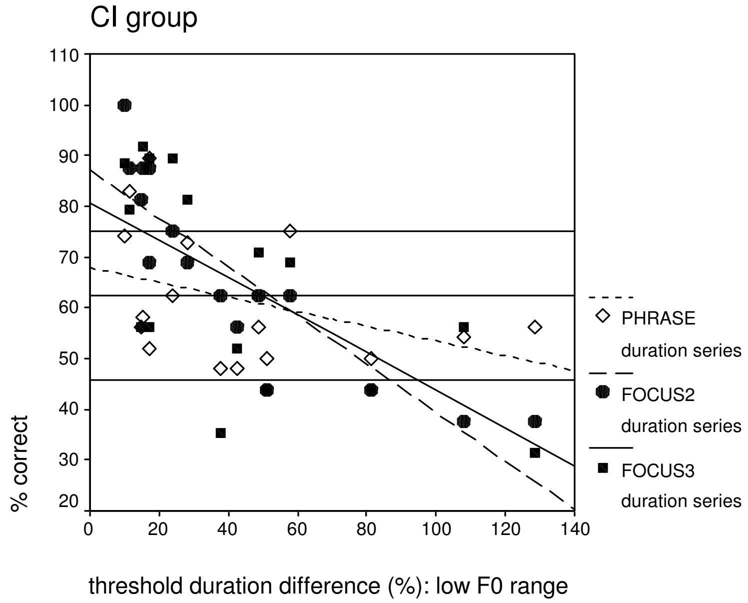


Figure 3.6 *Duration thresholds in Experiment I and Phrase, Focus 2 and Focus 3 test scores in Experiment II for the CI subjects in the low F_0 range only. Reference lines at 62.5% (Phrase), 75% (Focus 2) and 45.8% (Focus 3) indicate where we can be 95% confident that subjects were not responding randomly to the stimuli in the three tests.*

3.4.3. Correlations between amplitude discrimination (Experiment I) and Phrase, Focus 2 and Focus 3 scores (Experiment II)

Amplitude thresholds correlated with Focus 2 and Focus 3 scores ($p = 0.008$) in Table 3.4 when they were averaged together (MFocus) but when analysed separately the correlation with performance in Focus 3 only with Bonferroni correction was approaching significance ($p = 0.007$). When age was partialled out the correlation disappeared indicating a developmental effect (see Table 3.6).

The scattergraph in Figure 3.7 shows that amplitude difference thresholds in the low F_0 range varied for individual CI subjects who were performing significantly greater than chance in all three Experiment II tests and some of them could only hear amplitude differences greater than 9 dB. However, the variability in results suggests that some subjects might be able to make use of amplitude cues in the perception of compound vs. phrase stress and focus.

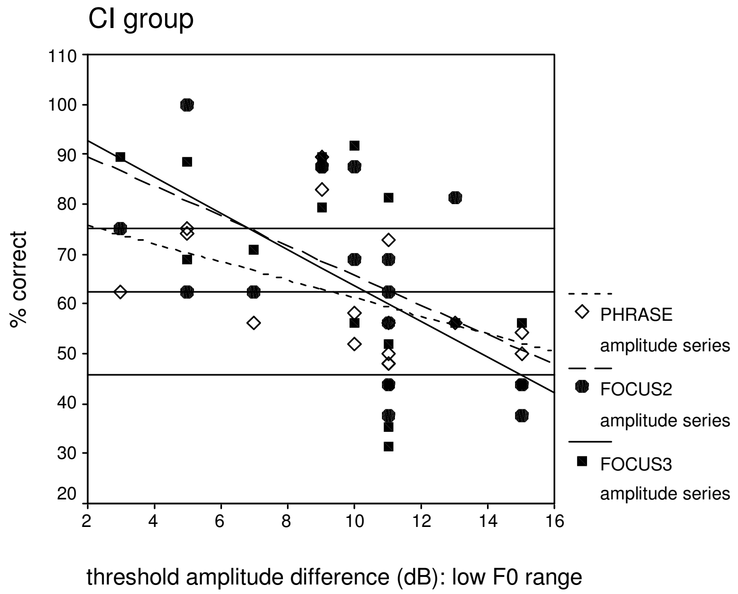


Figure 3.7 *Amplitude difference thresholds in Experiment I and Phrase, Focus 2 and Focus 3 test scores in Experiment II for the CI subjects in the low F_0 range only. Reference lines at 62.5% Phrase), 75% (Focus 2) and 45.8% (Focus 3) respectively indicate where we can be 95% confident that subjects were not responding randomly to the stimuli in the three tests.*

3.4.4 Summary

In summary when age was controlled negative correlations remained between F_0 thresholds in the high and low F_0 range and performance in Focus 2 and Focus 3. These results indicate that ability to hear linguistic focus is linked with ability to hear smaller F_0 differences. However, individual results as shown in Figure 3.5 indicate that some subjects who performed significantly greater than chance in the linguistic tests could only hear F_0 differences greater than the maximum difference (84%) which means they must be relying on other cues such as duration or amplitude. However, when age was partialled out a correlation between duration thresholds and Focus 3 scores disappeared but a correlation remained for Focus 2 which suggests that performance in Focus 2 depended on ability to hear smaller duration differences. However, individual results for all three tests and duration thresholds in the scattergraph in Figure 3.6 show that most subjects could hear duration differences of 60% or less so duration must have been a more reliable cue than F_0 for some subjects. A weak correlation between amplitude thresholds and Focus 3 test disappeared when age was controlled but variability in individual results as seen in Figure 3.7 indicates that some individual subjects may use amplitude as a cue to stress and intonation

.

3.5 Discussion and conclusions

3.5.1 Overall performance in Experiment II by CI group

The results of the perception tests involving natural speech stimuli in Experiments II in the Phrase (48% - 90%), Focus 2 (38% -100%) and Focus 3 (31% - 93%) tests above show variability across CI subjects with some individuals performing at or just below chance, and others obtaining scores above 90%. In all three tests (see Figure 3.1 and Figure 3.2) there were individual CI subjects who performed significantly above chance levels at 62.5% (6), 75% (6) and 45.8% (12) in Phrase, Focus 2 and Focus 3 tests respectively. These results indicate that some CI subjects seem to have acquired these contrasts despite the fact that in the low F_0 range in Experiment I (see Figures 2.3 and 2.4) most subjects were only able to hear F_0 differences greater than 0.5 octave and some subjects were unable to reliably hear the maximum difference of 84%. In the high F_0 range there were eight CI subjects who could hear smaller F_0 differences which were less than 0.5 octaves (see Figure 2.3), and this issue is discussed in more detail below.

3.5.1.1 Focus 2 vs. Focus 3 tests

As discussed in section 3.2.2 the difference between these two tests was not just the number of focus items and reduced memory load in the two element phrase. The Focus 2 task resembled the /baba/ test in Experiment I where listeners had to choose whether stress was on the first or second position. However, in Experiment I the acoustic parameters (F_0, duration and amplitude) were controlled in non-meaningful pairs of /baba/ syllables whereas Focus 2 stimuli (and also Focus 3 stimuli) were meaningful, the acoustic parameters were not controlled, and linguistic factors such as boundary markers and turn delimitation came into play on the final focus item. Focus 3 had more target focus items in pre-final position, with stressed and unstressed syllables in a longer sentence which had a gradual decline in F_0. Focus 2 and Focus 3 tests involved different sentence types i.e. adjective + noun vs. subject + verb + object) but despite these differences there was a similar range of scores overall for the CI subjects for both tests with not much difference between the medians (i.e. see boxplots in Figure 3.1 with the median score 62.5% and 65% for Focus 2 and Focus 3 respectively).

However, closer analysis shows that there were some differences in the results of these subtests. Focus 2 was less sensitive as a measure of perception ability as it involved fewer focus items to choose from and the number of items presented was lower. The chance level (1 in 2) was 50% and assuming a binomial distribution, with 16 trials, listeners would need a score of 75% to be significantly above chance. In Focus 3 there were three items to choose from so the chance level was 33.3% and listeners needed a score of 45.8% to be significantly above chance. This means that the median score was below chance for the Focus 2 test with only 6 of the 16 CI subjects scoring significantly above chance level whereas the median score was significantly above chance for the Focus 3 test with 12 CI subjects significantly above chance level. Further analysis of the median scores suggest that final focus position seems to have been a bit more difficult than the pre-final focus position in the Focus 2 test with poorer performance in final position (63%) than in pre-final position (75%).

In the absence of pitch cues for the CI subjects, boundary markers at the end of a phrase such as final lengthening or a drop in amplitude in some non-focus words might have obscured increased lengthening of pre-final focus words. As Experiment I results show us, pitch differences associated with such final lowering would not be accessible to most implant users unless they were greater than 0.5 octaves (6 semitones). As a result these listeners would more dependent on duration and amplitude cues which may have been insufficient to signal final focus to CI listeners in Focus 2 stimuli. It is also possible that competing prosodic functions in the final focus item (i.e. boundary markers vs. final focus) might be more challenging for implanted children in adjacent target syllables such as ***BLACK*** book vs. black ***BOOK*** or green ***DOOR*** vs. ***GREEN*** door. By comparison, inspection of median scores for the different focus positions in Focus 3 (i.e. 72%, 59%, and 66% for initial, medial and final position respectively) shows the lowest score for medial focus.

The three element SVO sentences (subject+ verb+ object) differed from Focus 2 as they had unstressed syllables occurring between three target word/syllables so they were not immediately adjacent to each other e.g. the ***BOY*** is painting the boat vs. the boy is ***PAINT*ing** the boat vs. the boy is painting the ***BOAT***. For normal hearing listeners boosting of F_0 in the target word/syllables might stand out especially in medial or final position because of a step up or pitch reset against the natural decline of F_0. However, as indicated by Experiment I results most CI listeners would have difficulty hearing F_0 changes of less than 0.5 octaves and would have to rely more on duration and amplitude cues. The boxplots in Appendix 3.3 show that the F_0 differences between medial focus words and neighbouring words (***PAINT*** *vs. boat*, ***BAKE*** *vs. cake,* and ***EAT*** vs. bone, and ***DRIVE*** *vs. car*) are greater than for other focus positions. Since these median F_0 differences were generally less than 0.5 octaves they would not be accessible to most implanted listeners as indicated by Experiment I F_0 thresholds. There were generally small F_0 differences between the final focus items and previous words (*paint vs.* ***BOAT****, bake vs.* ***CAKE****, drive vs.* ***CAR****, eat vs.* ***BONE***) but as indicated in the boxplots in Appendix 3.6, increases in the median duration for target words in two sentences (i.e. *the dog is eating a bone* and *the man is driving a car*) and a step up in the median amplitude in all four sentences as shown in the boxplots in Appendix 3.8 may have helped convey final focus to some implanted listeners. See section 3.5.4 for more detailed discussion of measurements of the Focus 3 stimuli.

3.5.1.2 Phrase Test

As mentioned in section 1.11.2 differences between compound and phrase stress may not be signalled in the same way by different adult speakers and pitch reset may not be as reliable as lengthening and pause (Peppé et al., 2000). If this is the case these contrasts should be accessible to cochlear implant listeners who because of device limitations have to rely on duration or amplitude cues. Figure 3.2 shows that scores varied from 48% to 90% with 6 CI subjects significantly above chance (62.5%) and 10 below. Closer analysis of the total scores for the CI group shows a preference for phrase (median = 73%) rather than compounds (median = 56%) but the total median score for the CI group as indicated in Figure 3.1 was 56% which was still just above chance level. However, as discussed in section 1.11.1 for normal hearing children the ability to discriminate between compound vs. phrase stress does not seem to be developed until later in the acquisition process and can continue developing in some cases up to 12;0 years and beyond. The relationship between performance in Experiment II tests and age at time of testing is discussed below in section 3.5.3.1.

Since the acoustic parameters F_0, duration and amplitude in these stimuli were not controlled in Experiment II it is difficult to ascertain which cues CI listeners were responding to but given that most median F_0 differences in these phrase materials were less than 0.5 octaves (see Appendix 3.3) it is likely that duration and amplitude were more reliable cues for most CI subjects. The relationship between ability to hear smaller differences in F_0, duration and amplitude in Experiment I and perception of linguistic contrasts in Experiment II is also discussed in greater detail for CI subjects below in section 3.5.4.

3.5.2 Do Experiment II results for the CI subjects support findings reported in the literature?

As discussed in Chapter One there are no available reports for CI children on the perception of the prosodic contrasts under investigation in the present study and what we know to date about pitch discrimination difficulties by implanted children is drawn from studies of Chinese tones (see sections 1.8 and 1.11.3). Although methodology and stimuli differ from the present investigation results of these studies vary but in general they suggest that limited pitch information affects the ability to discriminate between lexical tones. For example, Ciocca et al. (2002) reported identification of meaningful Cantonese tones was poor overall with group performance significantly above chance for only three out of eight contrasts, where one of each pair of tones was a high tone. It was suggested that CI listeners might have been helped by high amplitude associated with high tones. Peng et al. (2004) also report that a group of Mandarin speaking children with implants were significantly above chance at Mandarin tone identification. They concluded however, that the shorter duration of one Mandarin tone (T4) may have provided an additional duration cue for these listeners. Experiment II results in the current study shows that although there was considerable individual variability in scores, performance was better than found by Ciocca et al. with more individual CI subjects scoring significantly greater than chance in the three subtests (i.e. 6 in the Phrase test, 12 in Focus 3, and 6 in Focus 2).

As mentioned earlier, overall performance in the current study for the Focus 2 and Focus 3 tests was similar but because of the smaller number of items in the Focus 2 test there was a higher score required to demonstrate a significant difference from chance. The better performance in the Focus 3 test compared to the Phrase test could be because the concept of focus is acquired earlier than phrase vs. compound stress. As discussed in section 1.11.1, Cutler and Swinney (1987) suggest that focus seems to be acquired by 5;0 year normal hearing children whereas the ability to discriminate between compound and phrase stress seems to be acquired later in the acquisition process i.e. up to and beyond 12;0 years (Atkinson-King, 1973; Vogel and Raimy, 2002; Wells et al. 2004; Doherty et al., 1999). The effect of age at time of testing on performance in Experiment II is discussed further in section 3.5.3.1 below.

Although different skills were being tested in Experiment I and Experiment II it is possible that CI subjects' ability to hear F_0, duration and amplitude differences in Experiment I might be directly linked with performance in the linguistic tasks in Experiment II. However, changes in these acoustic cues in the natural speech contrasts presented in Experiment II might not have not have been big enough to be accessible to some CI listeners, and this issue is discussed in greater detail in section 3.5.4. It remains to be seen whether performance in Experiment II (i.e. perception of intonation contrasts) is directly lined with the ability to hear F_0, duration and amplitude in Experiment I. Pearson correlation tests between the two test results may indicate whether F_0 is a necessary cue to lexical stress and focus in the current study as in hypothesis (i) or whether F_0 is not a necessary cue and that CI listeners can rely on other cues such as duration and amplitude as in hypothesis (ii).

3.5.3 Comparisons between NH and CI groups

Performance in Experiment II also varied across the NH subjects (see Figure 3.1 and Appendix 3.10) in the Phrase (47% - 96%), Focus 2 (63% - 100%), and Focus 3 (65% - 100%) tests. As already mentioned in section 3.5.1.1 there were only two focus items to choose from in the Focus 2 test so that the chance level was 50% and listeners would need a score of 75% to be significantly above chance in this test. This made it less sensitive than Focus 3 as a measure of perception ability. In the Focus 3 test there were three items to choose from so the chance level was 33.3% and listeners would need a score of 48.5% to be significantly above chance level. All of the NH subjects performed significantly above chance (45.8%) in the Focus 3 test, and most subjects i.e. 17 subjects in the Phrase test and 17 subjects in Focus 2 test performed significantly above chance (62.5% and 75% respectively). In contrast with this only 6 of the 16 CI subjects in Phrase and Focus 2 performed significantly better than chance whereas performance was better for Focus 3 with 12 CI subjects significantly greater than chance. Further, the median score of the CI children for Focus 3 was very close to that for Focus 2 (see fig 3.1) despite the lower chance level for Focus 3. As discussed in section 3.5.1.1 there were also syntactic differences between Focus 2 and Focus 3 stimuli which may account for difference in performance for CI listeners. In Focus 2 test competing prosodic functions (i.e. boundary markers and final focus) in two adjacent target words (e.g. *a GREEN door* vs. *a green DOOR*) may have been challenging for CI listeners. In contrast, Focus 3 test had three target words with unstressed syllables occurring between them. Since the target words were not adjacent to each other, the focus items in this test may have been more perceptually salient to CI listeners. In the boxplots in Figure 3.1 median scores for the NH subjects for the three tests (84%, 94% and 91.7% for Phrase, Focus 2 and Focus 3 respectively) were significantly above chance. Median scores for the CI subjects were 56%, 66% and 62.5% for Phrase, Focus 2 and Focus 3 respectively but only the Focus 3 median score (62.5%) was significantly greater than chance.

Overall, NH subjects seem to have used whatever cues were available to them in the perception of focus and compound vs. phrase stress in Experiment II, and although most were significantly above chance there was some individual variation. The median scores for the NH group in Focus 2 for pre-final and final focus items show better performance for the NH group (97% and 100% respectively) on the final focus word than for the CI group (75% and 63% respectively). One possible reason is that an additional acoustic cue i.e. a step up or more striking fall in F_0 on the final item may have been a stronger cue to focus for the NH listeners when combined with duration and/or amplitude cues. In Focus 3, however, the two groups differed and median scores (93.8%, 93.8% and 87.5% for initial, medial and final focus position) indicate that performance was slightly worse for final focus position for the NH group but worse in medial focus position for the CI group (72%, 59% and 66%).

According to Peppé et al. ambiguity is not uncommon even amongst adult speakers (see section 1.11.1), and when focus was not perceived on some target words it may have been because changes in F_0, duration or increased amplitude in these words were insufficient to convey focus to listeners. For the CI listeners it is possible that the step up in F_0 (and/or duration and amplitude adjustments) on the target focus word in medial position were not salient to these listeners, and for the NH group the changes in the acoustic cues may have been less salient for the NH listeners in final position. The accessibility of the acoustic cues for the CI listeners in Focus 3 stimuli are discussed in greater detail in section 3.5.4.

3.5.3.1 Did scores in Experiment II improve with age for NH and CI subjects?

By 13;6 years, all test scores for the NH group were at or close to 100% (see Figure 3.2) whereas for the CI group test scores were all significantly above chance by 14;6 years but they are delayed compared to the NH group. The NH group improved rapidly between 6;0 and 10;0 years and thereafter obtain scores of almost 100%. The CI group on the other hand showed a more gradual improvement with age but in general did not achieve perfect scores even beyond 12;0 years. However, since only the age range matched for the two groups it is difficult to draw comparisons between individual NH and CI subjects. Future experiments should include more age-matched subjects but the present results are useful as they give us some indication of whether there is a delay in the acquisition of the linguistic contrasts under investigation in Experiment II by CI within the same age range.

The gradual acquisition of compound vs. phrase stress by NH subjects up to and beyond 12;0 years in the present study supports previous studies of normal hearing children (Atkinson-King, 1973; Vogel and Raimy, 2002; Wells et al., 2004). By 6;6 years all except one of the NH subjects in the present study were significantly above chance in the Focus 3 test which is comparable to data from Cutler and Swinney (1987). However, some CI subjects were still below chance in the Focus 2 stimuli up to 12;0 years. Wells et al., who studied a much larger population of NH children, reported that some of their subjects did not reach ceiling scores in some of their sub-tests even by 13;0 years, and according to Cruttenden (1997) some aspects of intonation may not be acquired by 10;0 years. The age range in the current study is greater than previous studies of normal hearing children and Experiment II results suggest that the acquisition process continues up to 17;0 years and beyond for the CI group.

A Pearson correlation test for the NH group in Table 3.2 shows that the relationship between age and percentage scores was statistically significant for performance in the

Phrase test, and for Focus 2 and Focus 3 tests averaged together (MFocus). When the Focus 2 and Focus 3 tests were analysed separately the correlation with age was significant for Focus 3 and only approaching significance for Focus 2 test with Bonferroni correction (p= 0.017). For the CI group, performance seemed to be more delayed across the age range and most subjects did not reach ceiling. Table 3.3 also shows that the correlation between age and performance in the Phrase test was significant for the CI group, and when the Focus 2 and Focus 3 scores were averaged together (MFocus) the correlation with age at testing was approaching significance with Bonferroni correction (0.008). However when results were analysed separately the correlation was significant for Focus 3 only. The correlation between age at switch-on and both focus tests averaged together (MFocus) was significant but when these subtests were analysed separately at the top of Table 3.3 the correlation was significant for Focus 3 only. Although some correlations were non-significant there seems to be sufficient indication that performance improves with age in both the NH and CI groups. These results are in contrast with Ciocca et al. (2002) who report that correlations between Cantonese tone identification and age at implantation or age at the time of testing were not significant for CI children.

3.5.4 How accessible are acoustic cues (F_0, duration and amplitude) to the subjects in the stimuli in Experiment II?

Figure 2.4 shows that most of the NH subjects in Experiment I could hear F_0 differences less than 10% in the low F_0 range and 15% in the high F_0 ranges so they would have no difficulty hearing F_0 changes associated with target focus words. However, as discussed earlier cues to stress and intonation contrasts such as lexical stress and focus may vary for CI subjects according to difference thresholds for F_0, duration and amplitude. In the absence of F_0 or amplitude cues, listeners may rely on duration. Given the wide age range of the subjects, age effects should be expected in the speech tests and some younger subjects may perform poorly because of this. Correlation tests were carried out to establish whether performance in the linguistic tests in Experiment II depended on individual subjects' ability to hear smaller differences in F_0, duration and amplitude.

3.5.4.1 Does performance in Experiment II depend on how well CI subjects hear F_0 *differences in Experiment I?*

In Experiment I, most CI subjects were unable to hear peak F_0 differences less than 40% (almost 0.5 of an octave) between synthetic /baba/ bisyllables in the low F_0 range. Median F_0 thresholds for these subjects were 57% and 77% for the low and high F_0 range respectively (see Figures 2.3 and 2.4). Results suggest that in Experiment II many CI subjects might not hear F_0 differences between the target focus word and the neighbouring unfocussed words if they are less than 0.5 of an octave and others may not hear even when there is almost an octave difference as the F_0 thresholds as Experiment I results suggest. Detailed analyses of acoustic measurements of target words are available for Focus 3 stimuli only in the current investigation.

Measurements presented in Appendix 3.2 show that F_0 differences between target focus words and neighbouring words rarely exceeded 0.5 of an octave and would not have been accessible to most CI listeners (for exceptions see Talker 2 for ***MAN****: drive* 11.88 semit., and Talker 3 for ***EAT****: bone* 16.37 semit, and in an extreme case paint: *paint*: ***BOAT***: 26.04 semit. which were possibly errors in F_0 extraction and measurements in PRAAT and discussed in section 4.2.4.1). As discussed in section 3.5.1.1 earlier the boxplots in Appendix 3.3 show that the F_0 difference between focus words and neighbouring words were generally less than 0.5 octaves (i.e. 6 semitones) and so would be inaccessible to most CI subjects. Appendix 3.4 summarizing the range of median F_0 differences for individual NH talkers shows that the median values of the largest F_0 change over the target syllables in each sentence were less than or only slightly above 0.5 octaves (i.e. 4.04 semit., 4.53 semit., 3.78 semit., 6.36 semit.) for Talkers 1, 2, and 3 and 4 respectively which would not be accessible to most CI listeners. Although in the high F_0 range in Experiment I the median F_0 threshold was 77% for the CI group, there were seven CI subjects (i.e. subjects 1, 3, 8, 11, 12, 13, and 17 who could reliably hear peak F_0 differences between 10% and 30% (see Figure 2.3) and it is possible that these subjects might have been able to hear smaller F_0 differences (i.e. less than 0.5 octaves) between focussed and neighbouring unfocussed words in Experiment II. Appendix 3.9 for the CI group shows the distribution of scores for individual NH talkers for male Talkers 1 (57%) and 3 (69%)

and for female Talkers 2 (66%) and 4 (67%) indicate no advantage for female Talker 4 who also had a higher production range than other talkers. These results would also suggest generally that the ability to hear smaller F_0 difference in the high F_0 range was not necessarily an advantage for these CI listeners.

As discussed in section 3.4.1, Pearson correlation tests were carried out to investigate whether ability to hear smaller F_0 differences in Experiment I was statistically linked with the ability to hear differences of stress and focus in Experiment II. Table 3.4 shows that an average of high and low F_0 range thresholds (MF_0) significantly correlated with the average of Focus 2 and Focus 3 tests (MFocus) and the correlation remained when age was controlled. When the low and high F_0 ranges and focus tests were correlated separately there were negative correlations between F_0 discrimination in both F_0 ranges (Experiment I) and performance in both Focus 2 and Focus 3 tests (Experiment II). When age was partialled out significant correlations remained between Focus 2 and Focus 3 tests and F_0 discrimination in both F_0 ranges. It would appear that performance in these focus tests correlated with ability to hear smaller F_0 differences. No correlations were found between F_0 discrimination and scores in the Phrase test and as indicated in Table 3.3 performance in this test correlated with age at time of testing. However, individual scores plotted in the scattergraphs in Figure 3.5 indicate that some individual CI subjects who were unable to hear peak F_0 differences at or close to the maximum peak F_0 difference level (84%) performed significantly above chance in the Focus 3 test and in the Phrase test indicating that that these subjects do not necessarily rely on F_0 cues to stress. These individual scores support hypothesis (ii) which suggests that F_0 is not a necessary cue to lexical stress and focus for CI listeners.

3.5.4.2 Does performance in Experiment II depend on how well CI subjects hear duration differences in Experiment I?

Figure 2.6 shows us that NH listeners varied in their ability to hear duration differences (i.e. between 10% and 48%) in the unprocessed condition in Experiment I but the median score was 25%. The boxplots in Appendix 3.6 shows that the median durations of most of the target focus words in the boxplots for the NH stimuli were more than 50% longer than in the neighbouring unfocussed position and these differences should be accessible to most of the NH listeners in Experiment II. The scattergraph in Figure 3.6 shows that the CI subjects who were able to hear duration differences less than 30% in Experiment I scored significantly above chance in the three sub-tests in Experiment II (i.e. seven children in Focus 3, two children in Phrase and five children in Focus 2). Most of the CI subjects who scored significantly above chance in the three tests were able to hear duration differences less than 60%. Since the median duration threshold for the group in Experiment I was 35% (see Figure 2.6) it is possible that for some CI children, duration may provide a stronger cue to stress than F_0.

Duration measurements in Appendix 3.5 and the boxplots in Appendix 3.6 show that the median durations for the target focus/syllables in three of the four stimulus sentences (i.e. all excepting *the girl is baking*) were longer when target words were in focus than when they were not in focus e.g. ***BOY*** (75%), ***DOG*** 75%) ***BONE*** (140%) ***DRIVE*** (80%) ***CAR*** (140%). These duration differences would be accessible to CI listeners with a median duration threshold of 35% and also to individual CI listeners who could hear duration differences less than 60% in Experiment I. Smaller durations differences such as ***PAINT*** (20%) or ***BOAT*** (20%) might be accessible to the eight CI listeners who could hear duration differences of less than 30% in Experiment I.

The range of duration differences between the minimum and maximum durations for the target words in each sentence are presented for individual talkers in Appendix 3.4. The medians of the largest durational change over the target syllables were 164 ms (Talker 1), 127 ms (Talker 2), 136 ms (Talker 3), and 101 ms (Talker 4). Appendix 3.9 shows the distribution of scores obtained by the CI group for individual NH talkers (i.e. 57%, 66%, 69% and 67% for Talkers 1, 2, 3 and 4 respectively). Talkers 1 and 3 were male and Talkers 2 and 4 were female and although Talker 1 had the largest median difference between the minimum and maximum durations for the target words (i.e. 164 ms) CI listeners did not perform better for this talker.

Pearson Correlation tests were carried out for the CI subjects to establish whether there was any statistical relationship between performance in the three Experiment II subtests and ability to hear duration differences in Experiment I. When Focus 2 and Focus 3 tests (MFocus) were averaged together in Table 3.4, there was a significant correlation with the ability to hear smaller duration differences even when age was partialled out in Table 3.6. When analysed separately negative correlations were also found between duration thresholds and performance in Focus 2 and Focus 3 tests, but when age was partialled out the correlation disappeared for Focus 3 suggesting a developmental effect. This is borne out in Table 3.3 which shows that the correlation between Focus 3 and age at testing was significant with Bonferroni correction ($p = 0.004$). A significant correlation remained between Focus 2 scores and duration difference thresholds (Table 3.6) which suggests that performance in this test depended on ability to hear duration differences. A similar correlation remained (Table 3.5) when age was partialled out for Focus 2 (and also Focus 3) and F_0 thresholds as discussed above. So it would appear that CI subjects' performance in Focus 2 test was linked with the ability to hear F_0 and/or duration cues.

As discussed in Chapter One (see sections 1.11.2 and 1.4.2) pause and lengthening were reported to be more reliable cues to compound vs. phrase stress than pitch cues so it is surprising that there was no evidence of a correlation between ability to hear duration differences and performance in the Phrase test. For Focus 2 it seems that the ability to hear focus is linked with the ability to hear smaller F_0 and duration differences, and since the median threshold for the CI group in Figure 2.6 was 35% most durational increases in the target focus words in the stimuli listed above would be accessible to them. The scattergraph in Figure 3.6 shows most CI listeners who could hear duration difference less than 60% were significantly above chance in Experiment II. Most of these listeners could hear duration differences less than 30% which lends support to hypothesis (ii) i.e. that F_0 is not a necessary cue to stress and intonation contrasts in the present study for CI listeners and that duration might provide a more reliable cue.

3.5.4.3 Does performance in Experiment II depend on how well CI and NH subjects hear amplitude differences in Experiment I?

As shown in Figure 2.8 the NH subjects who participated in Experiment I varied in their ability to hear amplitude differences in the unprocessed condition (i.e. between 1 dB and 10 dB) and the median threshold was 5 dB. The boxplots for the stimuli produced by the NH talkers in Appendix 3.8 show that amplitude changes in the target focus words and neighbouring words ranged between <1 dB and 10 dB. Experiment I results suggest that it is possible that some of the smaller amplitude changes might not be accessible to the NH listeners who participated in Experiment

II. For the CI group amplitude thresholds in Experiment I ranged from 3 dB up to a maximum difference of 15 dB. The boxplots in Figure 2.8 show the median amplitude threshold for the group of CI listeners was 11 dB. The scattergraph in Figure 3.7 shows that even for CI children with large amplitude thresholds there was a wide range in performance in the Phrase, Focus 2 and Focus 3 tasks, so prosodic perception could not be entirely due to the use of amplitude cues. The scattergraphs also show that ability to hear amplitude differences varied for CI subjects who were significantly above chance in all tests but some were only able to hear amplitude differences greater than 9 dB.

The boxplots in Appendix 3.8 for the Experiment II stimuli produced by the four NH talkers show that the median amplitude differences for the target words in focus and neighbouring unfocussed positions for each of the stimulus sentences ranged between <1 and 5 dB for initial position, between 1 dB and 10 dB for medial position, and 4 dB and 9 dB for final focus position. It is possible that amplitude might provide a more accessible and reliable cue to focus than F_0 (see 2.4) for some CI listeners, but since the median amplitude threshold for the group of CI listeners was 11 dB, the amplitude differences in initial and final focus position might be less accessible to some CI listeners. Appendix 3.4 shows that for individual NH talkers the median of the largest changes in amplitude across the target syllables in the Experiment II stimuli were 9 dB, 8 dB, 8 dB and 9 dB for Talkers 1, 2, 3 and 4 respectively which was less than the median amplitude threshold (i.e. 11 dB) for the CI group. Talkers 1 and 4 had larger median changes in amplitude (9 dB) across target syllables than the other talkers, and as discussed in sections 3.5.4.2 and 3.5.4.3, Talker 1 had the largest median durational change (164 ms) and Talker 4 had the largest median F_0 change (6.48 semit.). However, CI listeners did not perform better for these talkers (see Appendix 3.9) in Experiment II, and this could be because the F_0 durational and amplitude changes might not have been accessible to some CI listeners.

To investigate whether ability to hear amplitude changes in Experiment I was statistically linked with performance in the Experiment II tests Pearson Correlation tests were carried out. When the focus tests were averaged together (MFocus in Table 3.4) the correlation with amplitude threshold disappeared when age was controlled. When the focus sub-tests were correlated individually no correlations were found between amplitude discrimination and Focus 2 or Phrase scores, but the correlation between Focus 3 and amplitude thresholds was approaching significance. However, when age was controlled this correlation disappeared suggesting some developmental effects. Although there was no evidence of a correlation between the ability to hear amplitude differences and performance in Experiment II tests, the variability in results suggests that some individual CI subjects might be able to use amplitude as a cue to lexical stress and focus. These results support hypothesis (ii) which suggests that F_0 is not a necessary cue to stress and intonation.

3.5.5 Effect of duration of implant use on CI performance in Experiment II

As mentioned earlier there was much more individual variation across the age spectrum for the CI group even up to 16;11 years but there was no evidence of a correlation between performance in Experiment II and duration of implant use. The results in previous studies vary. For example, Ciocca et al. (2002) found that correlations with post-operative use of CI were not significant in their study of Cantonese tones. In contrast with Ciocca and with the results of the present study, Peng et al. (2004) report that Mandarin tone identification scores for their subjects correlated with duration of implant use.

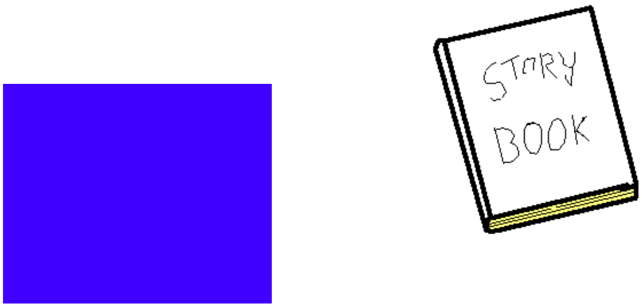
3.5.6 Effects of stimulation rate on CI performance in Experiment II

A Pearson Correlation test was carried out to establish whether performance was better for subjects using a faster stimulation rate. The CI children in the current investigation used Nucleus speech processors with either SPEAK (250 pps) or ACE (600-1800 pps) speech processing strategies but no correlations were found stimulation rate and performance in the Phrase or focus tests. There were some individual ACE and SPEAK users performing significantly above chance (75% and 45.8% respectively) in the Focus 2 and Focus 3 tests. In the Phrase test, however, some SPEAK users performed significantly above chance (62.5%) whereas most ACE users performed below this level. These results support some of the findings in the literature. For example, Barry et al. (2002a) found no significant difference between ACE and SPEAK users in the recognition of lexical tone and average performance was below chance for four tonal contrasts with SPEAK and below chance for seven contrasts with ACE (total number of contrasts was 15). Overall, it is reported that the SPEAK group performed better and the additional stimulation provided by ACE was not found to be an advantage. In a follow-up study by Barry et al. (2002b) considerable variation was found for ACE users and the higher stimulation rates seemed to provide more information about pitch direction (contour) than pitch height which is reported to play a crucial role in the identification of Chinese tones.

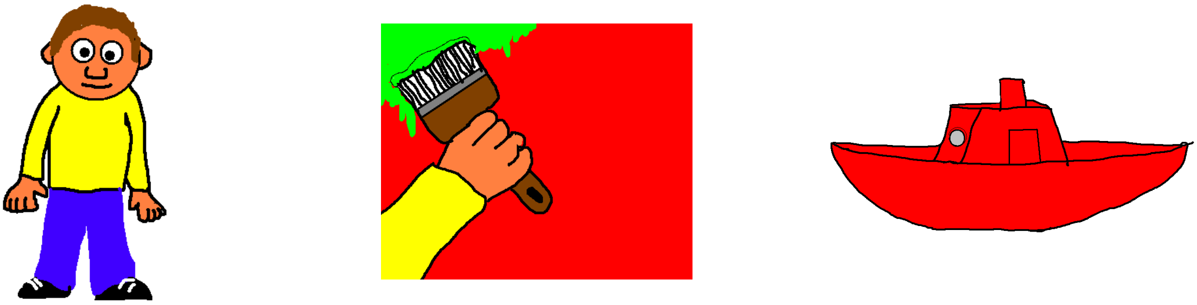
3.5.7 Concluding comments

Analysis of the acoustic cues used in the Focus 2 stimuli would also be useful for comparison with Focus 3 and will be investigated in the future. Data from additional NH and CI subjects at the different ages in the age range would be helpful for comparison with other normative studies. However, the results of the current study suggest that the gradual improvement in performance in Experiment II across the age range suggests that CI listeners must have stored representations of the prosodic contrasts but development of perceptual skills are delayed for these subjects compared to the NH subjects. As indicated in Table 3.3 performance in Focus 3 correlated with age at switch-on but there was no correlation between performance in the perception tests and duration of implant use or stimulation rate. It is possible that in addition to age there may be other influencing factors such as placement of electrodes or neural survival but they are beyond the scope of the present study**.** Variables such as age at testing, age at switch-on, duration of implant use and stimulation rate will be considered again in Chapter Four in the discussion of the acoustic measurements in the production of focus by the same group of CI subjects.

a.



b.



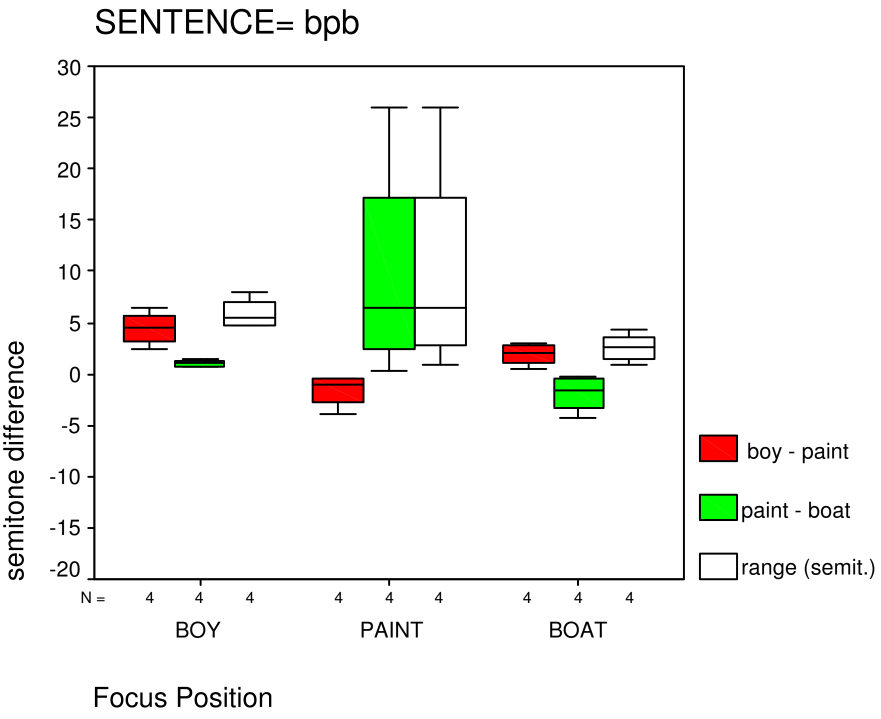
c.

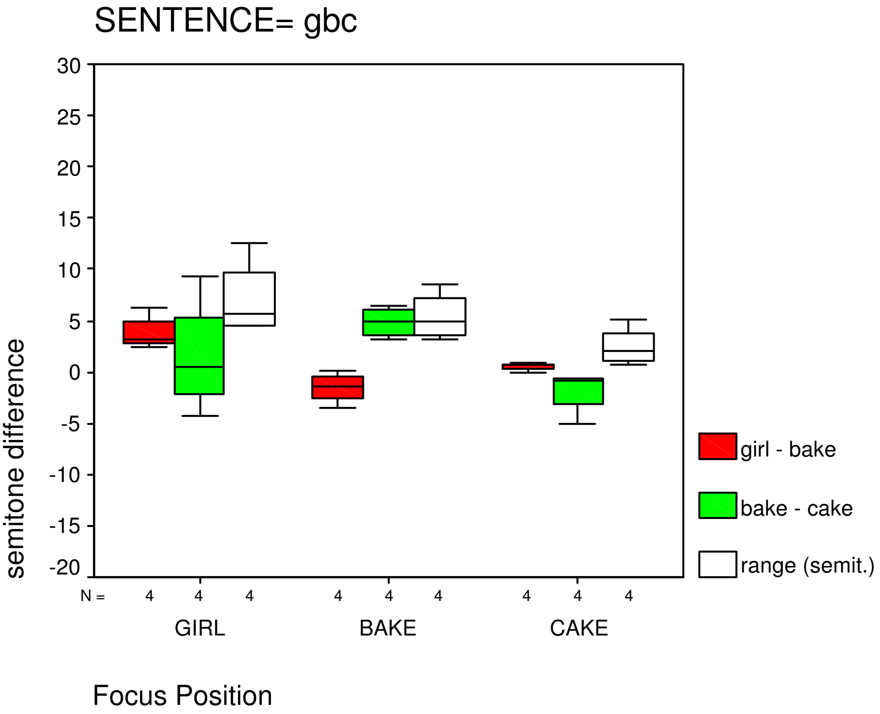
Appendix 3.1 *Examples of picture prompts (created by Barry O'Halpin) which were presented to the subjects with the natural speech stimuli in Experiment II for the Phrase Test (a) Focus 2 Test (b), and Focus 3 Test (c).*

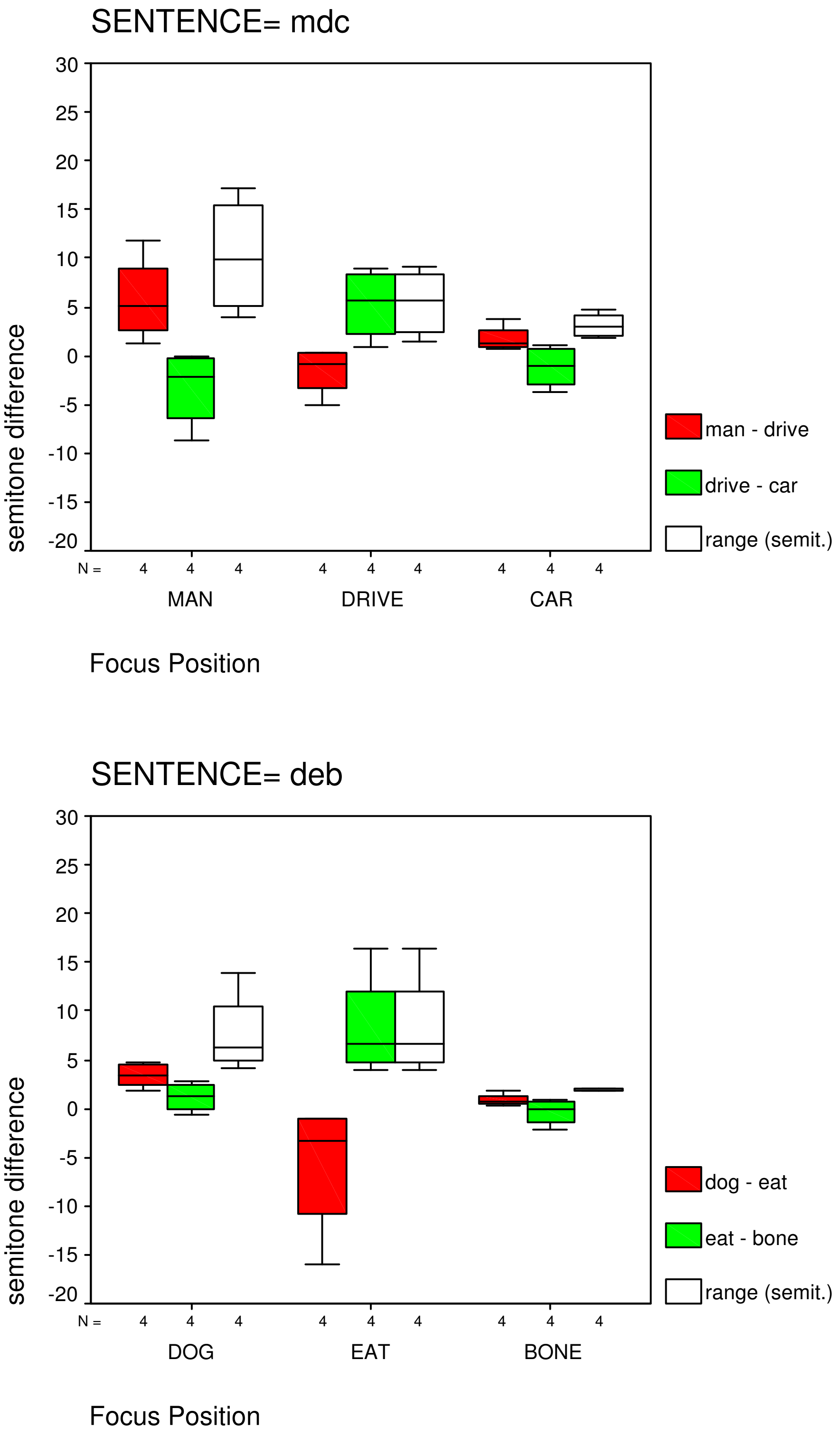
talkerid	**sentence**	**focus**	**boy**	**paint**	**boy:paint semit.**	**ing**	**boat**	**paint:boat semit**	**focus**	**min**	**max**	**range (Hz)**	**range (semit.)**
1	bpb	BOY	109	83	4.72	89	77	1.30	1	77	109	32	6.02
2	bpb	BOY	160	126	4.14	129	121	0.70	1	121	160	39	4.84
3	bpb	BOY	102	89	2.36	111	85	0.80	1	85	111	26	4.62
4	bpb	BOY	335	230	6.51	233	212	1.41	1	212	335	124	7.92
1	bpb	PAINT	103	112	-1.45	91	86	4.57	2	86	112	26	4.57
2	bpb	PAINT	147	153	-0.69	138	34	26.04	2	34	153	120	26.04
3	bpb	PAINT	90	92	-0.38	95	90	0.38	2	90	95	5	0.94
4	bpb	PAINT	271	339	-3.88	263	211	8.21	2	211	339	128	8.21
1	bpb	BOAT	107	90	3.00	95	104	-2.50	3	90	107	16	3.00
2	bpb	BOAT	160	145	1.70	142	148	-0.35	3	142	160	18	2.07
3	bpb	BOAT	94	91	0.56	90	95	-0.74	3	90	95	5	0.94
4	bpb	BOAT	267	231	2.51	255	295	-4.23	3	231	295	64	4.23
talkerid	**sentence**	**focus**	**dog**	**eat**	**dog:eat semit.**	**ing**	**bone**	**eat:bone semit**	**focus**	**min**	**max**	**range (Hz)**	**range (semit.)**
1	deb	DOG	102	87	2.75	84	74	2.80	1	74	102	28	5.56
2	deb	DOG	150	118	4.15	122	122	-0.58	1	118	150	31	4.15
3	deb	DOG	96	86	1.90	43	84	0.41	1	43	96	54	13.90
4	deb	DOG	311	236	4.78	219	209	2.10	1	209	311	102	6.88
1	deb	EAT	95	101	-1.06	90	80	4.04	2	80	101	22	4.04
2	deb	EAT	150	160	-1.12	146	117	5.42	2	117	160	42	5.42
3	deb	EAT	89	224	-15.98	101	87	16.37	2	87	224	137	16.37
4	deb	EAT	229	318	-5.68	283	203	7.77	2	203	318	115	7.77
1	deb	BONE	98	96	0.36	88	93	0.55	3	88	98	10	1.86
2	deb	BONE	144	137	0.86	142	130	0.91	3	130	144	15	1.77
3	deb	BONE	92	83	1.78	82	86	-0.61	3	82	92	9	1.99
4	deb	BONE	239	231	0.59	231	261	-2.11	3	231	261	29	2.11
talkerid	**sentence**	**focus**	**girl**	**bak**	**girl:bak semit.**	**ing**	**cake**	**bak:cake semit.**	**focus**	**min**	**max**	**range (Hz)**	**range (semit.)**
1	gbc	GIRL	104	90	2.50	80	84	1.19	1	80	104	24	4.54
2	gbc	GIRL	141	116	3.38	115	68	9.25	1	68	141	72	12.63
3	gbc	GIRL	102	85	3.16	110	109	-4.31	1	85	110	25	4.46
4	gbc	GIRL	314	218	6.32	214	220	-0.16	1	214	314	100	6.64
1	gbc	BAKE	102	101	0.17	115	70	6.35	2	70	115	44	8.59
2	gbc	BAKE	138	148	-1.21	140	116	4.22	2	116	148	32	4.22
3	gbc	BAKE	89	98	-1.67	91	82	3.09	2	82	98	16	3.09
4	gbc	BAKE	245	299	-3.45	248	216	5.63	2	216	299	83	5.63
1	gbc	CAKE	104	99	0.85	108	103	-0.69	3	99	108	9	1.51
2	gbc	CAKE	145	141	0.48	140	146	-0.60	3	140	146	6	0.73
3	gbc	CAKE	88	88	0.00	101	94	-1.14	3	88	101	13	2.39
4	gbc	CAKE	225	216	0.71	240	289	-5.04	3	216	289	73	5.04

talkerid	sentence	focus	man	driv	man:driv semit.	ing	car	driv:car semit	focus	min	max	range (Hz)	range (semit.)
1	mdc	MAN	103	82	3.95	83	84	-0.42	1	82	103	21	3.95
2	mdc	MAN	149	75	11.88	67	124	-8.70	1	67	149	82	13.84
3	mdc	MAN	87	81	1.24	38	102	-3.99	1	38	102	65	17.09
4	mdc	MAN	294	207	6.07	212	209	-0.17	1	207	294	87	6.07
1	mdc	DRIVE	93	103	-1.77	84	84	3.53	2	84	103	19	3.53
2	mdc	DRIVE	145	144	0.12	142	86	8.92	2	86	145	59	9.04
3	mdc	DRIVE	88	86	0.40	81	82	0.82	2	81	88	7	1.43
4	mdc	DRIVE	241	322	-5.02	268	205	7.82	2	205	322	118	7.82
1	mdc	CAR	103	97	1.04	90	96	0.18	3	90	103	13	2.34
2	mdc	CAR	144	138	0.74	134	130	1.03	3	130	144	15	1.77
3	mdc	CAR	102	82	3.78	78	93	-2.18	3	78	102	24	4.64
4	mdc	CAR	236	217	1.45	232	269	-3.72	3	217	269	52	3.72

Appendix 3.2 *Mean F_0 measurements for target words/syllables in focussed and unfocussed positions in Experiment II stimuli. Four different talkers produced the four target sentences: bpb (the boy is paining a boat); deb (the dog is eating a bone); gbc (the girl is baking a cake); and mdc (the man is driving a car). The range in the largest change in average F_0 over the target syllables is expressed in Hz and semitones for each sentence. Differences between target focus words and neighbouring words are also expressed in semitones.*


SENTENCE= bpb


SENTENCE= gbc



Appendix 3.3 *Boxplots showing semitone differences between target focus words and neighbouring words for initial medial and final focus position in each of the stimulus sentences presented in Experiment II i.e. gbc (the girl is baking a cake); mdc (the man is driving a car); bpb (the boy is painting a boat); deb (the dog is eating a bone).*

NH Talkers	Talker 1 (male)	Talker 2 (female)	Talker 3 (male)	Talker 4 (female)
Range (semit.)				
	4.04	4.53	3.78	6.37
Range median duration (msecs)				
	164	127	136	101
	Range in amplitude (dB)			
	9	8	8	9

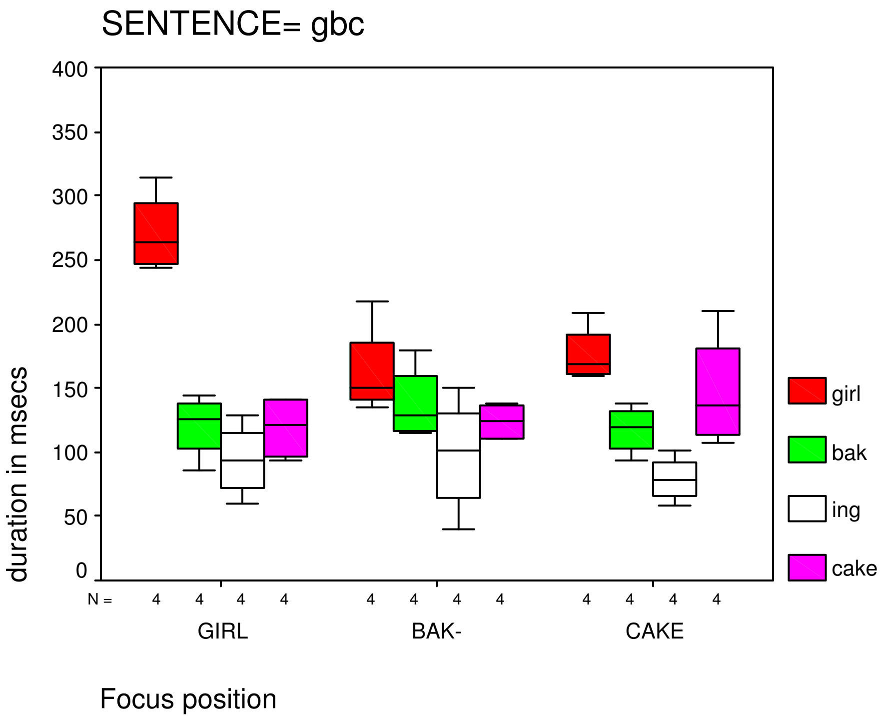
Appendix 3.4 *The median range of semitone differences between target focus and neighbouring words are presented for the NH talkers who produced the Focus 3 stimuli in Experiment II. The medians of the largest change in duration and amplitude are also presented for these talkers.*

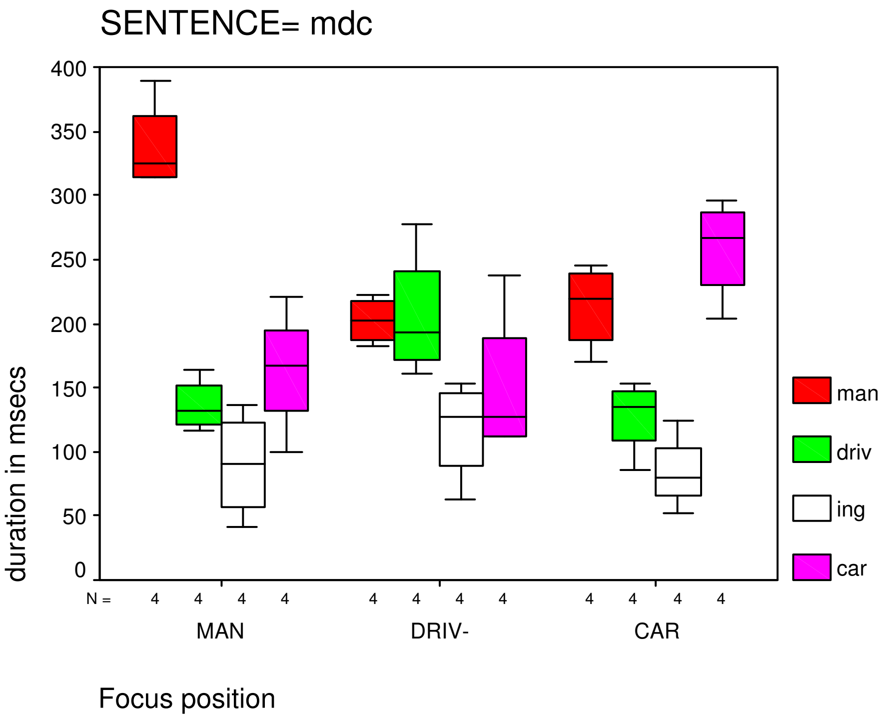
Duration Measurements (msecs)for target words/syllables in Focus 3 stimuli in Experiment II

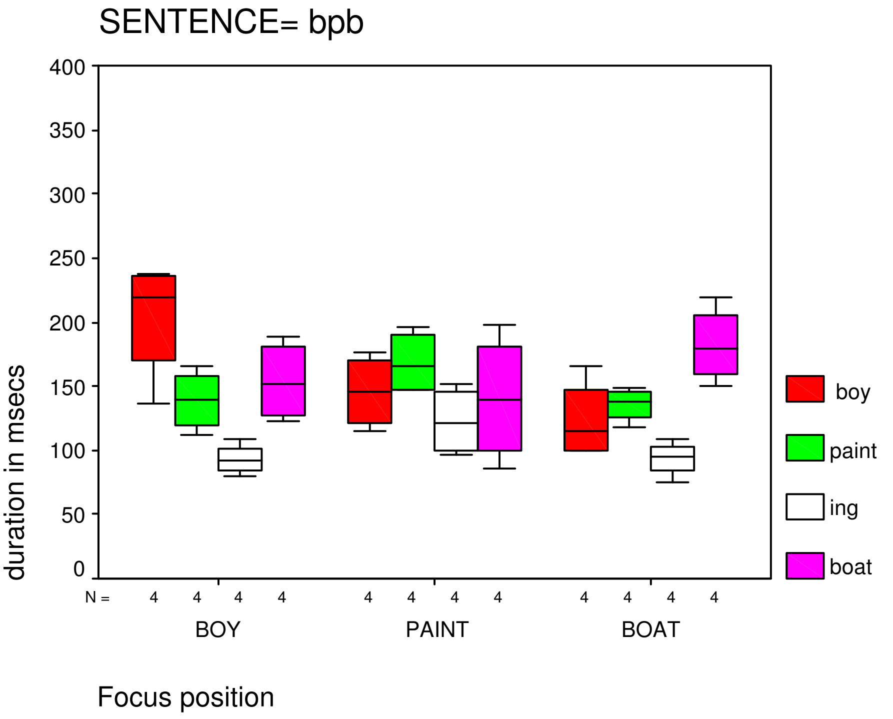
TALKERID	SENTENCE	FOCUS	boy	paint	ing	boat	FPOS	min	max	range
1	bpb	BOY	203	111	108	174	1	108	203	95
2	bpb	BOY	137	127	80	123	1	80	137	57
3	bpb	BOY	235	166	89	188	1	89	235	146
4	bpb	BOY	237	151	95	130	1	95	237	143
1	bpb	PAINT	164	148	140	165	2	140	165	25
2	bpb	PAINT	115	147	96	86	2	86	147	60
3	bpb	PAINT	176	197	104	198	2	104	198	95
4	bpb	PAINT	127	183	151	114	2	114	183	68
1	bpb	BOAT	166	149	75	191	3	75	191	115
2	bpb	BOAT	100	118	96	167	3	96	167	71
3	bpb	BOAT	128	134	93	219	3	93	219	126
4	bpb	BOAT	101	141	110	151	3	101	151	50
			dog	**eat**	**ing**	**bone**				
1	deb	DOG	266	78	90	299	1	78	299	221
2	deb	DOG	229	94	107	188	1	94	229	135
3	deb	DOG	293	64	52	136	1	52	293	241
4	deb	DOG	237	61	99	186	1	61	237	177
1	deb	EAT	223	140	87	231	2	87	231	144
2	deb	EAT	154	116	54	214	2	54	214	160
3	deb	EAT	168	187	108	199	2	108	199	90
4	deb	EAT	175	129	121	182	2	121	182	61
1	deb	BONE	185	108	52	326	3	52	326	274
2	deb	BONE	162	81	61	343	3	61	343	282
3	deb	BONE	252	112	52	346	3	52	346	293
4	deb	BONE	211	96	96	254	3	96	254	158
			girl	**bak**	**ing**	**cake**				
1	gbc	GIRL	276	131	60	141	1	60	276	216
2	gbc	GIRL	250	85	85	93	1	85	250	165
3	gbc	GIRL	314	144	129	142	1	129	314	185
4	gbc	GIRL	244	121	102	100	1	100	244	144
1	gbc	BAK	218	119	40	138	2	40	218	178
2	gbc	BAK	152	115	89	110	2	89	152	63
3	gbc	BAK	149	179	112	136	2	112	179	67
4	gbc	BAK	134	138	150	111	2	111	150	38
1	gbc	CAKE	209	126	59	152	3	59	209	150
2	gbc	CAKE	159	94	72	107	3	72	159	87
3	gbc	CAKE	164	138	83	209	3	83	209	126
4	gbc	CAKE	175	112	101	120	3	101	175	73
			man	**driv**	**ing**	**car**				
1	mdc	MAN	313	124	73	221	1	73	313	241
2	mdc	MAN	316	117	41	165	1	41	316	275
3	mdc	MAN	389	139	107	169	1	107	389	282
4	mdc	MAN	335	164	137	99	1	99	335	235
1	mdc	DRIVE	222	203	116	139	2	116	222	106
2	mdc	DRIVE	182	161	63	114	2	63	182	119
3	mdc	DRIVE	214	277	154	237	2	154	277	124

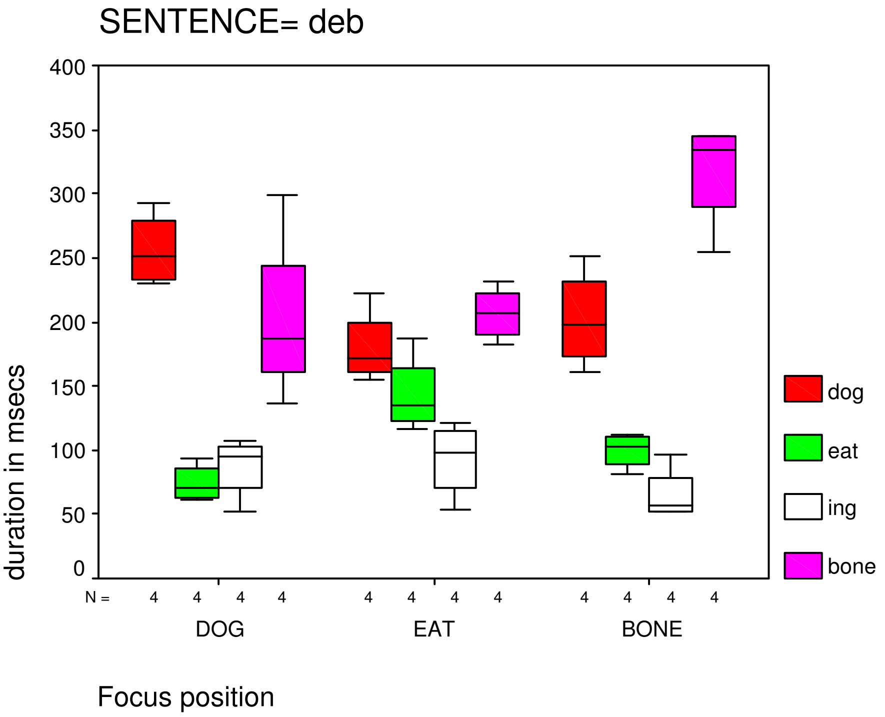
4	mdc	DRIVE	192	182	138	111	2	111	192	80
1	mdc	CAR	170	140	53	277	3	53	277	225
2	mdc	CAR	234	86	82	257	3	82	257	175
3	mdc	CAR	204	154	79	296	3	79	296	217
4	mdc	CAR	245	130	125	203	3	125	245	121
bpb	**the boy is painting a boat**									
deb	**the dog is eating a bone**									
gbc	**the girl is baking a cake**									
mdc	**the man is driving a car**									

Appendix 3.5 *Duration measurements in msecs for the target words/syllables in focussed and unfocussed position in Experiment II stimuli. Four different sentences (bpb, deb, gbc, and mdc) were produced by four talkers.*


SENTENCE= gbc
400
350
300
250
200
150
100
50
0
duration in msecs
N =
4
4
4
4
4
4
4
4
4
4
4
4
GIRL
BAK-
CAKE
Focus position
🟥 girl
🟩 bak
☐ ing
🟪 cake


SENTENCE= mdc
400
350
300
250
200
150
100
50
0
duration in msecs
N =
4
4
4
4
4
4
4
4
4
4
4
4
MAN
DRIV-
CAR
Focus position
🟥 man
🟩 driv
☐ ing
🟪 car



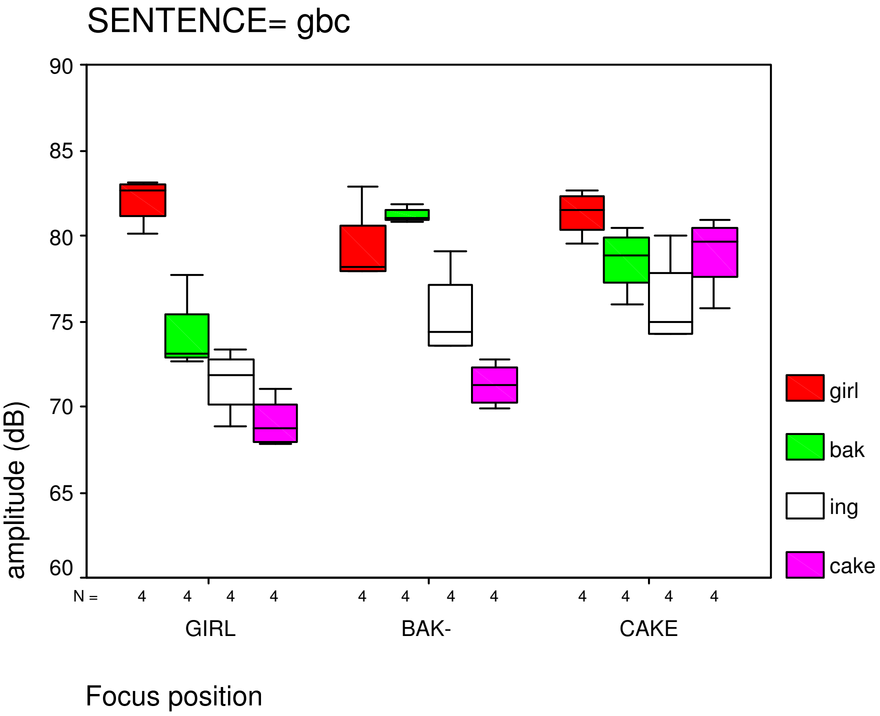


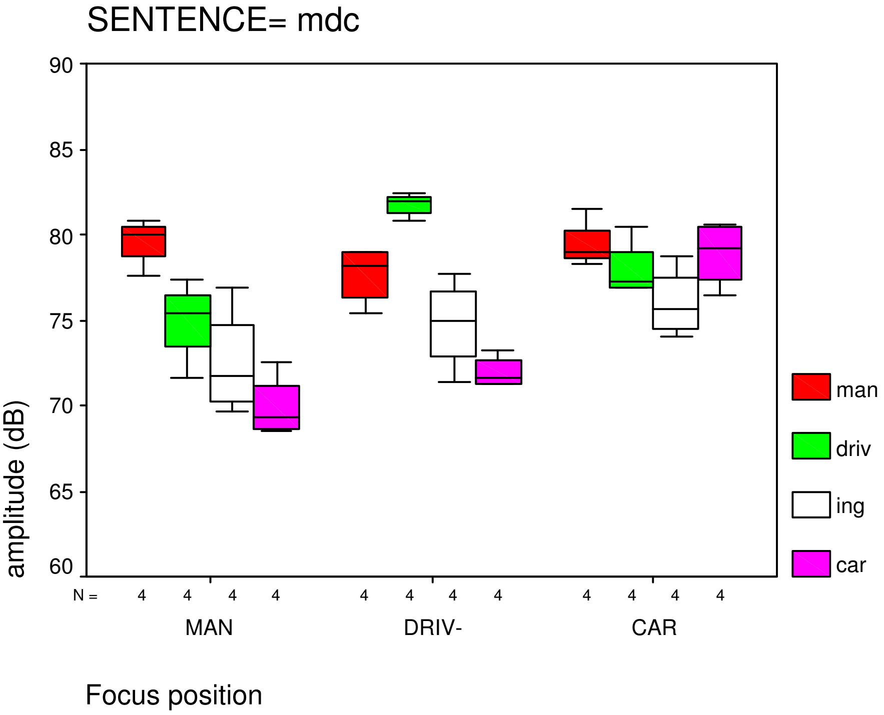
Appendix 3.6 *Boxplots for the NH stimuli in Experiment II showing durations of target focus words/ syllables in different focus position for the four stimulus sentences bpb (the boy is painting a boat); gbc (the girl is baking a cake); mdc (the man is driving a car); deb (the dog is eating a bone).*

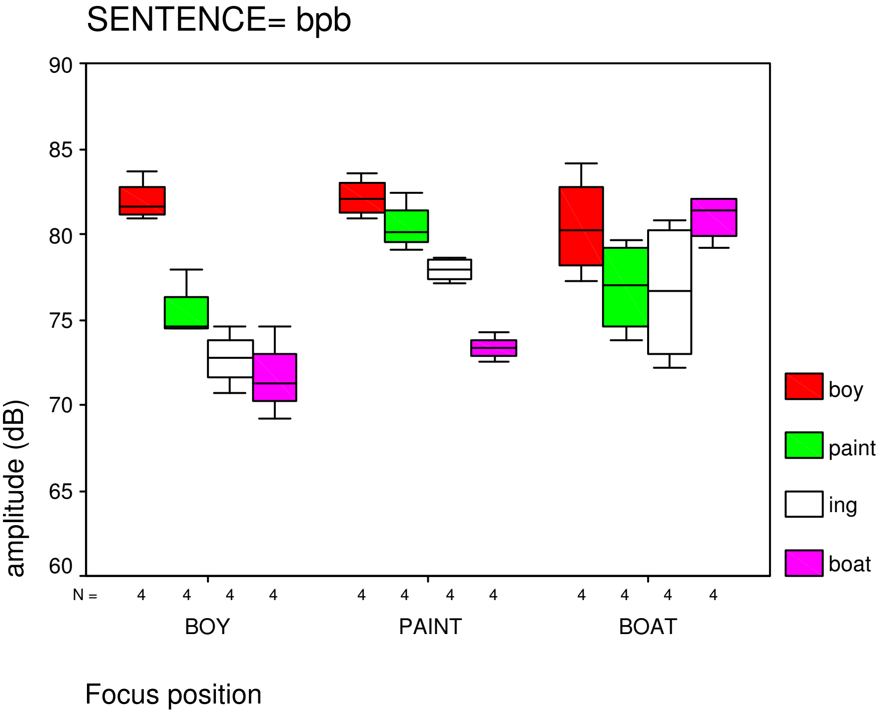
Amplitude Measurements (dB) target words/syllables in Focus 3 stimuli in Experiment II										
TALKERID	FOCUS	SENTENCE	boy	paint	ing	boat	FPOS	min	max	range
1	BOY	bpb	84	74	72	71	1	71	84	12
2	BOY	bpb	82	74	71	69	1	69	82	13
3	BOY	bpb	81	78	73	75	1	73	81	8
4	BOY	bpb	81	75	75	71	1	71	81	10
1	PAINT	bpb	84	82	78	73	2	73	84	10
2	PAINT	bpb	82	79	77	72	2	72	82	9
3	PAINT	bpb	81	80	77	73	2	73	81	7
4	PAINT	bpb	82	80	79	74	2	74	82	8
1	BOAT	bpb	84	79	80	82	3	79	84	5
2	BOAT	bpb	77	74	74	79	3	74	79	6
3	BOAT	bpb	79	75	72	81	3	72	81	8
4	BOAT	bpb	81	80	81	82	3	80	82	3
			dog	**eat**	**ing**	**bone**				
1	DOG	deb	82	78	74	72	1	72	82	10
2	DOG	deb	80	73	71	66	1	66	80	14
3	DOG	deb	80	74	71	72	1	71	80	9
4	DOG	deb	80	75	73	70	1	70	80	10
1	EAT	deb	82	84	79	76	2	76	84	8
2	EAT	deb	80	79	79	75	2	75	80	5
3	EAT	deb	81	79	77	73	2	73	81	8
4	EAT	deb	80	75	77	73	2	73	80	7
1	BONE	deb	82	79	77	79	3	77	82	6
2	BONE	deb	80	77	76	78	3	76	80	4
3	BONE	deb	81	79	75	79	3	75	81	6
4	BONE	deb	80	77	78	80	3	77	80	3
			girl	**bak**	**ing**	**cake**				
1	GIRL	gbc	83	73	71	71	1	71	83	12
2	GIRL	gbc	82	73	72	69	1	69	82	13
3	GIRL	gbc	80	73	69	68	1	68	80	12
4	GIRL	gbc	83	78	73	68	1	68	83	15
1	BAKE	gbc	83	81	74	73	2	73	83	10
2	BAKE	gbc	78	81	74	72	2	72	81	9
3	BAKE	gbc	78	81	75	70	2	70	81	11
4	BAKE	gbc	78	82	79	71	2	71	82	11
1	CAKE	gbc	83	78	76	81	3	76	83	7
2	CAKE	gbc	82	76	74	76	3	74	82	7
3	CAKE	gbc	80	79	74	79	3	74	80	5
4	CAKE	gbc	81	80	80	80	3	80	81	1
			man	**driv**	**ing**	**car**				
1	MAN	mdc	80	75	71	70	1	70	80	10
2	MAN	mdc	80	76	72	73	1	72	80	7
3	MAN	mdc	78	72	70	69	1	69	78	9
4	MAN	mdc	81	77	77	69	1	69	81	12
1	DRIV	mdc	75	82	71	71	2	71	82	11
2	DRIV	mdc	77	82	76	71	2	71	82	11
3	DRIV	mdc	79	81	74	73	2	73	81	8
4	DRIV	mdc	79	82	78	72	2	72	82	10
1	CAR	mdc	78	80	76	81	3	76	81	4
3	CAR	mdc	79	77	75	78	3	75	79	4
3	CAR	mdc	79	77	74	76	3	74	79	5
4	CAR	mdc	82	77	79	80	3	77	82	5

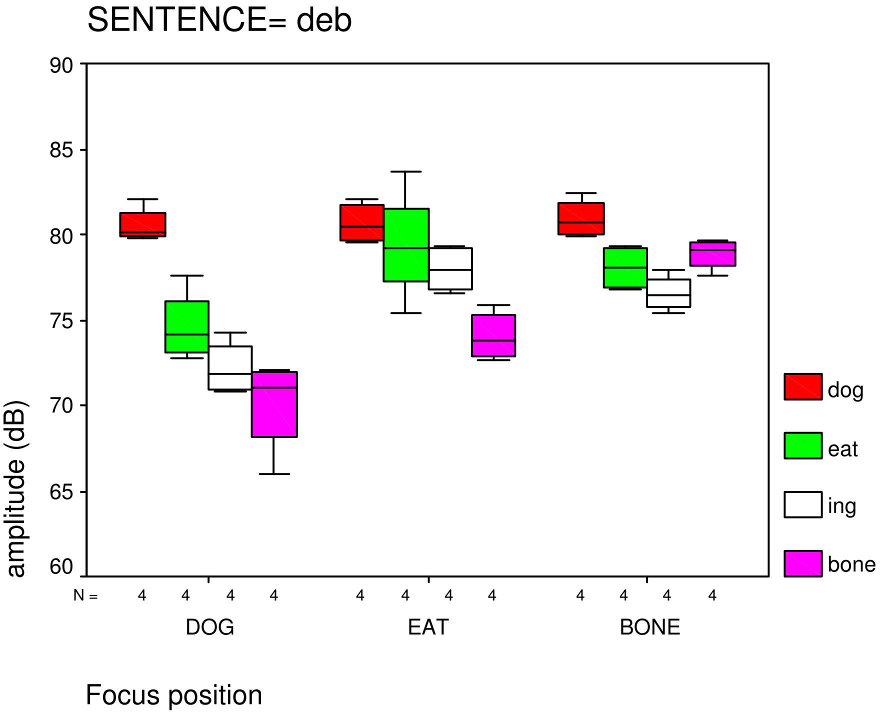
bpb	**the boy is painting a boat**
deb	**the dog is eating a bone**
gbc	**the girl is baking a cake**
mdc	**the man is driving a car**

Appendix 3.7 *Amplitude measurements of the target focus words in four sentences (bpb, deb, gbc and mdc) in the perception stimuli in Experiment II***.**


SENTENCE= gbc
90
85
80
75
70
65
60
amplitude (dB)
N =
4
4
4
4
4
4
4
4
4
4
4
4
GIRL
BAK-
CAKE
Focus position
girl
bak
ing
cake


SENTENCE= mdc
90
85
80
75
70
65
60
amplitude (dB)
N =
4
4
4
4
4
4
4
4
4
4
4
4
MAN
DRIV-
CAR
Focus position
man
driv
ing
car





Appendix 3.8 *Boxplots showing amplitude measurements for each of the target focus words in initial, medial and final position presented in the four stimulus sentences bpb (the boy is painting a boat); gbc (the girl is baking a cake); deb (the dog is eating a bone); mdc (the man is driving a car).*

Scores (%) talkers in Experiment II stimuli					
CI subjects	**Talker 1 (male)**	**Talker 2 (female)**	**Talker 3 (male)**	**Talker 4 (female)**	**Average**
1	67	88	100	100	88.5
2	42	25	33	42	35.4
3	42	75	42	67	56.3
4	33	50	33	58	43.8
6	42	42	67	75	56.3
7	67	92	92	67	79.2
8	75	100	100	83	89.5
9	42	42	50	42	43.8
10	83	83	83	75	81.3
11	42	50	83	50	56.3
12	67	83	83	50	70.8
13	83	92	100	92	91.7
14	42	58	58	50	50.1
15	67	75	58	75	68.8
16	25	25	25	50	31.3
17	92	83	92	92	89.6
Average	**57**	**66**	**69**	**67**	**65**

Appendix 3.9 *Distribution of CI individual and group scores (%) for each of the four talkers in Focus 3 stimuli in Experiment II.*

NH & CI Perception Scores in Experiment II			Subtests		
	% scores	**Median scores %**			
	Range	**Phrase Total**	**Phrase**	**Compound**	
NH	47 -96	84	88	88	
CI	48 - 90	56	56	73	
	Range	**Focus 2 Total**	**Focus position 1**	**Focus position 2**	
NH	63 - 100	94	97	100	
CI	38 - 100	66	75	63	
	Range	**Focus 3 Total**	**Focus position 1**	**Focus position 2**	**Focus position 3**
NH	65 - 100	91.7	94	84	88
CI	31 -93	62.5	72	59	66

Appendix 3.10 *Summary of the range (%) and median (%) scores for the NH and CI subjects in Phrase, Focus 2 and Focus 3 tests in Experiment II. Median scores (%) for subtests are also presented.*

CHAPTER FOUR

THE PRODUCTION OF FOCUS BY CI AND NH TALKERS: ACOUSTIC MEASUREMENTS OF F_0, AMPLITUDE AND DURATION

4.1 Introduction

As discussed in Chapter One (section 1.4) the physical parameters of stress (F_0, duration and intensity) contribute to the perception of stress by normal hearing listeners, and recent studies have indicated that F_0 might not always provide the most important cue (Kochanski et al., 2005; Peppé et al., 2000). Limitations of current speech processors in delivering adequate pitch information (see section 1.7) have implications for how stress is perceived by cochlear implant users, and it is possible that they can rely on other perceptual cues such as timing and loudness.

Experiment II results show that individual CI subjects who had higher scores in the speech perception tests were able to hear smaller differences in F_0 and duration in synthetic bisyllables in Experiment I. However as indicated in Figure 3.5 some of the CI children could only hear F_0 differences at or close to the maximum difference at 0.84 octaves (e.g. five in the low F_0 range and two in the high F_0 range) yet they performed significantly above chance in the Focus 3 test. This suggests that some CI children may not necessarily rely on F_0 cues to stress.

Since all except one of the CI subjects who performed above chance in the focus tests could hear duration differences less than 60 % where the maximum difference level was at 138% (see Figure 3.6) it is possible that duration provided a more salient cue to stress than F_0.

Subjects who were not hearing amplitude differences of less than 10 dB (see Figure 3.7) had a wide range in performance in the phrase and focus tests. Five such subjects performed significantly above chance in the three element focus test (Focus 3) which suggests that these CI subjects may not rely on amplitude cues to stress.

In the absence of F_0 or amplitude cues to linguistic focus, duration may be a more reliable cue for CI listeners. In Chapter Four, detailed acoustic analysis is carried out on F_0, duration and amplitude measurements for multiple tokens of a three element sentence (*the boy is painting the boat*) produced by the CI subjects. This sentence was one of four sentences (produced by four normal hearing talkers) which were presented to the CI subjects in the Focus 3 test in Experiment II.

The aims of Experiment III are to establish:

a) are there F_0 contours WITHIN sentences associated with different focus positions and are they similar to or different from to patterns produced by the four NH talkers?

b) what cues are used to convey focus in the target words by CI and NH talkers?

c) are there any differences in the use of F_0, duration and amplitude in the target words (*boy, paint(ing), boat*) ACROSS sentences types in focus and unfocussed positions?

d) are there any correlations between appropriate production of F_0 and duration, F_0 and amplitude, or duration and amplitude?

e) are there any correlations between F_0, duration or amplitude production and stimulation rate, age at production, age at switch-on, or duration of implant use?

4.2 Methods

4.2.1 Talkers

Sixteen implanted (CI) children who participated in Experiment II were also in Experiment III and comparisons could be made between their perception and production performance. Subject information is presented in Table 2.1 in Chapter One but one of the participants (C5) was unable to attend. The four NH talkers were those who recorded the stimuli for Experiment II.

4.2.2 Data

4.2.2.1 Cochlear implant production data

Recordings for the CI talkers were carried out in a quiet room either at home or in the hospital using a Tascam DA-PI Portable Digital Audio Tape Recorder (DAT) with two Sennheiser Evolution pocket receiver systems (Ew 122-p) and pocket transmitters with ME 2 omni clip-on microphones. As described in greater detail in section 3.2.2 four picture prompts were presented. Prior to recording, the children were familiarized with the vocabulary and the task with a few practice items to ensure they understood the task, the vocabulary, the sentence structure as well as the concept of the most important word in sentences with different focus positions. The 16 CI subjects were asked questions designed to elicit focus (contrast) on specific words in a three element sentence e.g. *the boy is painting the boat.* Sometimes the question was repeated to highlight the target focus position but in the recorded data no help was given otherwise. As outlined in section 3.2.2 full rather than elliptical sentences were elicited for consistency.

Q. Is the *GIRL* painting the boat?
R. No, the ***BOY*** is painting the boat.
Q. Is the boy *WASHing* the boat?
*R. No, the boy is **PAINTing** the boat*
Q. Is the boy painting the CAR?
*R. No, the boy is painting the **BOAT***

The procedure was repeated at least five times for each focus word using different sets of questions (total =240 utterances). The order of the questions varied in each set so that the target focus word was not predictable. SVO type sentences as used in Experiment II in the Focus 3 test had two pre-final focus items which did not compete with boundary markers. Unstressed syllables in between the target words might indicate whether CI talkers are able to make appropriate adjustments in F_0, duration or amplitude. To facilitate detailed acoustic analysis this sentence was chosen because the target words *boy, painting,* and *boat* with initial stop consonants which could be segmented easily. One or two sets of prompts for *the boy is painting the boat* were alternated with other sentences and stress tasks which were recorded for future analysis but not included in the present investigation. Preparation of the production materials for acoustic analysis required far more manual intervention than that been expected and due to time constraints it was not possible to analyse the additional recorded data. In the following discussion, sentences where the target words ***BOY, PAINT**ing* and ***BOAT*** are in focus are referred to respectively as Focus position 1, Focus position 2 and Focus position 3 type sentences.

4.2.2.2 Normal hearing production data

Recording procedures for the four NH talkers' production of the natural speech stimuli in Experiment II were described in section 3.2.2 in Chapter Three. Detailed analyses of three tokens of *the boy is painting the boat* with focus on different target *words (boy, paint(ing), boat*) were carried out for the four talkers (total = 36 utterances) who differed in age and gender.

4.2.3 Procedure

Digital files for each sentence were prepared using Cool Edit '96 (Syntrillium Software Corporation). All sentences were processed in PRAAT (www.praat.org Boersma & Weenink, 2005) and normalised to have the same peak amplitude. F_0, duration and amplitude measurements were carried out as follows for all the data.

4.2.3.1 Fundamental frequency (F_0)

A custom-written PRAAT script was used to carry out F_0 extraction and measurements. The waveform, spectrogram and label window were automatically displayed for each sentence, and segment intervals were labelled manually. The voiced intervals of the target words were marked and labelled on one tier and word segmentations on another so that the mean durations could be obtained for both. Another window displayed vocal pulse markings which were generated by PRAAT, and any missing pulses or double markings were corrected manually. A trimming algorithm was applied to remove local spikes from the F_0 contours (Xu, 1999) and to generate time-normalised F_0 tracks. In each syllable, the initial 15 ms was excluded from F_0 analysis to avoid the most dramatic portion of consonant perturbations (Xu and Wallace, 2004). Finally, mean F_0 for each interval was saved in a text file for each sentence.

4.2.3.2 Duration

The same PRAAT script was used to obtain duration values. Broadband spectrograms with the pitch trace and speech waveform for individual sentences were segmented and labelled by hand in PRAAT and segmentation was carried out as follows:

a) **overall sentence duration** between the release of initial 'the' and the end of devoicing in 'boat'

b) **overall duration of the target words** and syllables (*boy, painting, boat*) in focus and non-focus position as follows:
   - (i) *boy /b/* release to the end of the diphthong /ɔɪ/
   - (ii*)* *paint /p/* release to the point of closure for /t/
   - *(iii)* *ing:* onset of voicing in /ɪ/ to end of the nasal /ŋ/

(iv) boat /b/ release to the end of devoicing after the release of the final /t/. The voiceless stop /t/ is realised as a fricative [s̱] by some CI children which is not unusual for speakers of Southern Hiberno English (i.e. Irish English).

c) durations of other time points between

(v) end of ‘boy’ and the point of release of /p/ in ‘paint’

(vi) point of closure for /t/ at the end of ‘paint’ to the beginning of ‘ing’

(vii) end of ‘ing’ and the point of release of /b/ in ‘boat’

4.2.3.3 Amplitude

The same algorithms referred to above for PRAAT calculated mean amplitude for all the labelled voiced intervals for the target words in focus and non-focus position.

4.3 Results

Rationale for the analysis of the production data

The relationship between F_0, duration and amplitude

As discussed in Chapter One (see sections 1.2 and 1.4) narrow focus can be expressed by a change in pitch height or configuration (i.e. compression or expansion of F_0) in focus or post focus words or by durational and amplitude adjustments (Xu and Xu, 2005). The theoretical basis for acoustic analysis of the production data has been described in detail in section 1.4.3. As discussed in section 1.4.4 Southern Hiberno English (SHE) and Southern British English (SBE) have similar falling intonation contours in declarative sentences. Wells et al. (2004), however, report that there may be individual differences in how narrow focus is signalled in Southern British English. Although a falling glide was reported for most of their subjects there were differences in how other phonetic exponents were used e.g. silence, lengthening, loudness and pitch reset. Some studies suggest that natural speech may differ from laboratory controlled conditions and Kochanski et al. (2005) found in their study that accented syllables which were perceived as prominent by listeners were marked by duration and loudness cues and that F_0 played a minor role. But these results are not conclusive as they did not look at specific contrasts such as focus. As stated in sections 1.11.2 and 1.3.1.2 there may also be a physiological link between F_0, duration and amplitude i.e. the tension associated with an interest in the target focus word could lead to an increase in F_0 which might be accompanied by an increase in amplitude and/or duration. This would suggest that CI talkers might be able to produce appropriate increases in F_0 even though they do not have access to F_0 information through their implants. However, they would need to have acquired an abstract phonological representation of focus to be aware of the appropriate target focus word. To date there are no available reports on acquisition of focus by CI children generally and it is not yet clear whether F_0 is a necessary cue to focus (see hypothesis (i) or whether they can rely on other cues such as duration and/or amplitude.

Auditory Judgements

Auditory judgements of whether focus was conveyed on the target words are based on the impressions of a trained listener (i.e. the present investigator).

Appropriate adjustments of F_0, duration and amplitude WITHIN sentences

The line graphs in section 4.3 plotting F_0, duration and amplitude for the target words/syllables ***BOY, PAINTing*** and ***BOAT*** produced by the NH talkers *WITHIN* sentences (Figures 4.1, 4.6 and 4. 10) provided a reference point for what was considered to be visually appropriate for the CI talkers in the line graphs in Figures 4.3, 4.8 and 4.12). For example, tokens (T1, T2, …) for CI talkers for the focus word ***BOY*** were considered appropriate if they approximated any of the NH patterns which had a fall in F_0 followed by a gradual decline in F_0 or level F_0 with a rise in some cases to the post - focus syllables *paint* or *boat* (see line graphs in Figures 4.1 and 4.3). The schematic diagram in Figure 4.2 is a visual summary of the typical F_0 contours observed in the line graphs for the NH and CI subjects in sentences where ***BOY*** is the target focus word. The dashed lines represent F_0 patterns not typically produced by the NH talkers. The other target focus words/syllables ***PAINT*** and ***BOAT*** were analysed in a similar way in Figures 4.4 and 4.5. The schematic diagrams provide a visual summary for each focus position in Figures 4.2, 4.4 and 4.5. They are not in real time and the solid and dashed lines are not based on quantitative measurements. They capture the direction of intonation contours observed in the line graphs for the NH and CI talkers using a simple notation as follows:

a. *H, H+, H- L, L-* for higher or lower start F_0 points

b. *H* or *H+* for high or extra high F_0 peaks on target words or syllables

c. *F* and *R* for falling and rising F_0.

d. Other labels such as *Falling, Rising, Level, Fall to mid, High Fall, Step-up, Suspended* are used to indicate F_0 direction

Durations of the target focus words in most tokens for the NH talkers were longer relative to the average for the target words/syllables and in some cases they were only slightly longer than average. Similarly, durations of the target focus words for the CI talkers which were longer than average were considered appropriate even though some were only slightly longer than average (see Figure 4.8). Amplitude of the target focus syllables for NH was above average for most of the target focus words so for the CI talkers tokens which were greater than average amplitude were considered appropriate in the line graphs. The extent of the step - up in F_0 or fall from peak F_0 (H, H+ or H-), and the size of the increase in duration or amplitude in the target focus words varied for individual NH and CI talkers and were only considered inappropriate if the F_0 of surrounding target syllables were inappropriately boosted or not sufficiently deaccented, or if duration and amplitude of the focus words were the same or less than average for these words. However, what matters ultimately is whether focus on the appropriate target word is conveyed to a listener (i.e. the current investigator as discussed section 4.3.6.v).

F_0, duration and amplitude differences between target and neighbouring words/syllables ACROSS sentences

Additional measurements were also carried out for F_0 (Tables 4.4 – 4.11), duration (Tables 4.14 and 4.19), and amplitude (Tables 4.20 and 4.25) differences between the target words/syllables ***BOY***, ***PAINTing*** and ***BOAT*** and neighbouring words ACROSS sentences in focussed and non – focus positions for individual NH and CI talkers. Duration and amplitude were normalized so that comparisons could be drawn between different individual talkers. To normalise across NH and CI talkers with different F_0 ranges a logarithmic scale semitone scale was used to make it easier to draw comparisons between individual talkers and carry out acoustic analysis of the talkers.

Correlation tests

To establish if there are any statistical correlations between the appropriate production of F_0, duration or amplitude by CI subjects in the current study Pearson Correlation tests were carried out (see section 4.3.6. i-iii). If hypothesis (i) is supported (see section 1.1.2) and F_0 is a necessary cue to stress and intonation contrasts such as focus, the production of appropriate F_0 peaks may be accompanied by appropriate increases in amplitude and/or duration and correlations might be expected between appropriate F_0 and/or duration and amplitude. On the other hand if hypothesis (ii) is supported and F_0 is not a necessary cue to focus, there might be a correlation between the production of appropriate duration and amplitude adjustments on focus words but not with F_0. Other issues to be considered in the Pearson Correlation tests in the following section are whether there are any correlations between the production of appropriate F_0, duration or amplitude and variables such as rate of stimulation, duration of implant use, age at time of testing or age at switch-on. To date the only available reports are for CI children learning Chinese tones and the ages of the children and the results vary.

Individual subjects

The scattergraphs in Figure 4.14 provide more details on individual CI talkers than the correlation tests on the appropriate use of F_0 and amplitude, F_0 and duration, and duration and amplitude in Experiment III. The scattergraphs in Appendices 4.1 - 4.4 show for individual subjects the rate of appropriate production of F_0, duration and amplitude in relation to stimulation rate, duration of implant use, age at production, age at switch-on. Although some F_0, duration and amplitude increases look appropriate in the line graphs for many CI talkers, they may not manage to convey focus to a listener. The production of appropriate F_0, duration and amplitude for those individual CI talkers who managed to convey focus to the present investigator is of particular interest (see discussion in section 4.4).

4.3.1 Fundamental frequency (F_0) contour WITHIN sentences

As discussed in Chapter One (section 1.2), a speaker may wish to make a distinction between broad or narrow focus, given or new or contrastive information, or emphasise a particular word or syllable for grammatical purposes. Focus or contrast can be *explicit* in response to a question e.g. *is the GIRL painting the boat? no, the **BOY** is painting the boat* where ***BOY*** is highlighted and made prominent. Similarly as outlined above in 4.2.2.1 the words *paint*(ing) and *boat* can also be brought into focus in response to questions. In the discussion of the results below target sentences where ***BOY***, ***PAINT*** and ***BOAT*** are in focus are referred to as Focus position 1, Focus position 2, and Focus position 3 respectively.

As discussed in section 4.3 the line graphs showing mean F_0 for the target words in multiple tokens of each of the target sentences are presented in Figure 4.1 for the NH Talkers (i.e. N1, N2, N3 and N4), and in Figure 4.3 for the CI talkers. In the discussion of F_0 contours *WITHIN* each sentence type the terms step-up or step-down in F_0 are used to mean an increase (rise) or decrease (fall) in F_0. The terms level or suspended are used when F_0 remains at a similar level to the previous or following syllable(s). Tables 4.1, 4.2 and 4.3 summarise the number of appropriate F_0 contours for each CI talker. A 0.50 chance level of appropriate F_0 production (and also duration and amplitude) was chosen in the analysis of the results. The assumption that there was a 0.50 chance of an appropriate change of F_0, duration, or amplitude was arbitrary. It was not feasible to establish a priori probabilities for appropriate changes in a principled way, and a value of 0.50 was considered to be conservative. Assuming binomial variability and 15 sample sentences for each child, 12 samples need to be appropriate (i.e. 0.75) for the rate of appropriateness to be significantly higher than a 50% chance level. All of the CI subjects participated in Experiment III except for C5 who was unable to attend.

As explained earlier in section 4.3 the schematic diagrams presented in Figure 4.2, Figure 4.4, and Figure 4.5 connecting the target words/syllables only provide a visual rather than a quantitative summary of F_0 contours observed in the line graphs. The dashed lines represent F_0 patterns produced by some CI talkers which are not typically produced by the NH talkers except for a few individual cases. They are referred to below in the discussion of F_0 contours *WITHIN* Focus position 1 (***BOY***), Focus position 2 (***PAINT***), and Focus position 3 (***BOAT***) sentences.

Comparisons are also drawn *ACROSS* sentences for median F_0 values in target words *boy*, *paint* and *boat* in focussed and unfocussed positions. In Tables 4.4 – 4.11 the differences in the median F_0 for multiple repetitions of each target word and its neighbouring word(s) in focussed and unfocussed positions are expressed in Hertz (Hz) and semitones (semit.) for individual NH and CI talkers.

4.3.1.1 F_0 contour WITHIN Focus 1 sentences (B0Y)

NH talkers

The line graphs in Figure 4.1 show that a fall in F_0 from ***BOY*** to paint occurs consistently in three individual tokens for each of the NH talkers. A schematic summary of various possible F_0 contours in Figure 4.2 shows a fall (see solid lines) from higher and lower F_0 starting points for different talkers (i.e. H+, H, H-). There were some individual differences among talkers and tokens in the post-focus words with F_0 remaining almost level (Level e.g. N4:T2;T3[1]), or rising to *paint* or ing or *boat* (Rising e.g. N3:T1, N1:T1;T3, N2:T3), or declining gradually (Fall to Mid e.g. N1:T2; N2:T1;T2) or more strikingly to boat (High-Fall e.g. N3:T1).

---

[1] Individual tokens (T) for NH and CI talkers referred to as T1, T2, T3….

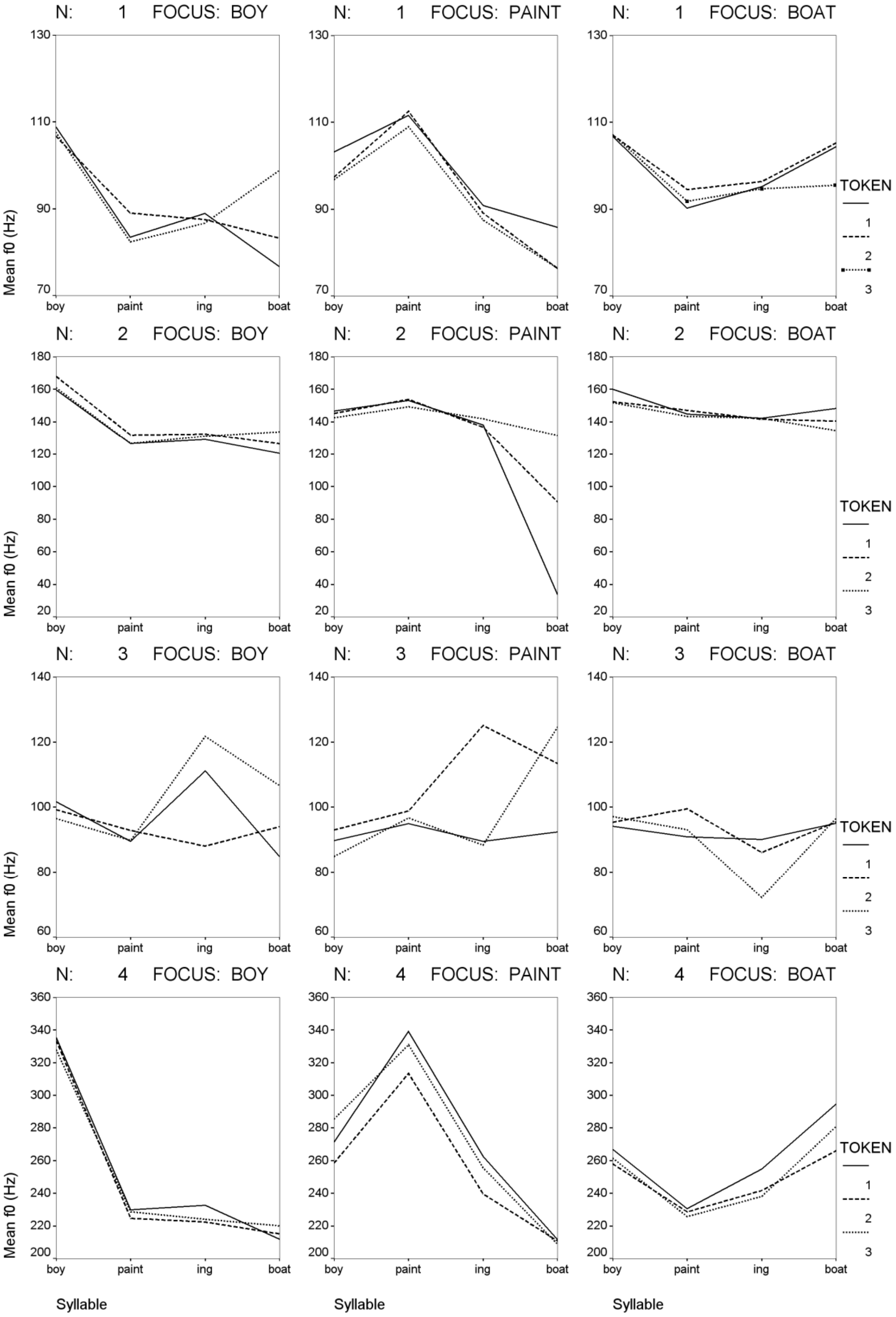


Figure 4.1 *Line graphs for the NH talkers showing mean* F_0 *for the target words* ***boy paint(ing)*** *and* ***boat*** *in multiple tokens of Focus position 1, Focus position 2 and Focus position 3 sentences. Individual tokens (1-3) are represented by different lines styles as indicated in the margin the right of the figure for each talker.*

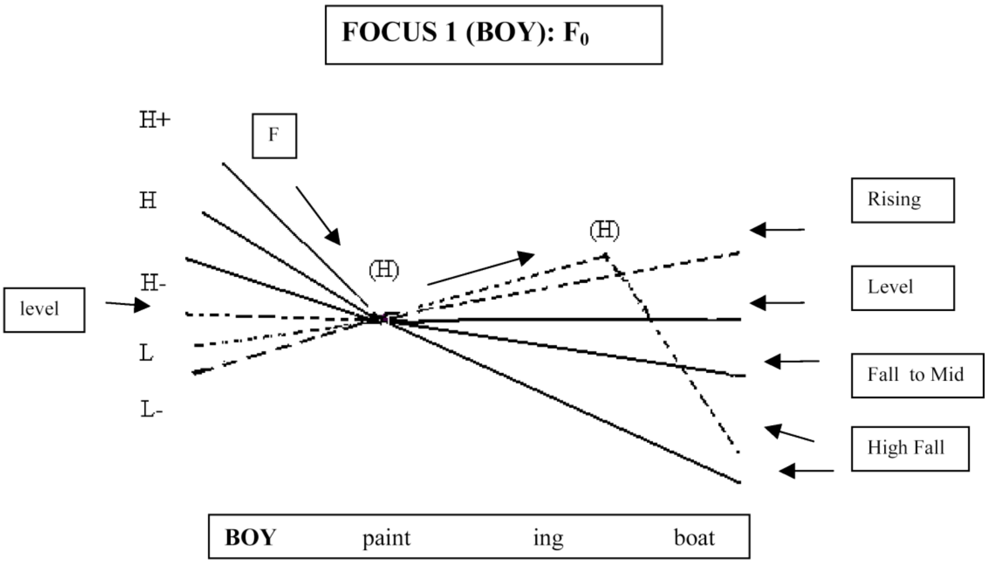


Figure 4.2 *Schematic diagram illustrating examples of F_0 contours for NH and CI talkers in Focus position 1 **(BOY)** sentences. The dashed lines illustrate F_0 patterns observed in the line graphs for CI talkers which are not typically produced by the NH talkers.*

CI talkers

The individual line graphs in Figure 4.3 show that a fall from ***BOY*** to *paint* occurs consistently across all individual tokens for only five talkers (C1, C11, C12, C14, and C15). For other talkers F_0 sometimes rises to *paint* or *ing* (e.g. C3:T4, C4:T1, C6:T5, C7:T2, C17:T1;T2) or remains almost level (e.g. C8:T1;T3, C9:T2;T5). These patterns are represented schematically in Figure 4.2 with higher starting F_0 points in the fall from ***BOY*** to paint as indicated by the solid lines (H+, H, H-). Dashed lines represent level or lower F_0 starting points (L or L-) and boosted F_0 peaks (H) on the post-focus syllables for some CI talkers which are not typically produced by NH talkers. Some individual talkers have boosted F_0 values (H) in individual tokens for *paint* and *ing* rather than deaccenting of post-focus syllables observed for three of the NH talkers (N1, N2 and N4), and the extent of the step-up in F_0 on these syllables varies. The line graphs and schematic diagram show different F_0 contours in the post-focus syllables such as a gradual decline (Fall to Mid or High Fall), or a high terminal rise on the non-focus word *boat* (Rising) or suspended F_0 (Level) which might obscure or contribute to the perception of focus on the target word ***BOY***.

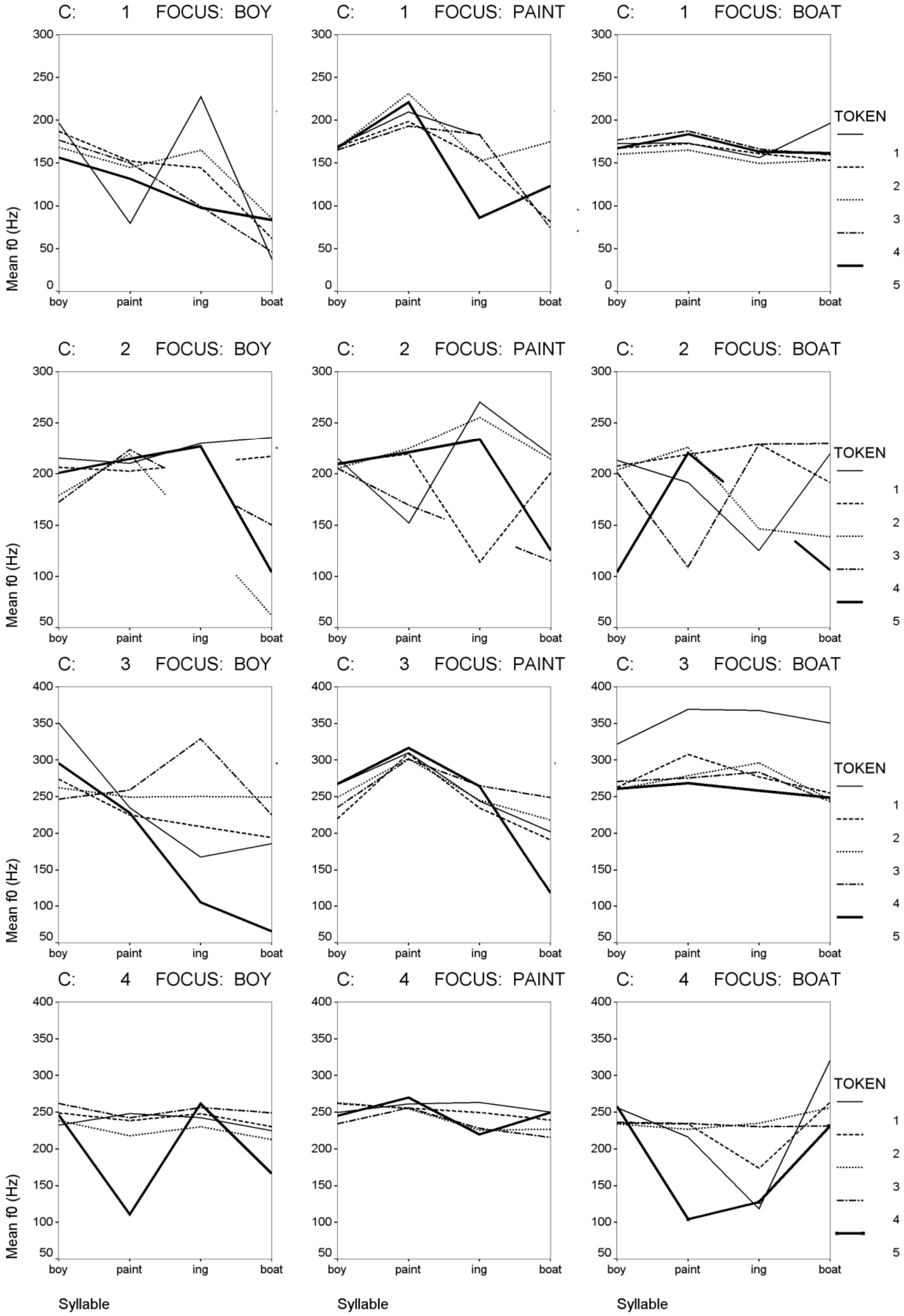
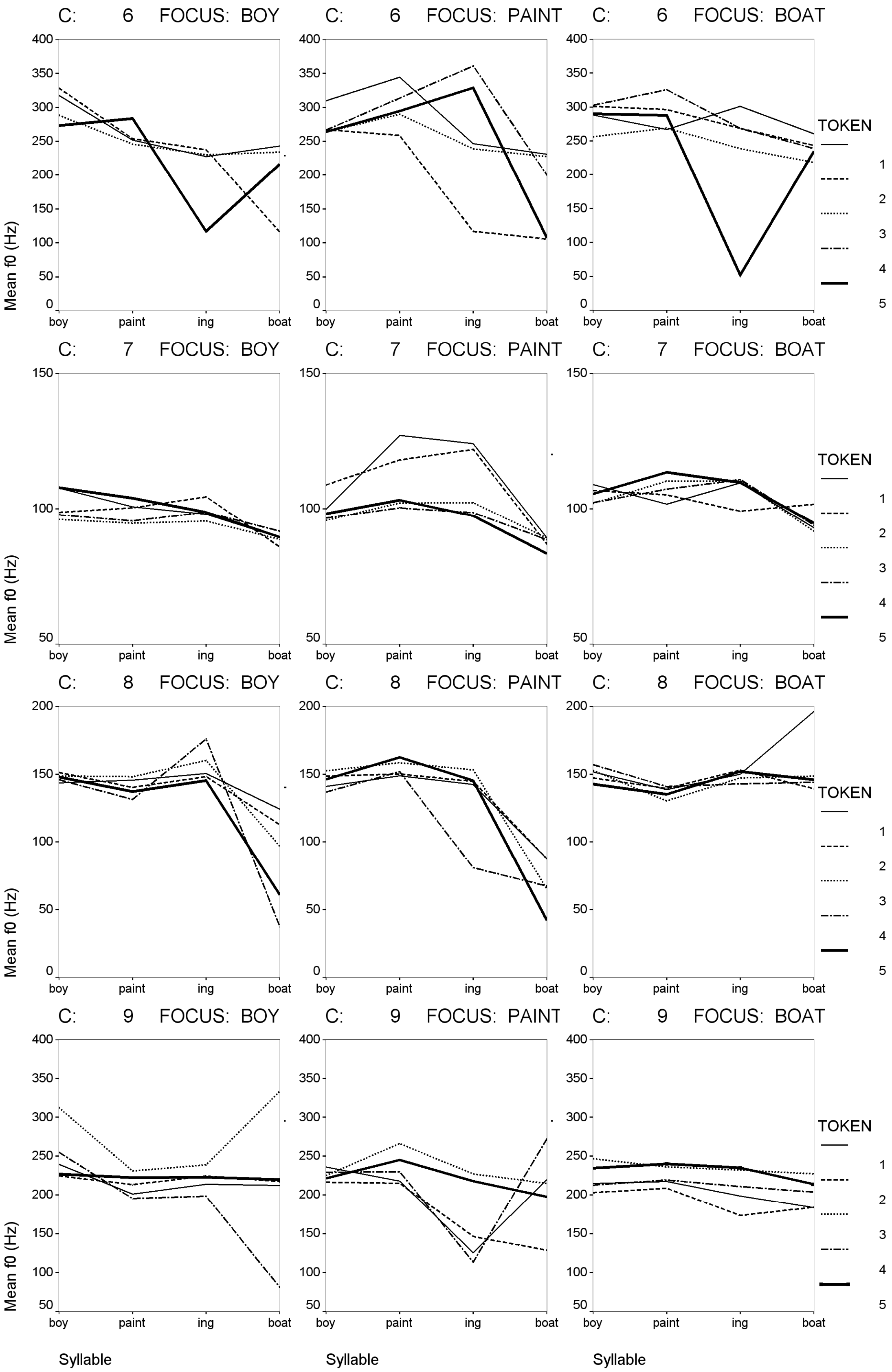
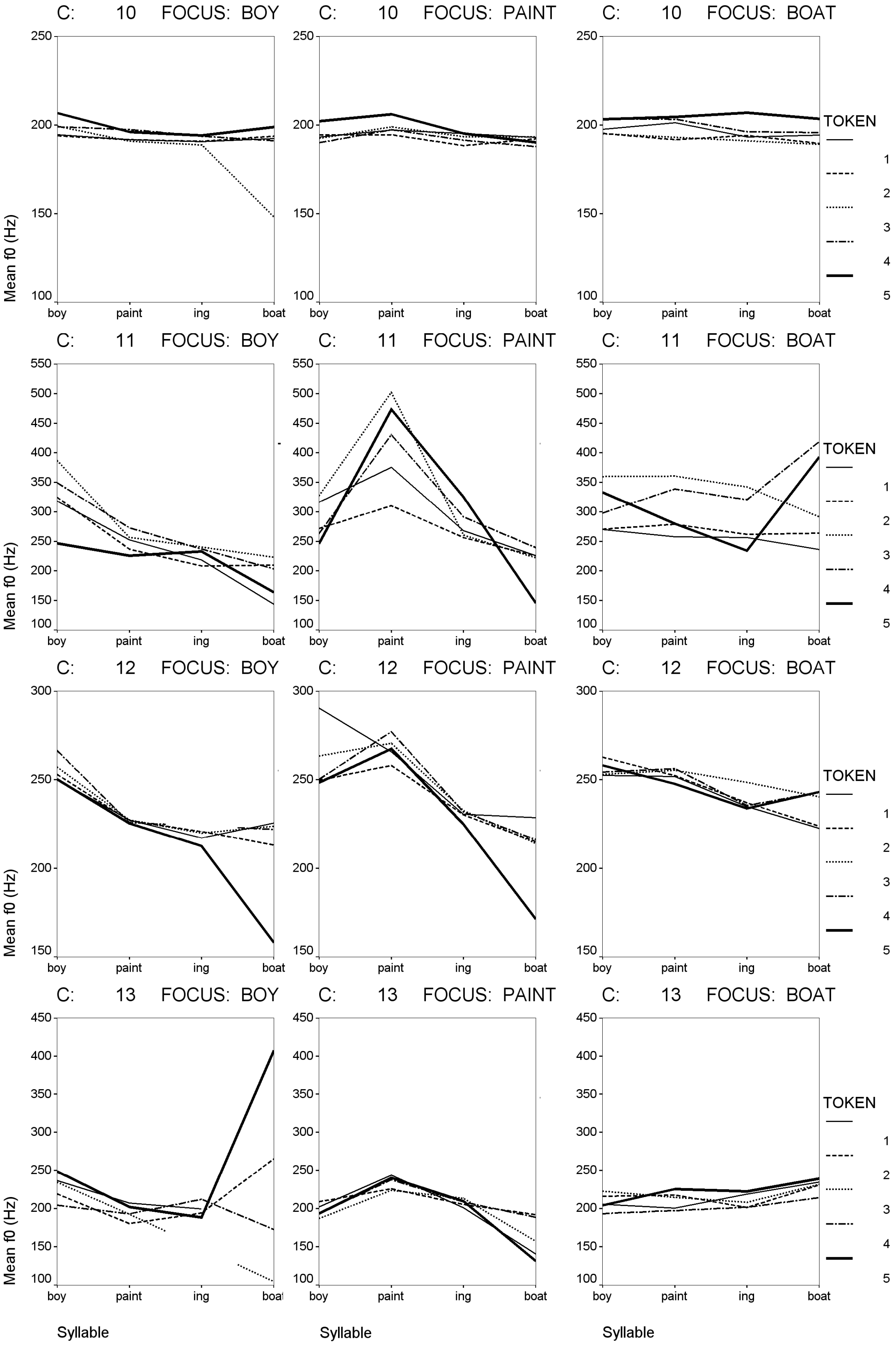


Figure 4.3 *Line graphs for the CI talkers showing mean* F_0 *for the target words boy, paint(ing) and boat in Focus position 1, Focus position 2 and Focus position 3 sentences. Individual tokens (1-5) are represented by different lines styles as indicated in the margin the right of the figure for each talker.*


C: 6 FOCUS: BOY
C: 6 FOCUS: PAINT
C: 6 FOCUS: BOAT
C: 7 FOCUS: BOY
C: 7 FOCUS: PAINT
C: 7 FOCUS: BOAT
C: 8 FOCUS: BOY
C: 8 FOCUS: PAINT
C: 8 FOCUS: BOAT
C: 9 FOCUS: BOY
C: 9 FOCUS: PAINT
C: 9 FOCUS: BOAT
Mean f0 (Hz)
0
50
100
150
200
250
300
350
400
boy
paint
ing
boat
Syllable
TOKEN
1
2
3
4
5


C: 10 FOCUS: BOY
C: 10 FOCUS: PAINT
C: 10 FOCUS: BOAT
C: 11 FOCUS: BOY
C: 11 FOCUS: PAINT
C: 11 FOCUS: BOAT
C: 12 FOCUS: BOY
C: 12 FOCUS: PAINT
C: 12 FOCUS: BOAT
C: 13 FOCUS: BOY
C: 13 FOCUS: PAINT
C: 13 FOCUS: BOAT
Mean f0 (Hz)
100
150
200
250
300
350
400
450
500
550
boy
paint
ing
boat
Syllable
TOKEN
1
2
3
4
5

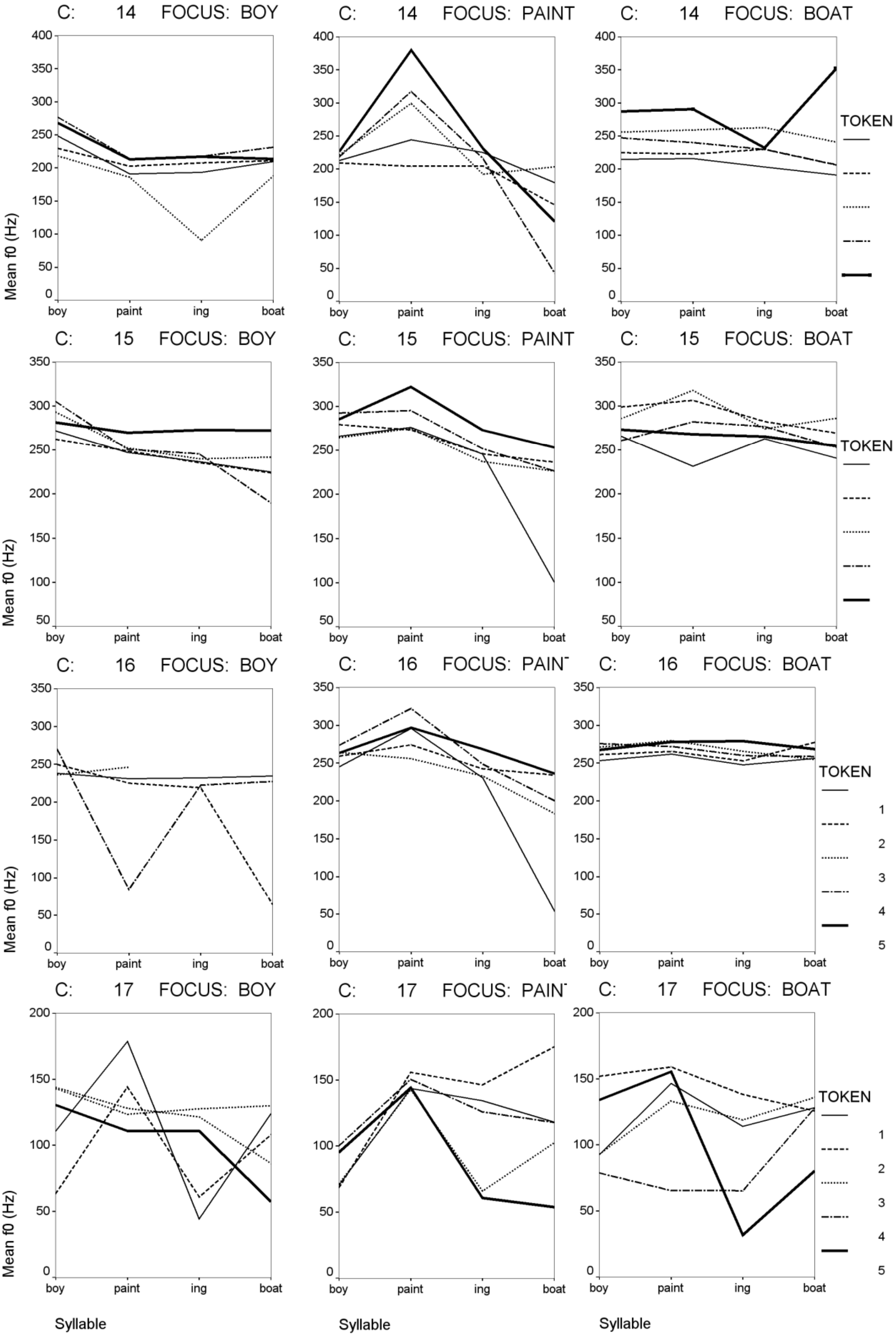


Figure 4.3 *(Continued)*

Table 4.1 below summarises for each CI talker the number of tokens with appropriate looking overall F_0 contours for Focus position 1 (***BOY***) sentences in the line graphs. Tokens which were considered appropriate approximated the patterns in the majority of the NH tokens such as a fall in F_0 from ***BOY*** to *pain*t followed by a sudden or a steady decline or levelling of F_0, sometimes with a slight rise on *ing* or *boat*. Tokens without a fall in F_0 on the focus word ***BOY*** or a levelling of F_0 throughout, or with excessive boosting of F_0 in the post-focus target syllables *paint* and *boat* were considered inappropriate. The maximum number of tokens for all talkers was five except for two talkers (C6 and C16) who had just four tokens. Table 4.1 shows that only five talkers (C1, C11, C12 C14 and C15) produced F_0 contours which were considered appropriate in all five tokens, and four talkers (C2, C4, C8, and C16) produced F_0 contours which were never considered appropriate.

BOY	***F_0 contours***				
CI Talkers	**Fall + decline**	**Fall + level**	**Fall + slight rise on *ing* or *boat***	**Appropriate tokens**	**Total tokens**
1	T4;T5		T1;T3;T2	**5**	5
2				0	5
3	T1;T2;T5			3	5
4				0	5
*6	T1; T2;T3			3	4
7	T1;T5			2	5
8				0	5
9	T4	T1		2	5
10	T5	T3		2	5
11	T1;T2;T3;T4;T5			**5**	5
12	T1;T2;T3;T4;T5			**5**	5
13	T3			1	5
14	T1;T2;T4;T5	T3		**5**	5
15	T1;T2;T3;T4	T5		**5**	5
*16				0	4
17	T3;T5	T4		3	5

Table 4.1 *Details of F_0 contours in individual tokens for the CI talkers in the line graphs in Figure 4.3 for Focus position 1* ***(BOY)*** *sentences.*

4.3.1.2 F_0 contour WITHIN *Focus position 2 sentences (PAINT)*

NH talkers

The individual line graphs in Figure 4.1 for individual tokens show that the differences in the step-up in F_0 to ***PAINT*** from *boy* are more striking for N1, N3 and N4 than for N2, and these patterns (H+, H, H-) are summarised in the schematic diagram (see solid lines) in Figure 4.4. The line graphs show that F_0 sometimes declines more dramatically after the focus word (High Fall e.g. N4) and one talker (N3) has some variation in the post-focus syllables. The rise-fall F_0 contour on *paint* (i.e. step-up followed by a fall) occurs consistently for N1, N2 and N4 although the extent varies for each talker.

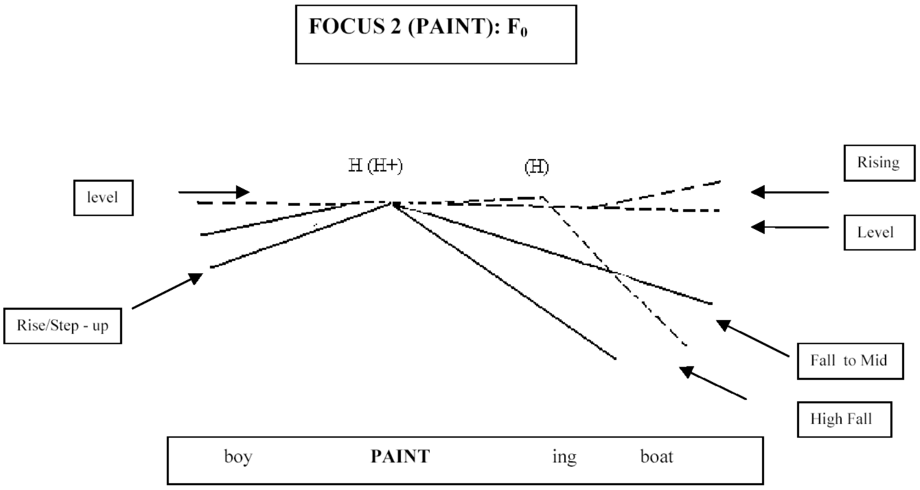


Figure 4.4 *Schematic diagram illustrating F_0 contours in the line graphs for NH and CI talkers for Focus position 2* ***(PAINT)*** *sentences. The dashed lines illustrate F_0 patterns observed in the line graphs for CI talkers which are not typically produced by the NH talkers.*

CI talkers

Line graphs for the CI talkers Figure 4.3 show that the rise or step-up in F_0 to ***PAINT*** occurred consistently in all tokens for six talkers (C1, C3, C7, C11, C13, C17) and a fall from ***PAINT*** to *boat* occurred in all tokens for seven talkers (C1, C3, C6, C8, C11, C13, C15). The step-up in F_0 to the target focus word ***PAINT,*** which was greater in some individual tokens for C3 and C11 and for other talkers (e.g. C1: T3, C7: T1, C9: T3, C12: T4), is indicated schematically by H (H+) on ***PAINT*** and (H) *ing* in Figure 4.4. The line graphs and the dashed lines in the schematic diagram also show that F_0 can sometimes remain almost level from *boy* to ***PAINT*** or from ***PAINT*** to *ing* (Level). This can be followed by a high terminal fall in F_0 after ***PAINT***, or *ing* (High

Fall), a slight decline in F_0 (Fall to Mid), or a terminal rise in F_0 to *boat* (Rising). Some of these patterns could obscure the perception of focus on the target word ***PAINT***.

Table 4.2 below summarises the number of tokens with overall F_0 contours in the line graphs for CI talkers which were considered appropriate for Focus position 2 sentences. The maximum number of tokens was five for each CI talker. Contours which approximated most NH tokens with patterns such as a rise-fall in F_0 (H or H+) on the syllables ***PAINT*** or ing, or high F_0 on *boy* and *paint* with a fall on ***PAINT*** or ing were considered appropriate. Tokens with boosted F_0 peaks on pre- or post focus syllables (*boy* or *boat*) or suspended F_0 throughout the entire sentence were not considered appropriate. Seven talkers (C1, C3, C6, C8, C11 C13, C15) had F_0 contours which were considered appropriate in all five tokens and two talkers (C12, C16) in four out of five tokens.

PAINT	**F_0 Contours**					
CI Talkers	**rise-fall on PAINT**	**rise on PAINT+ fall on ing**	**level on boy + fall on PAINT**	**(rise)-fall on PAINT+ slight rise on boat**	**Appropriate tokens**	**Total Tokens**
1	T1;T2			T3;T4;T5	**5**	5
2				T2	1	5
3	T1;T2;T3;T4;T5				**5**	5
4	T4			T5	2	5
6	T1;T2;T3	T4;T5			**5**	5
7		T1;T2			2	5
8	T4;T5	T1;T2;T3			**5**	5
9	T3;T5		T2		3	5
10	T4;T5			T2	3	5
11	T1;T2;T3;T4;T5				**5**	5
12	T2;T3;T4;T5				4	5
13	T1;T2;T3;T4;T5				**5**	5
14	T3;T4;T5				3	5
15	T1;T3;T5		T2;T4		**5**	5
16	T1;T4;T5;T2				4	5
17	T3;T4;T5				3	5

Table 4.2 *Details of F_0 contours in individual tokens for the CI talkers in the line graphs in Figure 4.3 for Focus position 2* ***(PAINT)*** *sentences.*

4.3.1.3 F_0 contour WITHIN *Focus position 3 (BOAT) sentences*

NH talkers

The line graphs in Figure 4.1 (represented schematically in Figure 4.5) show that most tokens for the NH talkers had a terminal rise to the target focus word ***BOAT*** after a fall (F) in F_0 from *boy* to *paint.* There were some differences between talkers in the extent of the terminal rise to ***BOAT*** (Step-up or Rise e.g. N1, N3 and N4) as illustrated by the solid lines in the schematic diagram in Figure 4.5, and in a few individual tokens F_0 remained level or suspended towards the end of the sentence (e.g. N1:T3, N2:T2).

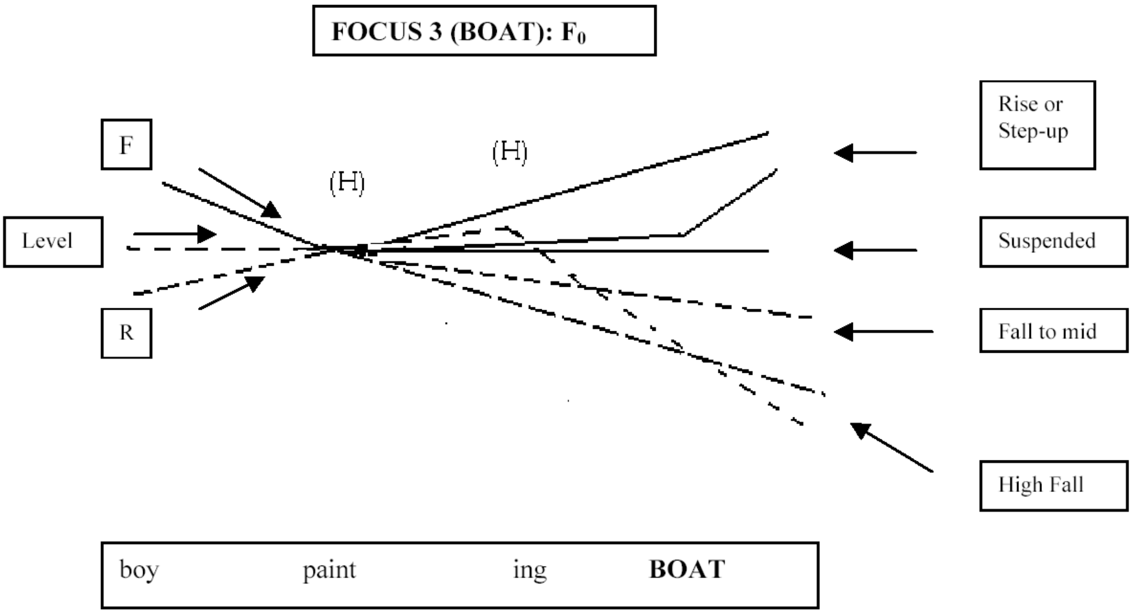


Figure 4.5 *Schematic diagram showing examples of F_0 contours for Focus position 3 (**BOAT**) sentences for NH and CI talkers.*

CI talkers

In the line graphs in Figure 4.1. One talker (C13) had an appropriate but not very striking terminal rise to ***BOAT*** in all tokens whereas it occurred only in some individual tokens for other talkers (e.g. C4, C8, C11, C17).

The schematic summary in Figure 4.5 illustrates how F_0 on the target focus word ***BOAT*** can rise, but also remain level or suspended for both the NH and CI talkers (solid lines). Sometimes a gradual (Fall to mid) or more striking fall (High Fall) in F_0 can occur for the CI talkers (dashed lines) following the pre-focus words *paint* (H) or *ing* (H) syllables. A suspended fall in F_0 generally or a more striking fall on *ing* or *boat* might also convey focus on ***BOAT*** (see dashed line in schematic diagram in Figure 4.5). Inappropriate boosting or insufficient deaccenting of F_0 the pre-focus syllables *boy, paint* and *ing* could obscure focus on the target word.

Table 4.3 below summarises appropriateness of the overall F_0 contour for five tokens of Focus (***BOAT***) sentences in the line graphs for the CI talkers. Tokens which approximated F_0 contours produced by the NH talkers were considered appropriate such as a boosted terminal rise on ***BOAT*** following smaller F_0 peaks in the pre-focus syllables, or suspended F_0 on ***BOAT***. A very striking fall after the *ing* syllable might also convey focus on ***BOAT.*** Only four talkers (C1, C4, C8, C13) had F_0 contours which were considered appropriate in all five tokens and three talkers (C9, C11, C14) in four out of five tokens.

BOAT	**F_0 contours**				
CI Talkers	**suspended F_0**	**terminal rise on *BOAT***	**suspended terminal fall on *BOAT***	**Appropriate tokens**	**Total tokens**
1		T1	T2, T3, T4, T5	**5**	5
2	T2	T1;T4		3	5
3	T1		T4, T5	3	5
4	T4	T1,T2,T3, T5		**5**	5
6			T1	1	5
7	T2			1	5
8	T2;T3;T4;T5	T1		**5**	5
9	T2 T3 ,T4;T5			4	5
10	T2;T5			2	5
11	T1;T2	T4;T5		4	5
12			T3	1	5
13		T1;T2;T3;T4;T5		**5**	5
14	T2 ;T3;T4	T5		4	5
15		T1		1	5
16	T2;T5			2	5
17	T2	T3;T4		3	5

Table 4.3 *Details of F_0 contours in individual tokens for the CI talkers in the line graphs for Focus position 3* ***(BOAT)*** *sentences.*

The term *appropriate* as used above (see also section 4.3) does not necessarily mean that F_0 contours were always identical to those produced by the NH talkers. In some cases F_0 patterns may have been approaching what was typical for the NH talkers in the present study.

As discussed in section 1.2, contrast or focus may be a process of boosting or deaccenting of new or old information (Ladd, 1996) rather than mapping of particular acoustic correlates (e.g. F_0) onto the target syllable, or there may be expansion or compression of F_0 peaks respectively on the focus words and post-focus words (Xu and Xu, 2005). Although the NH talkers had either a rise or suspended F_0 on the target focus word ***BOAT***, it is also possible that focus or prominence might also be conveyed by a striking fall in F_0. However, in a few cases where focus was not heard on the target focus word insufficient boosting or deaccenting of amplitude or duration might have obscured appropriate F_0 contours for some CI talkers. This issue is discussed again in section 4.3.6 (vii).

4.3.2 Comparisons of target words ACROSS Focus position 1, Focus position 2 and Focus position 3 sentences: Fundamental frequency (F_0)

To normalise across NH and CI talkers with different F_0 ranges, F_0 measurements are expressed below using a logarithmic scale (i.e. semitones) in addition to a linear scale (i.e. Hertz). The logarithmic scales relate more to the perception of pitch and make it easier to draw comparisons between different talkers. In the following sections the difference between the median F_0 values (in Hz and semitones) for target words ***BOY, PAINT,*** and ***BOAT*** and neighbouring word (in focus and non - focus positions) are presented for individual NH talkers and for individual CI talkers in Tables 4.4 – 4.11 below. T tests were carried for the CI talkers only.

4.3.2.1 Focus position 1 (BOY: paint) and Focus position 3 (boy: paint)

NH Talkers

Table 4.4 and the line graphs in Figure 4.1 show that all four NH talkers (N1, N2, N3 and N4) had a bigger step-down in median F_0 from ***BOY*** to *paint* in Focus position 1 sentences of between 10 - 105 Hz or 1.82 -6.51 semit.[2] than in Focus position 3 sentences (range 3Hz - 31 Hz or 0.56 - 2 .78 semit.) The step-down or fall in F_0 and the difference between Focus position 1 and Focus position 3 was greatest for N4 and smallest for N3.

	Focus position 1					**Focus position 3**				
	BOY	*paint*	***BOY:*** *paint*			*boy*	*paint*	*boy: paint*		
NH Talkers	**Hz**	**Hz**	**diff in Hz**	**diff in semit.**	**F_0 contour**	**Hz**	**Hz**	**diff in Hz**	**diff in semit.**	**F_0 contour**
1	108	84	24	4.35	***fall***	108	92	16	2.78	***fall***
2	160	125	35	4.27	***fall***	152	145	7	0.82	***fall***
3	100	90	10	1.82	***fall***	95	92	3	0.56	***fall***
4	335	230	105	6.51	***fall***	260	229	31	2.20	***fall***

Table 4.4 *Differences in the median F_0 in Hz and semitones for* ***BOY:*** *paint (Focus position 1) and boy: paint (Focus position 3) in focussed and unfocussed positions respectively for the NH talkers.*

---

[2] the word semitones is abbreviated to semit

CI Talkers

Table 4.5 shows that there was a fall or step-down in all median F_0 values ranging from 5 Hz -75 Hz or 0.36 - 4.54 semitones from the target focus word ***BOY*** to *paint* for all except two talkers (C2 and C7). T tests for the group in Table 4.5 show that this fall was highly statistically significant.

In Focus position 3 sentences where *boy* was not in focus there was smaller decline in F_0 from *boy* to *paint* for six of the talkers (C4, C6, C8, C11, C14, C12) ranging from 2 Hz – 20 Hz (.14 –1.19 semitones). However, only four of these talkers (C6, C11, C12, C14) (in Table 4.5) showed patterns resembling the NH talkers with a more striking fall in median F_0 from ***BOY*** in Focus position 1 sentences (see underlined entries). Across the group there was no significant decline in F_0 from *boy* to *paint* when *boy* was not in focus.

	Focus position 1					**Focus position 3**				
	BOY	*paint*	***BOY:*** *paint*			*boy*	*paint*	*boy: paint*		
CI Talkers	**Hz**	**Hz**	**diff in Hz**	**diff in semit.**	**F_0 contour**	**Hz**	**Hz**	**diff in Hz**	**diff in semit.**	**F_0 contour**
1	175	145	30	3.26	***fall***	170	175	-5	-0.50	***rise***
2	200	215	-15	-1.25	rise	205	220	-15	-1.22	***rise***
3	275	235	40	2.72	***fall***	265	280	-15	-0.95	***rise***
4	245	240	*5*	*0.36*	***fall***	235	225	*10*	*0.75*	fall
6	305	255	***50***	***3.10***	***fall***	295	290	***5***	***0.30***	fall
7	98	100	-2	-0.35	rise	105	107	-2	-0.33	***rise***
8	148	140	8	0.96	***fall***	151	140	11	1.31	fall
9	240	220	20	1.51	***fall***	216	220	-4	-0.32	***rise***
10	198	192	6	0.53	***fall***	196	202	-6	-0.52	***rise***
11	325	250	***75***	***4.54***	***fall***	300	280	***20***	***1.19***	fall
12	257	227	***30***	***2.15***	***fall***	254	252	***2***	***0.14***	fall
13	235	192	43	3.50	***fall***	205	215	-10	-0.82	***rise***
14	248	202	***46***	***3.55***	***fall***	245	240	***5***	***0.36***	fall
15	280	255	25	1.62	***fall***	270	280	-10	-0.63	***rise***
16	245	230	15	1.09	***fall***	265	270	-5	-0.32	***rise***
17	130	128	2	0.27	***fall***	92	145	-53	-7.88	***rise***
mean			23.6	1.7				-4.5	-0.6	
var			544	2				261	4	
t			4.05	4.25				-1.11	-1.14	
df			15	15				15	15	
sig			***0.0005***	***0.0004***				***0.1414***	***0.1369***	

Table 4.5 *Difference in the median F_0 in Hz and semitones for five tokens of* ***BOY:*** *paint (Focus position 1) and boy: paint (Focus position 3) in focussed and unfocussed positions respectively for the CI talkers.*

4.3.2.2 Focus position 2 (boy: PAINT) and Focus position 3 (boy: paint)

NH talkers

As shown by the results displayed Table 4.6 all four NH talkers (N1, N2, N3, and N4) had a step-up or rise in median F_0 from *boy* to ***PAINT*** in Focus position 2 (5 – 60 Hz or 0.82 – 3.47 semit.) and a step -down or fall from *boy* to *paint* in Focus position 3 (2- 30Hz or 0.37 – 2.12 semit.). The step-up in F_0 to ***PAINT*** in Focus position 2, and also the step-down to the non-focus word *paint* in Focus position 3 was greater for N4 and N1 than for the other two talkers.

	Focus position 2					***Focus position 3***				
	boy	***PAINT***	*boy:* ***PAINT***			*boy*	*paint*	*boy: paint*		
NH Talkers	*Hz*	*Hz*	***diff in Hz***	***diff in semit.***	***F_0 contour***	*Hz*	*Hz*	***diff in Hz***	***diff in semit.***	***F_0 contour***
1	98	112	-14	-2.31	***rise***	108	98	10	1.68	***fall***
2	145	152	-7	-0.82	***rise***	151	145	6	0.70	***fall***
3	90	95	-5	-0.94	***rise***	94	92	2	0.37	***fall***
4	270	330	-60	-3.47	***rise***	260	230	30	2.12	***fall***

Table 4.6 *Difference in the median F_0 in Hz and semitones for boy:* ***PAINT*** *(Focus position 2) and boy: paint (Focus position 3) for the NH talkers.*

CI Talkers

Fifteen CI talkers in Table 4.7 had a rise in the median F_0 value from *boy* to the target focus word ***PAINT*** in Focus position 2 sentences ranging from 3 Hz -80 Hz or 0.34 – 12.37 semitones. T tests show that this rise was significant for the group as a whole. Five of the CI talkers (C1, C3, C13, C16, C17) had a greater increase in F_0 when ***PAINT*** was in focus than when it was not in focus (see underlined and bold entries below). The rest of the CI talkers had a fall from *boy* to *paint* when *paint* was not in focus like the NH talkers.

	Focus position 2					**Focus position 3**				
	boy	***PAINT***	*boy:* ***PAINT***			*boy*	*paint*	*boy: paint*		
CI Talkers	**Hz**	**Hz**	**diff in Hz**	**diff in semit.**	**F_0 contour**	**Hz**	**Hz**	**diff in Hz**	**diff in semit.**	**F_0 contour**
1	168	210	***-42***	***-3.86***	***rise***	170	175	***-5***	***-0.50***	***rise***
2	208	218	-10	-0.81	***rise***	205	220	-15	-1.22	***rise***
3	250	310	***-60***	***-3.72***	***rise***	265	280	***-15***	***-0.95***	***rise***
4	250	255	-5	-0.34	***rise***	235	225	10	0.75	fall
6	98	103	-5	-0.86	***rise***	295	290	5	0.30	fall
7	97	103	-6	-1.04	***rise***	105	107	-2	-0.33	***rise***
8	147	150	-3	-0.35	***rise***	151	140	11	1.31	fall
9	225	230	-5	-0.38	***rise***	216	220	-4	-0.32	***rise***
10	193	197	-4	-0.36	***rise***	196	202	-6	-0.52	***rise***
11	270	430	-160	-8.06	***rise***	300	280	20	1.19	*fall*
12	250	268	-18	-1.20	***rise***	254	252	2	0.14	fall
13	195	240	***-45***	***-3.59***	***rise***	205	215	***-10***	***-0.82***	***rise***
14	220	300	-80	-5.37	***rise***	245	240	5	0.36	fall
15	280	275	5	0.31	fall	270	280	-10	-0.63	***rise***
16	265	295	***-30***	***-1.86***	***rise***	265	270	***-5***	***-0.32***	***rise***
17	70	143	***-73***	***-12.37***	***rise***	92	145	***-53***	***-7.88***	***rise***
mean			-33.8	-2.7				-4.5	-0.6	
var			1860	11				261	4	
t			-3.14	-3.2				-1.11	-1.14	
df			15	15				15	15	
sig			***0.0034***	***0.0030***				***0.1414***	***0.1369***	

Table 4.7 *Differences in the median F_0 in Hz and semitones for boy:* ***PAINT*** *(Focus position 2) and boy: paint (Focus position 3) for the CI talkers.*

4.3.2.3 Focus position 2 (PAINT: boat) and Focus position 1 (paint: boat)

NH talkers

Table 4.8 shows that in Focus position 2 sentences there was a high fall in the median F_0 (37-120 Hz or 6.94 – 9.19 semit.) from the target focus word ***PAINT*** to *boat* for three talkers (N1, N3 and N4). One talker (N4) had a bigger fall in median F_0 for Focus position 2 (120 Hz or 7.82 semit.) than for Focus position 1 (14 Hz or 1.09 semit), whereas the median F_0 was already low on *paint* and *boat* in Focus position 1 following the focus on *boy* for two talkers (N1 and N2). For the fourth talker (N3) there was a rise or step-up in median F_0 from ***PAINT*** to *boat* in Focus position 2 (18 Hz or 3.00 semit.) with little change in F_0 from *paint* to *boat* in Focus position 1 (4Hz or .75 semit.).

	Focus position 1					**Focus position 2**				
	paint	*boat*	*paint: boat*			***PAINT***	*boat*	***PAINT**: boat*		
NH Talkers	**Hz**	**Hz**	**diff in Hz**	**diff in semit.**	**F_0 contour**	**Hz**	**Hz**	**diff in Hz**	**diff in semit.**	**F_0 contour**
1	83	83	0.00	0.00	***low level***	112	75	37	6.94	***fall***
2	125	125	0.00	0.00	***low level***	153	90	63	9.19	***fall***
3	90	94	-4	-0.75	***rise***	95	113	-18	-3	***rise***
4	229	215	14	1.09	***fall***	330	210	120	7.82	***fall***

Table 4.8 *Differences in the median F_0 in Hz and semitones for **PAINT**: boat (Focus position 2) and paint: boat (Focus position 1) for the NH talkers.*

CI talkers

Table 4.9 shows that all CI talkers had a fall in the median F_0 from ***PAINT*** to *boat* in Focus position 2 (4 Hz – 205 Hz or .36 – 13.93 semit.) and t test show that this was significant for the group as a whole. Eight of these talkers (C3, C6, C7, C8, C11, C12, C15, C17) who had a fall in F_0 in both sentence types had greater fall in Focus position 2 following the focus word (see underlined and bold entries in Table 4.9).

Focus position 1						**Focus position 2**				
	paint	*boat*	*paint: boat*			***PAINT***	*boat*	***PAINT:*** *boat*		
CI Talkers	**Hz**	**Hz**	**diff in Hz**	**diff in semit.**	**F_0 contour**	**Hz**	**Hz**	**diff in Hz**	**diff in semit.**	**F_0 contour**
1	145	60	85	15.28	***fall***	210	102	108	12.50	***fall***
2	215	150	65	6.23	***fall***	220	200	20	1.65	***fall***
3	235	195	**40**	**3.23**	***fall***	305	200	**105**	**7.31**	***fall***
4	240	220	20	1.51	***fall***	255	240	15	1.05	***fall***
6	255	225	**30**	**2.17**	***fall***	295	200	**95**	**6.73**	***fall***
7	100	88	**12**	**2.21**	***fall***	103	88	**15**	**2.72**	***fall***
8	140	95	**45**	**6.71**	***fall***	152	68	**84**	**13.93**	***fall***
9	213	217	-4	-0.32	rise	230	215	15	1.17	***fall***
10	92	92	0	0.00	level	197	193	4	0.36	***fall***
11	255	205	**50**	**3.78**	***fall***	**430**	**225**	**205**	**11.21**	***fall***
12	227	223	**4**	**0.31**	***fall***	**268**	**215**	**53**	**3.81**	***fall***
13	190	220	-30	-2.54	rise	240	160	80	7.02	***fall***
14	200	210	-10	-0.84	rise	300	145	155	12.59	***fall***
15	250	225	**25**	**1.82**	***fall***	**275**	**230**	**45**	**3.09**	***fall***
16	230	230	0	0.00	level	295	200	95	6.73	***fall***
17	128	118	**10**	**1.41**	***fall***	**143**	**118**	**25**	**3.33**	***fall***
mean			21.38	2.6				69.9	5.9	
var			897	17				3235	20	
t			2.85	2.45				4.92	4.54	
df			15	15				15	15	
sig			***0.0060***	***0.0135***				***<0.0001***	***<0.0001***	

Table 4.9 *Differences in the median F_0 in Hz and semitones for* ***PAINT:*** *boat (Focus position 2) and paint: boat (Focus position 1) for the CI talkers.*

4.3.2.4 Focus position 1 (paint: boat) and Focus position 3 (paint: BOAT)

NH talkers

Data for the Focus position 3 sentences are shown in Table 4.10. Three NH talkers (N1, N3 and N4) had a rise in the median F_0 from *paint* to the target focus word ***BOAT*** (4 - 40 Hz or .73 – 2.67 semit.) whereas the fourth talker (N2) had a 5 Hz (0.61 semit.) fall. In Focus position 1 sentences when *boat* is not in focus F_0 falls after the focus on *boy* and remains low on *boat* for subjects N1 and N2, and F_0 continues to decline for subject N4 when *boat* is not in focus. There is very little difference between the increase in F_0 in Focus position 1 and Focus position 3 sentences for N3.

Focus position 1						Focus position 3				
	paint	*boat*	*paint: boat*			*paint*	***BOAT***	*paint:* ***BOAT***		
NH Talkers	**Hz**	**Hz**	**diff in Hz**	**diff in semit.**	**F_0 contour**	**Hz**	**Hz**	**diff in Hz**	**diff in semit.**	**F_0 contour**
1	84	84	0	0.00	low level	91	105	-14	-2.48	rise
2	125	125	0	0.00	low level	145	140	5	0.61	fall
3	90	94	-4	-0.75	rise	93	97	-4	-0.73	rise
4	229	215	14	1.09	fall	240	280	-40	-2.67	rise

Table 4.10 *Differences in the median F_0 in Hz and semitones for paint:**BOAT** (Focus position 3) and paint: boat (Focus position 1) for the NH talkers.*

CI talkers

Table 4.11 shows that in Focus position 3 sentences only four CI talkers (C4, C8, C11, C13) had a terminal rise or step-up in median F_0 from *paint* to ***BOAT*** ranging from 10 – 20 Hz (0.61-1.41 semitones). Table 4.11 also shows that twelve of the CI talkers (C1, C2, C3, C6, C7, C9, C10, C12, C14, C15, C16, C17) had a fall in median F_0 from *paint* to the target focus word ***BOAT***. The fall in F_0 ranged from 4 - 50 Hz or .65 - 3.28 semitones and was significant for the CI group as a whole. However, five (C1, C2, C3, C7, C15) of the eight talkers who had a fall in median F_0 in both sentence types, had a reduced fall (but only slightly for C15) in Focus position 3 when ***BOAT*** was in focus (see underlined bold entries in Table 4.11 below). The presence of a terminal rise, or a more reduced or suspended fall or even a very striking fall from *ing* to ***BOAT*** observed for some talkers in Focus position 3 sentences might have contributed to the perception of focus on the target focus word ***BOAT*** in individual tokens of Focus position 3 sentences. A rise in median F_0 on ***BOAT*** observed for three of the NH talkers only occurred for four of the CI talkers, and the rest of the talkers had a fall. The t test shows that the CI group as a whole showed a significant fall in F_0 on ***BOAT***.

As ***BOAT*** was at the end of the sentence most CI talkers may have found it easier to produce a fall where F_0 was declining anyway. The reduced fall in F_0 for some talkers on the target focus word may have been an attempt to suspend the natural decline of F_0 to convey focus. On the other hand for the group of CI talkers the fall in F_0 at the end of a sentence with a natural declination had a weaker significance level than the fall from ***BOY*** at the start of the sentence from a higher F_0 starting point.

Focus position 1						**Focus position 3**				
	paint	*boat*	*paint:boat*			*paint*	***BOAT***	*paint:* ***BOAT***		
CI Talkers	**Hz**	**Hz**	**diff in Hz**	**diff in semit.**	**F_0 contour**	**Hz**	**Hz**	**diff in Hz**	**diff in semit.**	**F_0 contour**
1	145	60	**85**	**15.28**	***fall***	185	160	**25**	**2.51**	***fall***
2	215	150	**65**	**6.23**	***fal[***	215	190	**25**	**2.14**	***fall***
3	235	195	**40**	**3.23**	***fall***	275	245	**30**	**2.00**	***fall***
4	240	220	20	1.51	***fall***	235	255	-20	-1.41	rise
6	255	225	30	2.17	***fall***	290	240	50	3.28	***fall***
7	100	88	**12**	**2.21**	***fall***	107	103	**4**	**0.66**	***fall***
8	140	95	45	6.71	***fall***	135	145	-10	-1.24	rise
9	213	217	-4	-0.32	rise	220	205	15	1.22	***fall***
10	92	92	0	0.00	level	202	193	9	0.79	***fall***
11	255	205	50	3.78	***fall***	280	290	-10	-0.61	rise
12	227	223	4	0.31	***fall***	252	240	12	0.84	***fall***
13	190	220	-30	-2.54	rise	215	230	-15	-1.17	rise
14	200	210	-10	-0.84	rise	240	205	35	2.73	***fall***
15	250	225	**25**	**1.82**	***fall***	280	255	**25**	**1.62**	***fall***
16	230	230	0	0.00	level	270	260	10	0.65	***fall***
17	128	118	10	1.41	***fall***	145	130	15	1.89	***fall***
mean			21.4	2.6				12.5	1.0	
var			897	17				374	2	
t			2.85	2.45				2.58	2.7	
df			15	15				15	15	
sig			***0.0060***	***0.0135***				***0.0103***	***0.0081***	

Table 4.11 *Differences in the median F_0 in Hz and semitones for paint:* ***BOAT*** *(Focus position 3) and paint: boat (Focus position 1) for the CI talkers.*

4.3.3 F_0 WITHIN and ACROSS sentences: Summary and conclusion

Table 4.12 below summarizes for all CI talkers the number of tokens with F_0 contours *WITHIN* sentences for each of the target focus sentences Focus position 1 (***BOY***), Focus position 2 (***PAINT***) and Focus position 3 (***BOAT***) which approximated the NH talkers and were considered appropriate in the line graphs in Figure 4.1.

In Focus position 1 (***BOY)*** sentences in Table 4.1 five CI talkers (C1, C11, C12, C14, C15) were considered appropriate in all five tokens if there was there was a fall in F_0 followed by a decline or leveling of F_0 in the post-focus syllables. In Focus position 2 (***PAINT***) sentences in Table 4.2 seven talkers (C1, C3, C6, C8, C11, C13, C15) were considered appropriate in all five tokens (and C12, C16 in four out of five tokens) if there was a rise-fall in F_0 or a high F_0 on *boy* and ***PAINT*** followed by a fall. In Focus position 3 (***BOAT)*** sentences in Table 4.3 four talkers (C1, C4, C8, C13) were considered appropriate in all five tokens and three talkers (C9, C11, C14) in four out of five tokens if they had a terminal rise, or a suspended fall, or striking fall in F_0 on ***BOAT***.

Overall, however, only three of the CI talkers (C1, C11, C14) were significantly above chance (0.75 or 0.76[3]) in the production of appropriate F_0 contours in the three target focus words in Table 4.12 (see bold entries in the column under proportion correct).

CI Talker	**BOY (n = 5)**	**PAINT (n = 5)**	**BOAT (n = 5)**	**Total Appropriate**	**Total tokens**	**Proportion correct**
1	5	5	5	**15**	15	**1.00**
2	0	1	3	4	15	0.27
3	3	5	3	11	15	0.73
4	0	2	5	7	15	0.47
6	*3	5	1	9	14	0.64
7	2	2	1	5	15	0.33
8	0	5	5	10	15	0.67
9	2	3	4	9	15	0.60
10	2	3	2	7	15	0.47
11	5	5	4	**14**	15	**0.93**
12	5	4	1	10	15	0.67
13	1	5	5	11	15	0.73
14	5	3	4	**12**	15	**0.80**
15	5	5	1	11	15	0.73
16	*0	4	2	6	14	0.43
17	3	3	3	9	15	0.60
*** n = 4**	***Significant at 0.75 for 12 appropriate out of maximum of 15 and 0.76 for 11 out of 14***					

Table 4.12 *Summary of appropriate F_0 contours in Focus position 1, Focus position 2 and Focus position 3 sentences. All talkers had a maximum of 5 tokens except for C6* and C16* who had four.*

NH and CI talkers (except for C2 and C7) in Table 4.13 had a similar range in the fall in median F_0 on the target focus word ***BOY*** in Focus position 1 sentences which was significant ($p<0.0005$) for the group of CI talkers (see Table 4.5). However F_0

---

[3] Assuming a sig. proportion correct at 0.05 level i.e. 12 of 15 trials (0.75) or 11 of 14 trials (0.76)

measurements *ACROSS* sentences show that only four CI talkers had a greater fall when *boy* was in focus than when it was not in focus.

Although there was a rise - fall in F_0 on ***PAINT*** for both groups in Focus position 2 sentences which was statistically significant for the CI group, Table 4.13 shows that the CI talkers as a group had a bigger median F_0 range than NH talkers on the target focus word ***PAINT***. However, the rise in F_0 to ***PAINT*** in Table 4.7 for five CI talkers (i.e. C1, C3, C13, C16, C17) and the fall from ***PAINT*** in Table 4.9 for eight CI talkers (C3, C6, C7, C8, C11, C12, C15, C17) was greater when *paint* was in focus. This suggests a possible trend in the data for a greater rise and fall in F_0 on the target focus word.

In Focus position 3 sentences only four CI talkers (C4, C8, C11, C13) resembled the three NH talkers with a terminal rise in F_0 on ***BOAT*** in Table 4.11, whereas twelve CI talkers had a fall and the group as a whole showed a significant fall. However, five of the twelve talkers (C1, C2, C3, C7, C15) had a more reduced fall in F_0 (only slightly for C15) than when ***BOAT*** was in focus.

Target words	**NH**	**F_0 contour**	**CI**	**F_0 contour**
BOY: paint	1.82 – 6.51 semit	fall	0.36 – 4.54 semit.	fall (14 talkers)
boy: ***PAINT***	0.82 – 3.47 semit.	rise	1.86 - 12.37 semit	rise (15 talkers)
PAINT : boat	6.94 – 9.19 semit.	fall	.36 – 13.93 semit.	fall (16 talkers)
paint: ***BOAT***	.73 – 2.67 semit.	rise	.61 – 1.41 semit.	rise (4 talkers)
			.65 - .3.28 semit.	fall (12 talkers)

Table 4.13 *The range of median F_0 differences between the target focus words **BOY, PAINT** and **BOAT** and their neighbouring words for the NH and CI groups in Experiment III.*

4.3.4 Word durations

Word durations for the NH talkers are presented in the line graphs in Figures 4.6 and in the boxplots in Figure 4.7 and in Table 4.14. Durations for the CI talkers are presented in the line graphs in Figures 4.8, in the boxplots in Figure 4.9, and in Table 4.19. To eliminate inherent word durations differences the data have been normalized for each word and talker and the values presented show the ratio of the word durations relative to the average (which is always expressed as 1.0). As discussed earlier in section 4.2.3.2 duration measurements are presented for entire target words *boy, painting* and *boat* in three tokens of Focus position 1 ***(BOY),*** Focus position 2 ***(PAINTing***) and Focus position 3 ***(BOAT)*** for the four NH talkers and in five tokens for the CI talkers.

Tables 4.15 – 4.18 summarise the number of tokens with appropriate lengthening of the focus words ***BOY, PAINTing*** and ***BOAT*** for individual CI talkers. Durations which were longer than the average for that word were considered appropriate, and durations which were the same or shorter than the average were considered inappropriate.

4.3.4.1 Durations of target focus words BOY, PAINTing, BOAT

NH talkers

The line graphs in Figure 4.6 for NH individual tokens show that in all three tokens ***BOY*** and ***PAINTing*** were longer than the average for these words. However a few individual tokens were only slightly longer than average (e.g. N1:T1 for ***BOY*** and N4:T3 for ***PAINTing***. There were also some individual ***BOAT*** tokens where durations were shorter than average for some talkers (e.g. N1:T3 and N4:T3). The boxplots in Figure 4.7 show for the group of NH talkers that the median durations of the three focus words ***BOY***, ***PAINTing*** and ***BOAT*** were longer than the average for each focus word.

Median duration values in Table 4.14 also show that for the four individual NH talkers the three target focus words were longer than the average duration for these words. Mean increases in duration for the group were 1.25 secs, 1.18 secs. and 1.18 secs. for ***BOY, PAINTing*** and ***BOAT*** respectively.

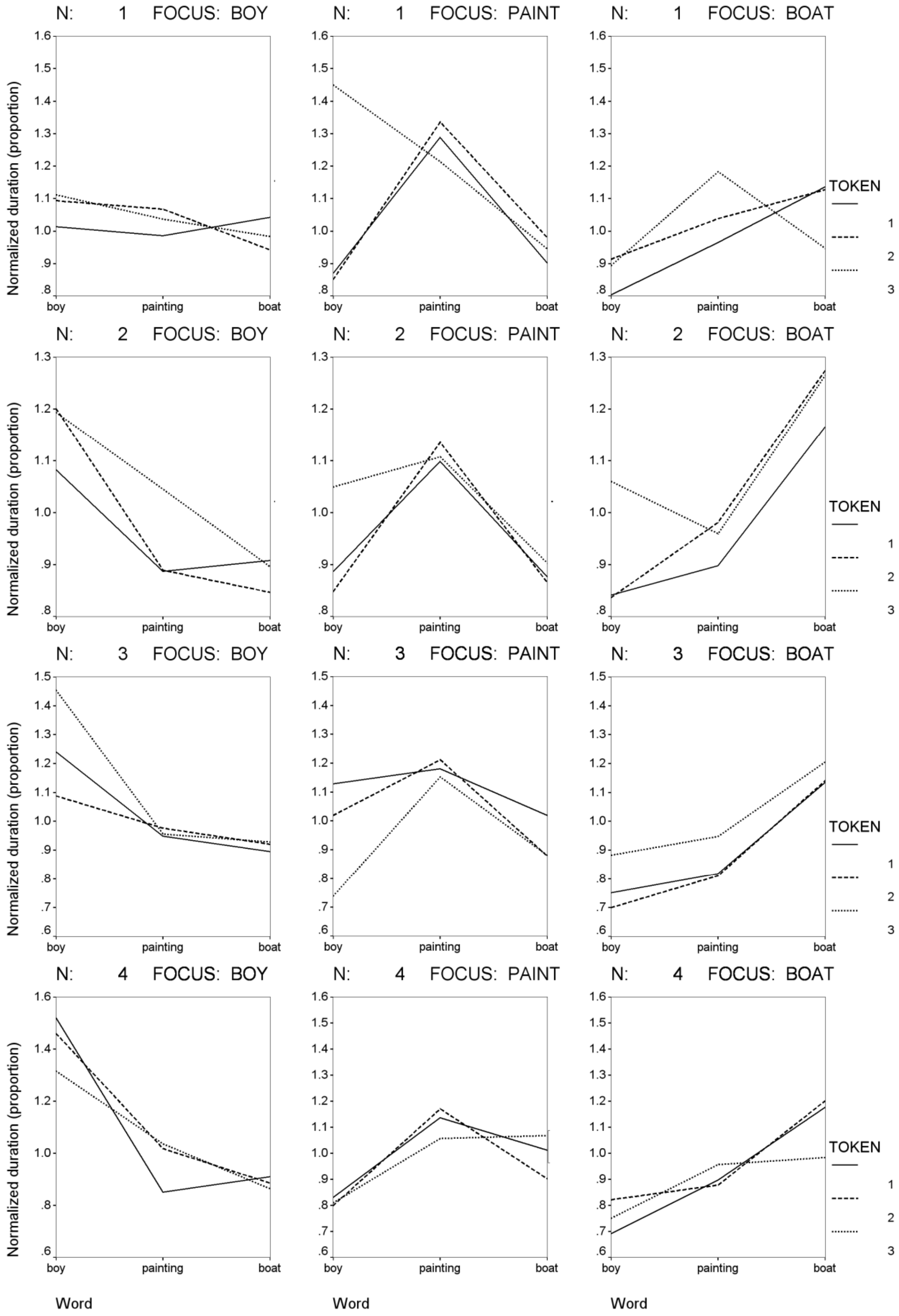


Figure 4.6 *Line graphs for the NH talkers showing mean duration for the target words* ***boy, paint(ing)*** *and* ***boat*** *in Focus position 1, Focus position 2 and Focus position 3 sentences.*

NH Talker	Focus position 1	Focus position 2	Focus position 3
	BOY	PAINTing	BOAT
	secs	secs	secs
1	1.09	1.29	1.13
2	1.19	1.11	1.26
3	1.24	1.18	1.14
4	1.46	1.14	1.18
mean	*1.25*	*1.18*	*1.18*

Table 4.14 *Ratios of word durations for **BOY, PAINTing** and **BOAT** relative to the average for these words for individual NH talkers in Focus position 1, Focus position 2 and Focus position 3 sentences.*

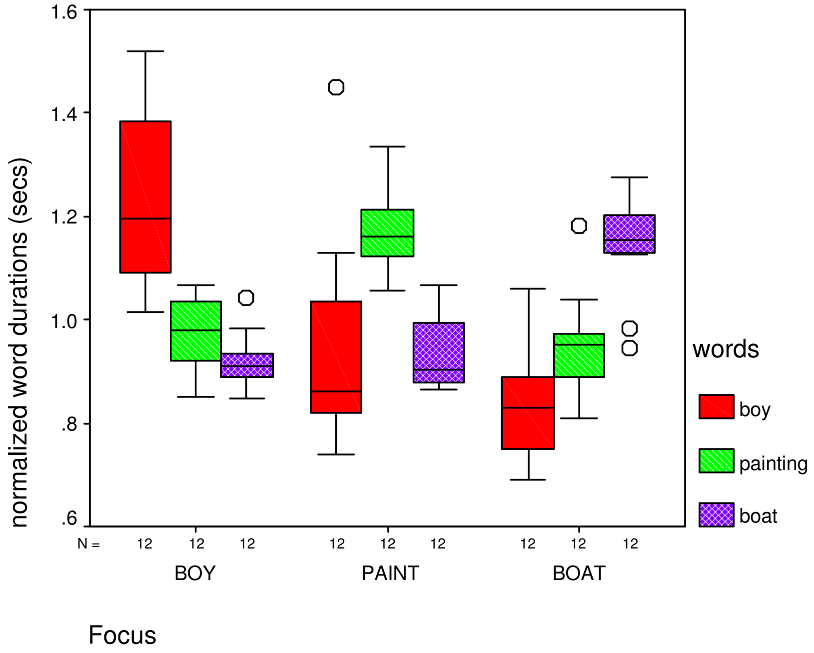


Figure 4.7 *Box and whisker plot of normalised word durations for each word and focus target for the NH talkers***.**

CI talkers

BOY

The line graphs in Figure 4.8 and Table 4.15 show that only four CI talkers (C8, C10, C12, C14) increased the duration relative to the average for ***BOY*** in all five tokens, and five other talkers (C3, C4, C7, C13, C15) had appropriate lengthening in four out of five tokens. All talkers had a maximum of five tokens except for C6* and C16* who had four.

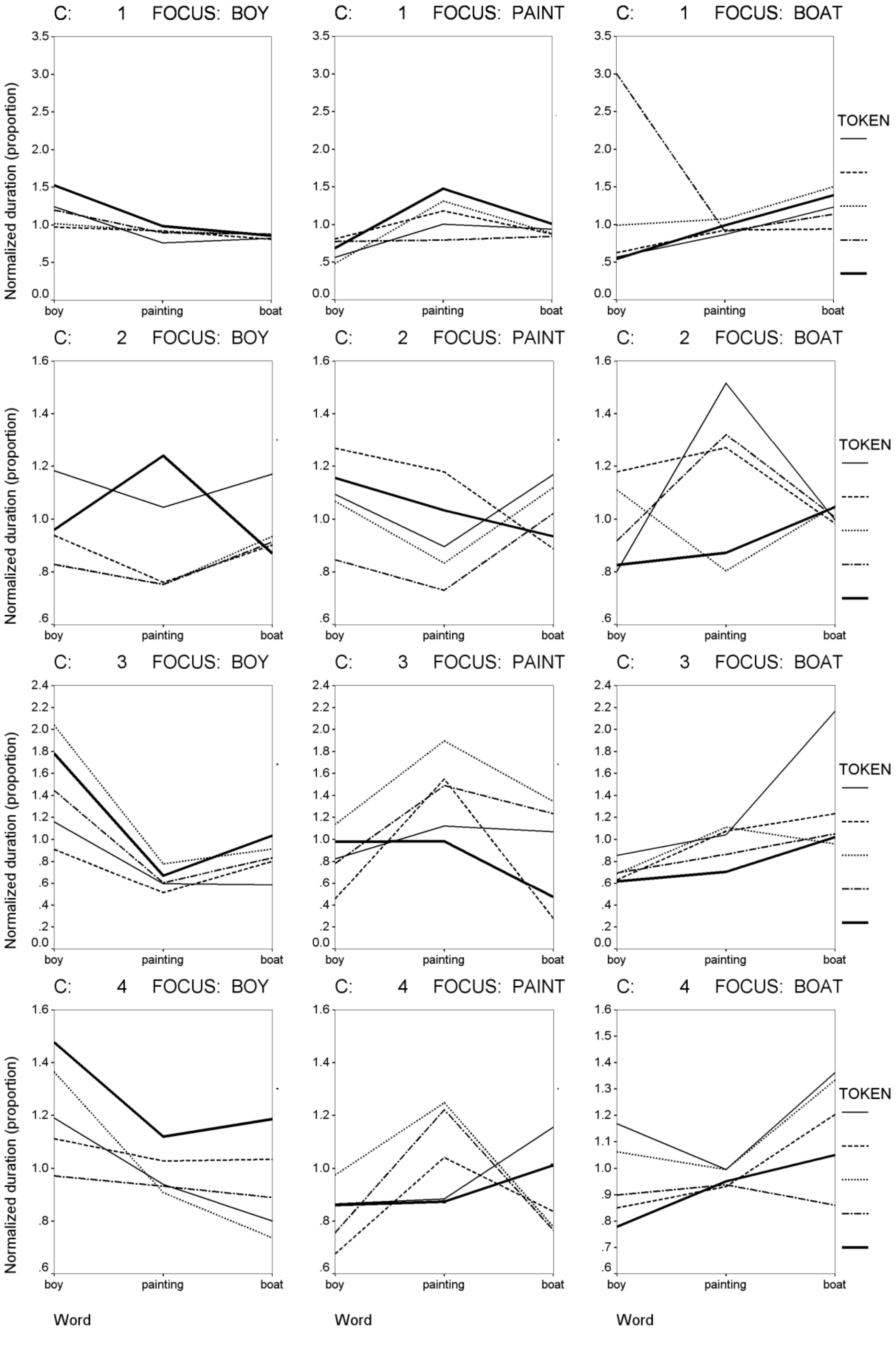


Figure 4.8 *Line graphs for the CI talkers showing mean durations for the target words boy, paint(ing) and boat in Focus position 1, Focus position 2, and Focus position 3 sentences. Individual tokens (1-5) are represented by different lines styles as indicated in the margin the right of the figure for each talker.*

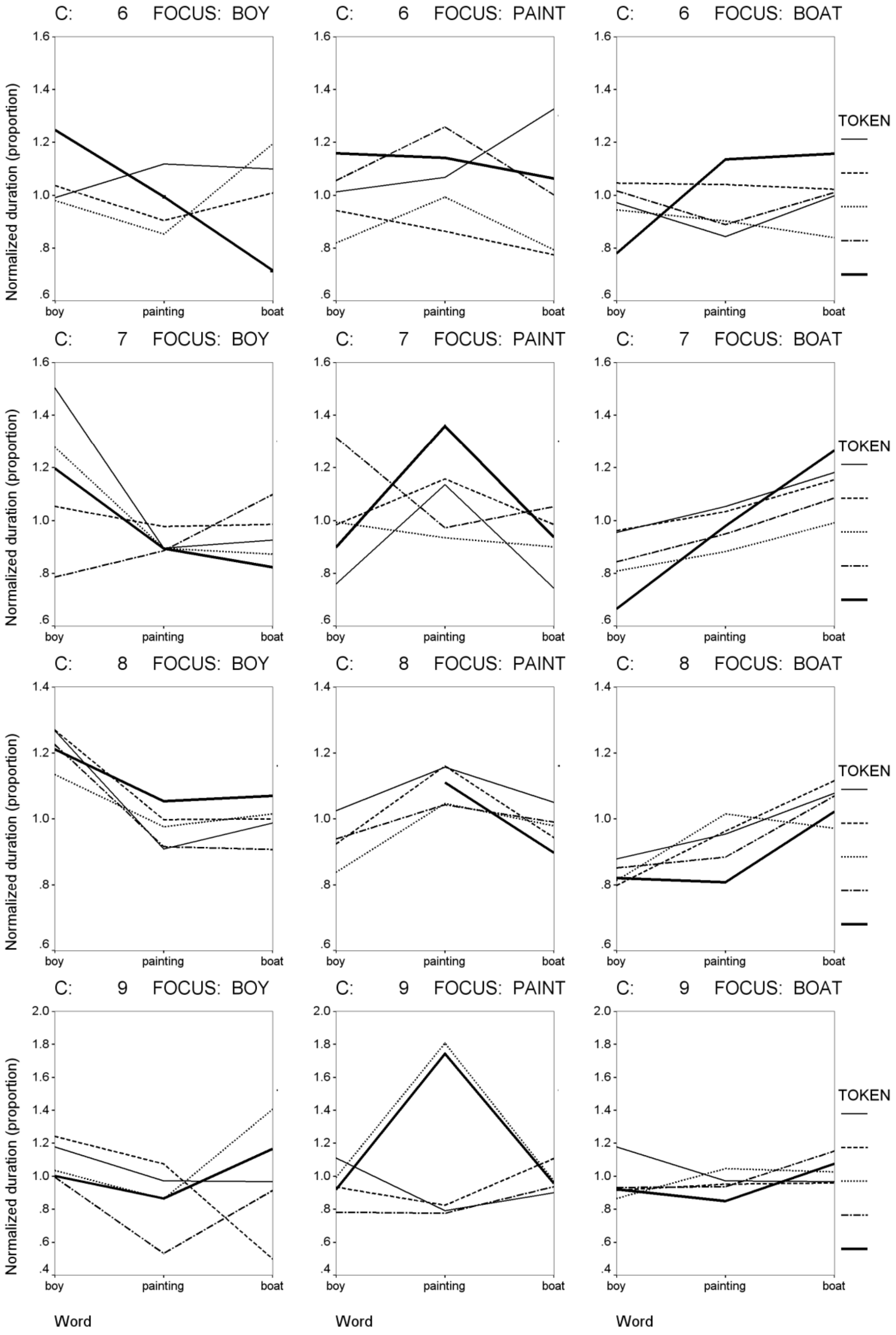


Figure 4.8 *(Continued)*

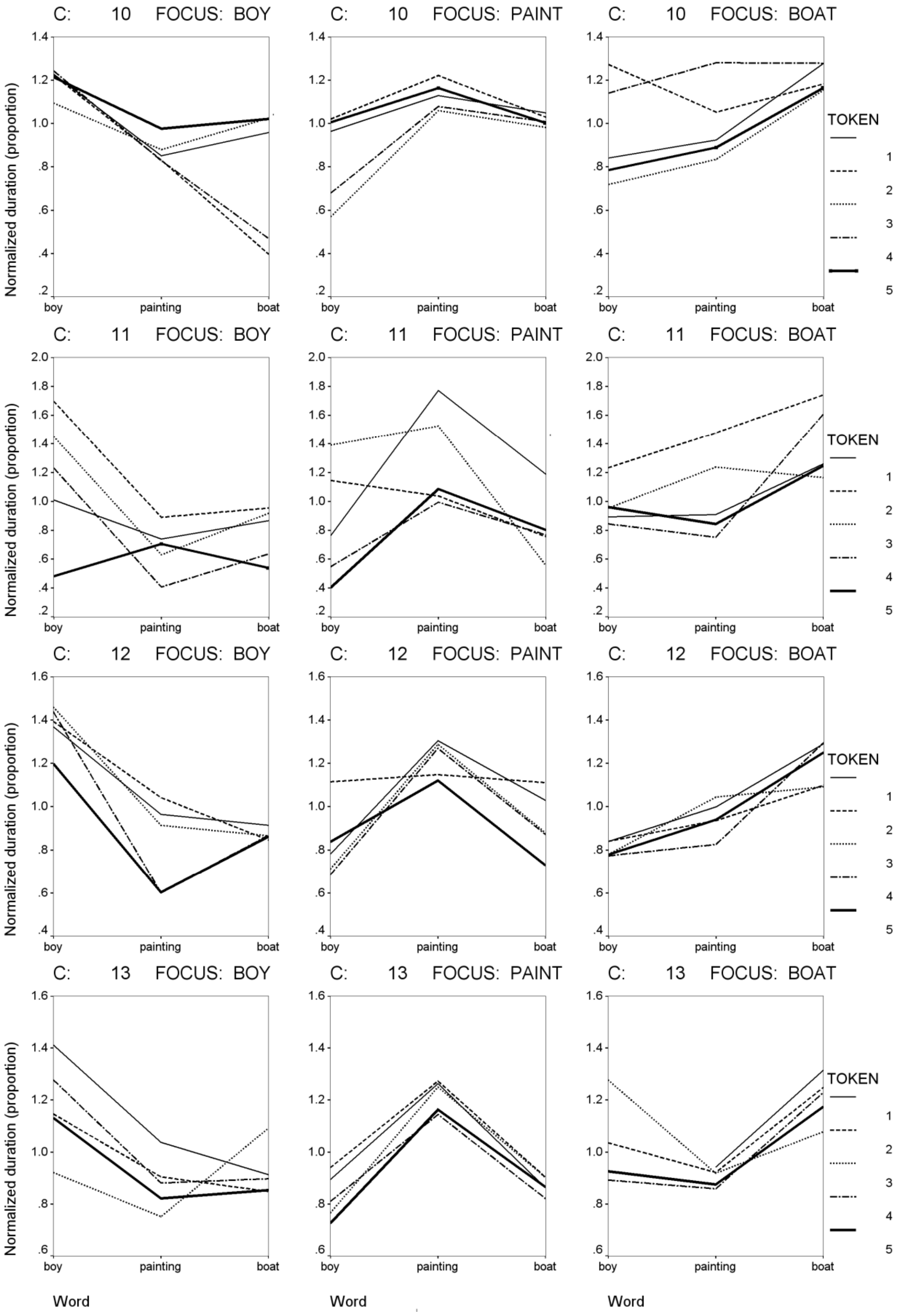


Figure 4.8 *(Continued)*

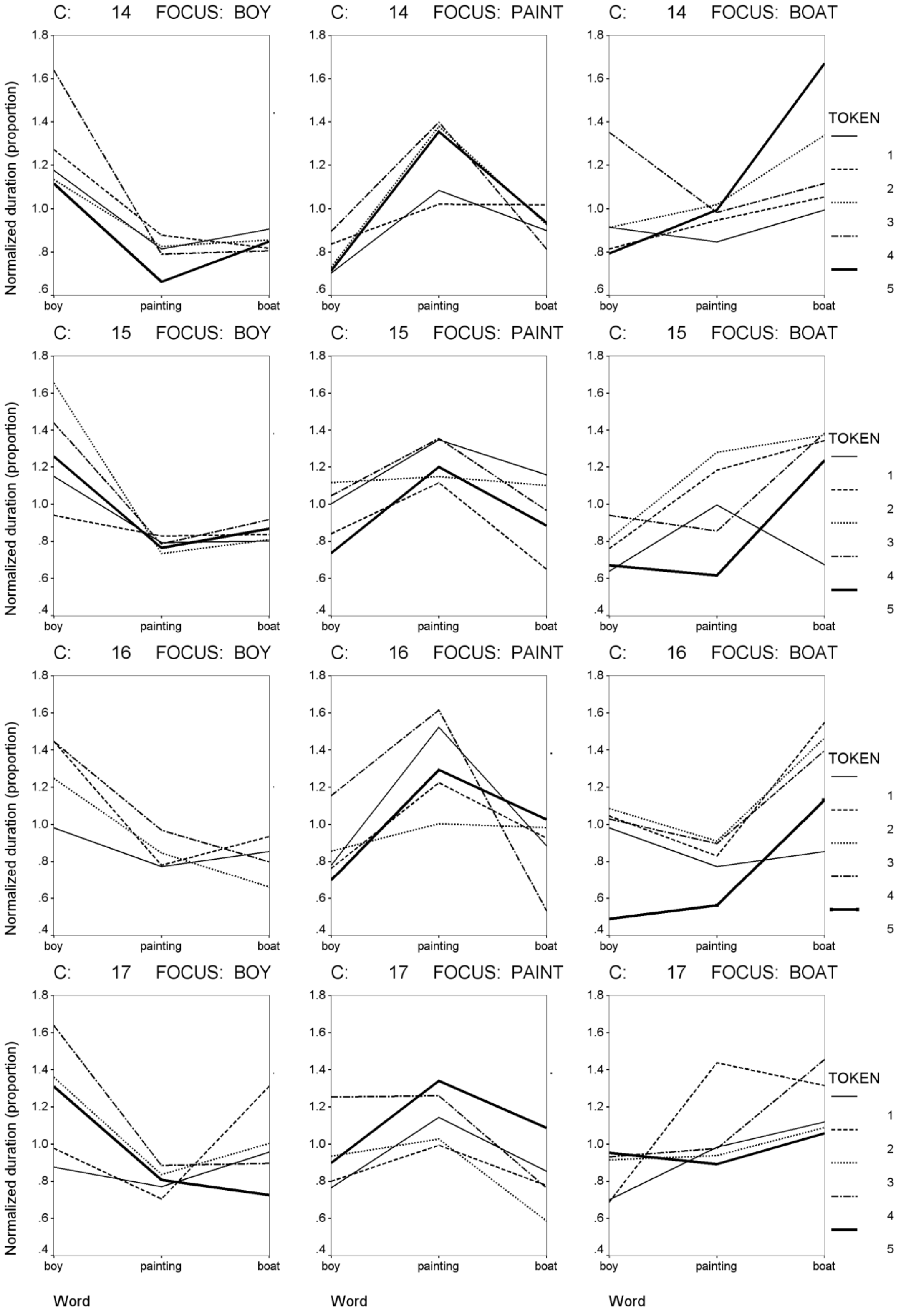


Figure 4.8 *(Continued)*

BOY	**Duration relative to Average (1.0)**				
Talker	**longer than average**	**same as average**	**shorter than average**	**number of appropriate tokens**	**Total tokens**
1	T1; T4;T5	T3	T2	3	5
2	T1		T2;T3;T4;T5	1	5
3	T1;T3;T4;T5		T2	4	5
4	T1;T2;T3;T5		T4	4	5
*6	T2;T5		T1;T3	2	4
7	T1;T2;T3;T5		T4	4	5
8	T1;T2;T3;T;T5			**5**	5
9	T1;T2;T3	T4;T5		3	5
10	T1;T2;T3;T4;T5			**5**	5
11	T2;T3;T4	T1	T5	3	5
12	T1;T2;T3;T4;T5			**5**	5
13	T1;T2;T4;T5		T3	4	5
14	T1;T2;T3;T4;T5			**5**	5
15	T1;T3;T4;T5		T2	4	5
*16	T2;T3;T5		T1	3	4
17	T3;T4;T5		T1;T2	3	5

Table 4.15 *Duration details in individual tokens for Focus position 1 (**BOY**) sentences in the line graphs for the CI talker.*

PAINTing

In Table 4.16 only five talkers (C8, C12, C13, C14, C15) consistently lengthened the focus word ***PAINTing*** relative to the average for that word, and four other talkers (C3, C11, C16, C17) lengthened ***PAINTing*** in four out of five tokens.

PAINTing	**Durations relative to the Average (1.0)**				
Talker	**longer than average**	**same as average**	**shorter than average**	**number of appropriate tokens**	**Total tokens**
1	T2; T3;T4		T1; T2	3	5
2	T2;T5		T1;T3;T4	2	5
3	T1;T2;T3;T4	T5		4	5
4	T2;T3;T4		T1;T5	3	5
6	T1;T4;T5		T2;T3	3	5
7	T1;T2;T5		T4;T3	3	5
8	T1;T2;T3;T4;T5			**5**	5
9	T3;T5		T1;T2;T4	2	5
10	T1;T2;T3;T4;T5			5	5
11	T1; T2; T3; T5	T4		4	5
12	T1;T2;T3;T4;T5			**5**	5
13	T1;T2;T3;T4;T5			**5**	5
14	T1;T2;T3;T4;T5			**5**	5
15	T1;T2;T3;T4;T5			**5**	5
16	T1;T2;T3;T5	T4		4	5
17	T1;T3;T4;T5	T2		4	5

Table 4.16 *Duration details in individual tokens for Focus position 2* ***(PAINT)*** *sentences in the line graphs for the CI talkers.*

BOAT

Table 4.17 below shows that only five talkers (C10, C11, C12, C13,C17) consistently lengthened the focus word ***BOAT*** relative to the average (1.0) for that word in five tokens, and seven other talkers (C1, C4, C8, C14, C15, C16) in four out of five tokens.

BOAT	**Durations relative to the Average (1.0)**				
Talker	**longer than average**	**same as average**	**shorter than average**	**number of appropriate tokens**	**Total tokens**
1	T1;T3;T4;T5		T2	4	5
2	T3;T5	T4;T2;T1		2	5
3	T1;T2;T4	T5	T3	3	5
4	T1;T3;T2;T5		T4	4	5
6	T2;T4;T5	T1	T3	3	5
7	T1;T2;T4;T5		T5	4	5
8	T1;T2;T4;T5		T3	4	5
9	T3;T4;T5		T1;T2	3	5
10	T1;T2;T3;T4;T5			**5**	5
11	T1;T2;T3;T4;T5			**5**	5
12	T1;T2;T3;T4;T5			**5**	5
13	T1;T2;T3;T4;T5			**5**	5
14	T2;T3;T4;T5	T1		4	5
15	T2;T3;T4;T5		T1	4	5
16	T2;T3;T4;T5		T1	4	5
17	T1;T2;T3;T4;T5			**5**	5

Table 4.17 *Duration details in individual tokens) for Focus position 1* ***(BOAT)*** *sentences in the line graphs for the CI talkers.*

Table 4.18 below summarises for individual talkers the number of tokens where appropriate increases in duration occurred in the production of focus in the three target focus words ***BOY***, ***PAINTing*** and ***BOAT.*** Nine talkers (C8, C10, C11, C12, C14, C15, C16, C17) significantly lengthened the target focus words in the production of appropriate duration (see bold entries).

Talker	BOY (n = 5)	PAINTing (n = 5)	BOAT (n = 5)	Total Appropriate	Total tokens	Proportion correct
1	3	3	4	10	15	0.67
2	1	2	2	5	15	0.33
3	4	4	3	11	15	0.73
4	4	3	4	11	15	0.73
6	*2	3	3	8	14	0.57
7	4	3	4	11	15	0.73
8	5	5	4	**14**	15	**0.93**
9	3	2	3	8	15	0.53
10	5	5	5	**15**	15	**1.00**
11	3	4	5	**12**	15	**0.80**
12	5	5	5	**15**	15	**1.00**
13	4	5	5	**14**	15	**0.93**
14	5	5	4	**14**	15	**0.93**
15	4	5	4	**13**	15	**0.87**
16	*3	4	4	**11**	14	**0.79**
17	3	4	5	**12**	15	**0.80**
*** n = 4**	***Significant at 0.75 for 12 appropriate out of maximum of 15 and 0.76 for 11 out of 14***					

Table 4.18 *Summary of appropriate durational increases in the target focus words* ***BOY, PAINTing,*** *and* ***BOAT*** *for the CI talkers.*

The boxplots for the group of CI talkers in Figure 4.9 show that the median durations of the target focus words ***BOY***, ***PAINTing*** and ***BOAT*** were longer than average (1.0) for the group of CI talkers.

Table 4.19 also shows that for most individual CI talkers the median duration of the target focus words were increased relative to the average duration for each word. Exceptions to this are C2 for ***BOY*** in Focus position 1 and C2 and C9 for ***PAINT*** in Focus position 2 (see underlined entries). T tests carried out for the whole group of CI talkers and shown in Table 4.19 indicate significant lengthening for each of ***BOY***, ***PAINT***, and **BOAT** when in focus.

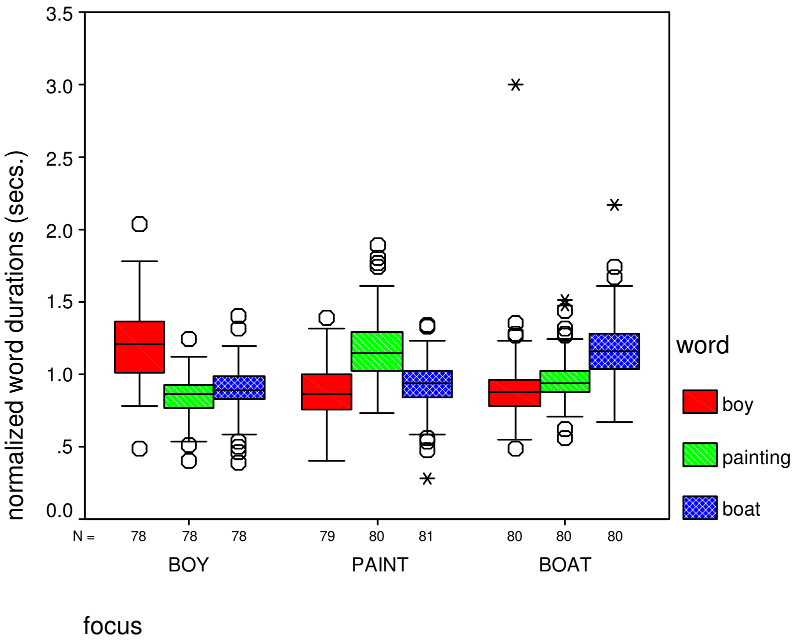


Figure 4.9 *Box and whisker plot of normalised word durations for each word and focus target for the CI group.*

CI Talker	**Focus position 1**	**Focus position 2**	**Focus position 3**
	BOY	**PAINT**	**BOAT**
	secs	secs	secs
C1	1.20	1.18	1.23
C2	0.94	0.90	1.01
C3	1.45	1.49	1.05
C4	1.19	1.04	1.20
C6	1.01	1.07	1.01
C7	1.20	1.13	1.15
C8	1.23	1.11	1.07
C9	1.04	0.82	1.03
C10	1.22	1.13	1.18
C11	1.23	1.09	1.26
C12	1.39	1.27	1.25
C13	1.15	1.25	1.23
C14	1.18	1.36	1.12
C15	1.26	1.20	1.34
C16	1.35	1.29	1.40
C17	1.31	1.14	1.12
mean	1.2	1.1546	1.1659
var	0.0180	0.0268	0.0139
t	.6.05	3.65	5.5
df	15	15	15
sig	**<0.0001**	**0.0012**	**<0.0001**

Table 4.19 *Median duration of the target focus words* ***BOY, PAINTing, BOAT*** *for individual CI talkers are presented above.*

4.3.4.2 Duration summary

NH Talkers

Median duration measurements in Table 4.14 and individual tokens in the line graphs (Figure 4.6) show that the target focus words ***BOY, PAINTing*** and ***BOAT*** were lengthened relative to the average duration (1.0) for most NH talkers. Exceptions to this were individual ***BOAT*** tokens N1;T3, and N4;T3. Mean durations in Table 4.14 and the median values in the boxplots (Figure 4.7) show that as a group the NH talkers lengthened the target focus words.

CI talkers

The line graphs in Figure 4.8 for the CI talkers three target focus words ***BOY, PAINTing*** and ***BOAT*** show that some individual tokens were the same duration as, longer or shorter than average (expressed as 1.0) for these words. Only those longer than average were considered appropriate and they are summarised in Table 4.18. Overall, nine CI talkers (C8, C10, C11, C12, C13, C14, C15, C16, C17) significantly lengthened the three focus words ***BOY***, ***PAINTing*** and ***BOAT***. Two of these talkers (C10 and C12) had appropriate lengthening in all tokens for the three target focus words.

Median duration measurements in Table 4.19 for individual CI talkers show increased lengthening of the target focus words for all except for C2 (***BOY***) and C2 and C9 (***PAINT***). However as a group, median durations of the focus words in the boxplots in figure 4.9 were longer than average and also t tests for the group of CI talkers show significantly lengthening of ***BOY***, ***PAINT*** and ***BOAT***. In summary, only nine CI talkers had appropriate lengthening of the target focus in the individual line graphs like most of the NH talkers. However, median duration measurements show increased lengthening for all except one talker in ***BOY*** sentences and two talkers in ***PAINT*** sentences. T tests show that as a group the CI talkers significantly lengthened ***BOY, PAINT*** and ***BOAT.***

4.3.5 Amplitude measurements

4.3.5.1 Amplitude for target focus words BOY, PAINTing, BOAT

To eliminate inherent amplitude of individual words and syllables the data have been normalized for each syllable and talker. The values presented in Tables 4.20 and 4.25

for the NH and CI talkers below are the median amplitudes (dB) relative to the average with the average expressed as 0 dB on the line graphs in Figure 4.10 and 4.12 for each word/syllable for individual tokens and talkers. Boxplots showing the distribution of normalised amplitudes for the NH and CI talkers are presented in Figures 4.11 and 4.13. Tables 4.21, 4.22 and 4.23 show that when amplitude of the target focus words ***BOY***, ***PAINTing*** and ***BOAT*** was above average (0 dB) in individual tokens for each CI talker they were considered appropriate and in the right direction. When amplitude was the same as average or below average, it was not considered appropriate.

NH talkers

The line graphs in Figure 4.10 for Focus position 1 sentences show that for the NH talkers amplitude of the focus word ***BOY***, ***PAINT*** and ***BOAT*** was above average (0 dB) in individual tokens except for N3:T2;T3, N4:T2 ***(BOY***), and N4:T3 ***(PAINT***). Table 4.20 below shows that when ***BOY*** is in focus in Focus position 1 sentences median amplitude was greater than the average (0 dB) but with a mean increase of less than 1 dB for three NH talkers. In Focus position 2 and Focus position 3 sentences the median amplitude is greater than average for all four talkers in ***PAINT*** with a mean increase of 2 dB and in ***BOAT*** with a mean increase of 6 dB. The increase in amplitude on ***BOY*** when less than 1 dB may not have been audible whereas 6 dB increases in amplitude for ***BOAT*** were likely to be more audible.

The boxplots in Figure 4.11 also show that median amplitude in target focus words ***BOY***, ***PAINT*** and ***BOAT*** for the group of NH talkers was greater than the average for each of the focus words but the median increase is much smaller for ***BOY*** (1 dB) and ***PAINT*** (2 dB) than for ***BOAT*** (6 dB)

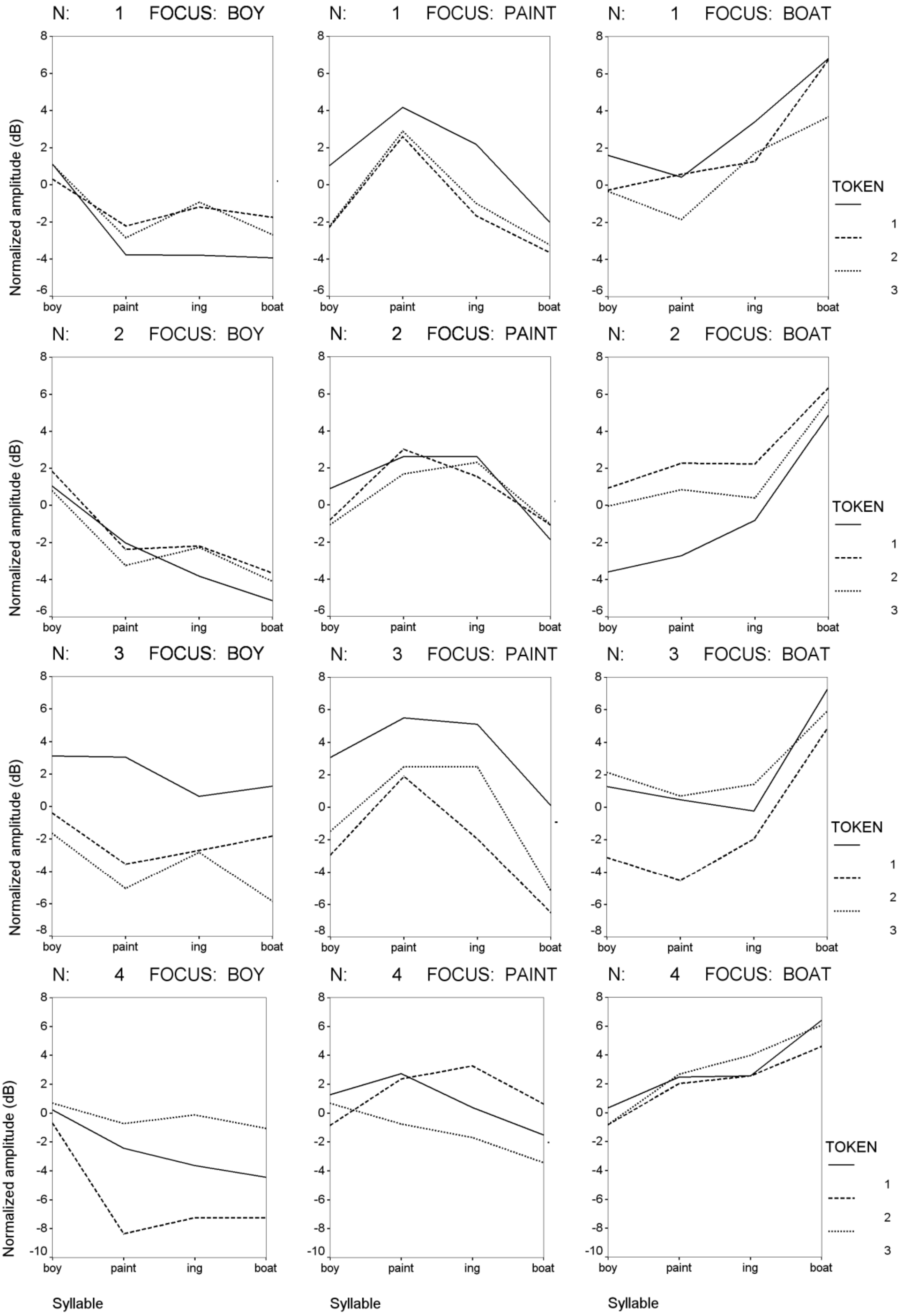


Figure 4.10 *Line graphs for NH talkers showing mean amplitude for the target words* ***boy, paint(ing)*** *and* ***boat*** *in Focus position 1, Focus position 2, and Focus position 3 sentences.*

NH Talker	Focus position 1	Focus position 2	Focus position 3
	BOY	**PAINT**	**BOAT**
	dB	dB	dB
1	1.09	2.89	6.76
2	1.05	2.61	5.68
3	-0.39	1.90	5.92
4	0.22	2.36	6.07
mean	***0.49***	***2.44***	***6.11***

Table 4.20 *Amplitude values (dB) for the NH talkers in the target focus words **BOY, PAINT** and **BOAT** relative to the average amplitude for these words.*

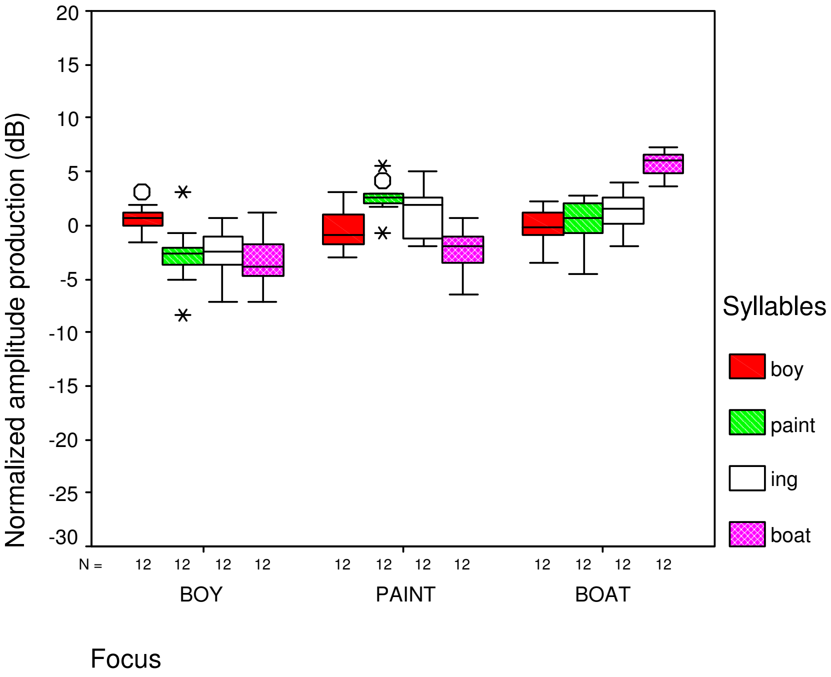


Figure 4.11 *Box and whisker plot of normalised amplitudes for each syllable and focus target for the NH group indicating median increase in amplitude for **BOY** (1 dB), **PAINT** (2 dB), and **BOAT** (6 dB).*

CI talkers

The line graphs in Figure 4.12 show which tokens had appropriate increases in amplitude and were considered appropriate for each talker and focus word. The maximum number of tokens was five except for C6 and C16 who had four. Details of individual tokens for each CI talker are presented below in Tables 4.21 – 4.24.

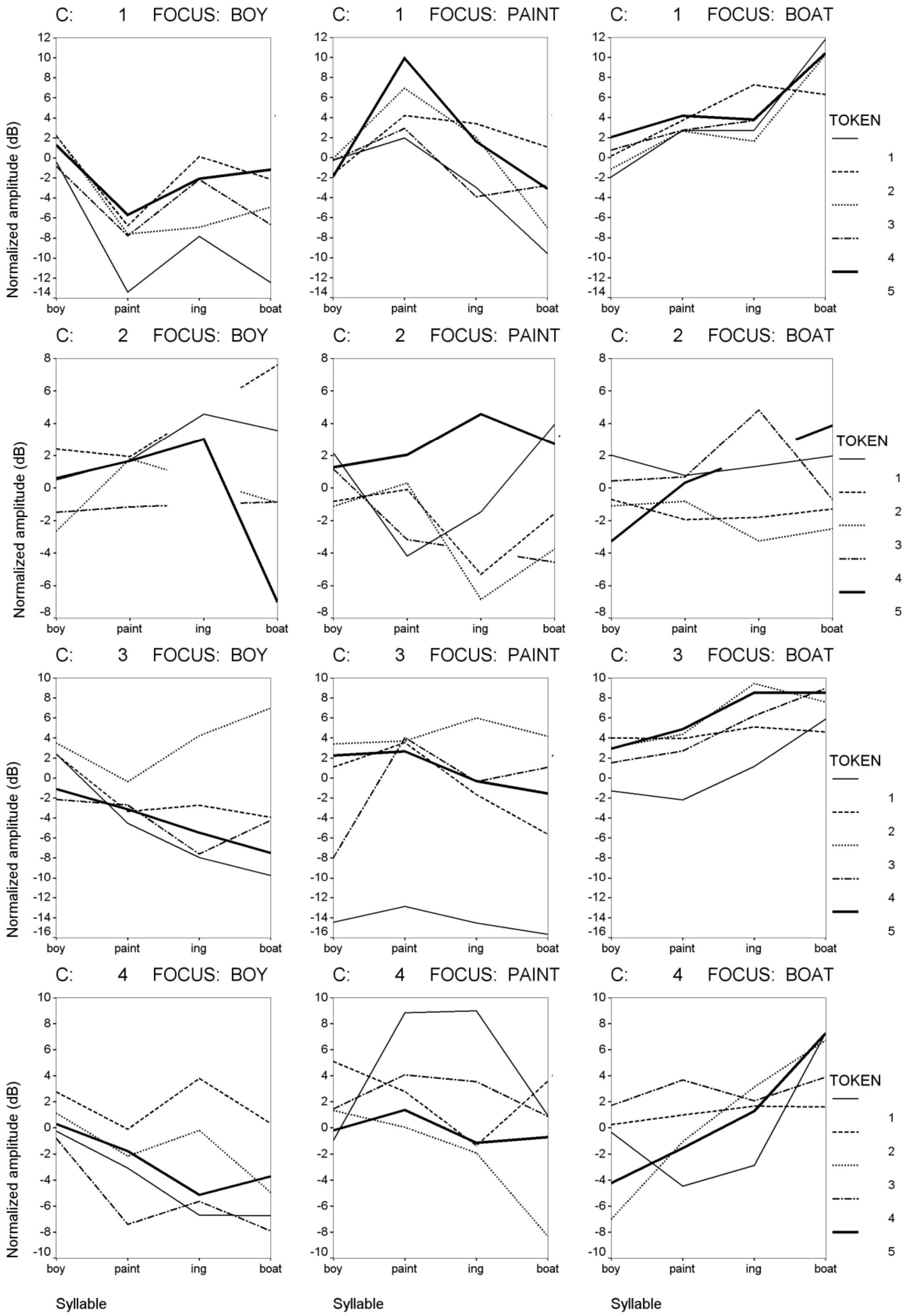


Figure 4.12 *Line graphs for the CI talkers showing mean amplitude for the target words* ***boy****,* ***paint(ing)*** *and* ***boat*** *in Focus position 1, Focus position 2 and Focus position 3 sentences. Individual tokens (1-5) are represented by different lines styles as indicated in the margin the right of the figure for each talker.*

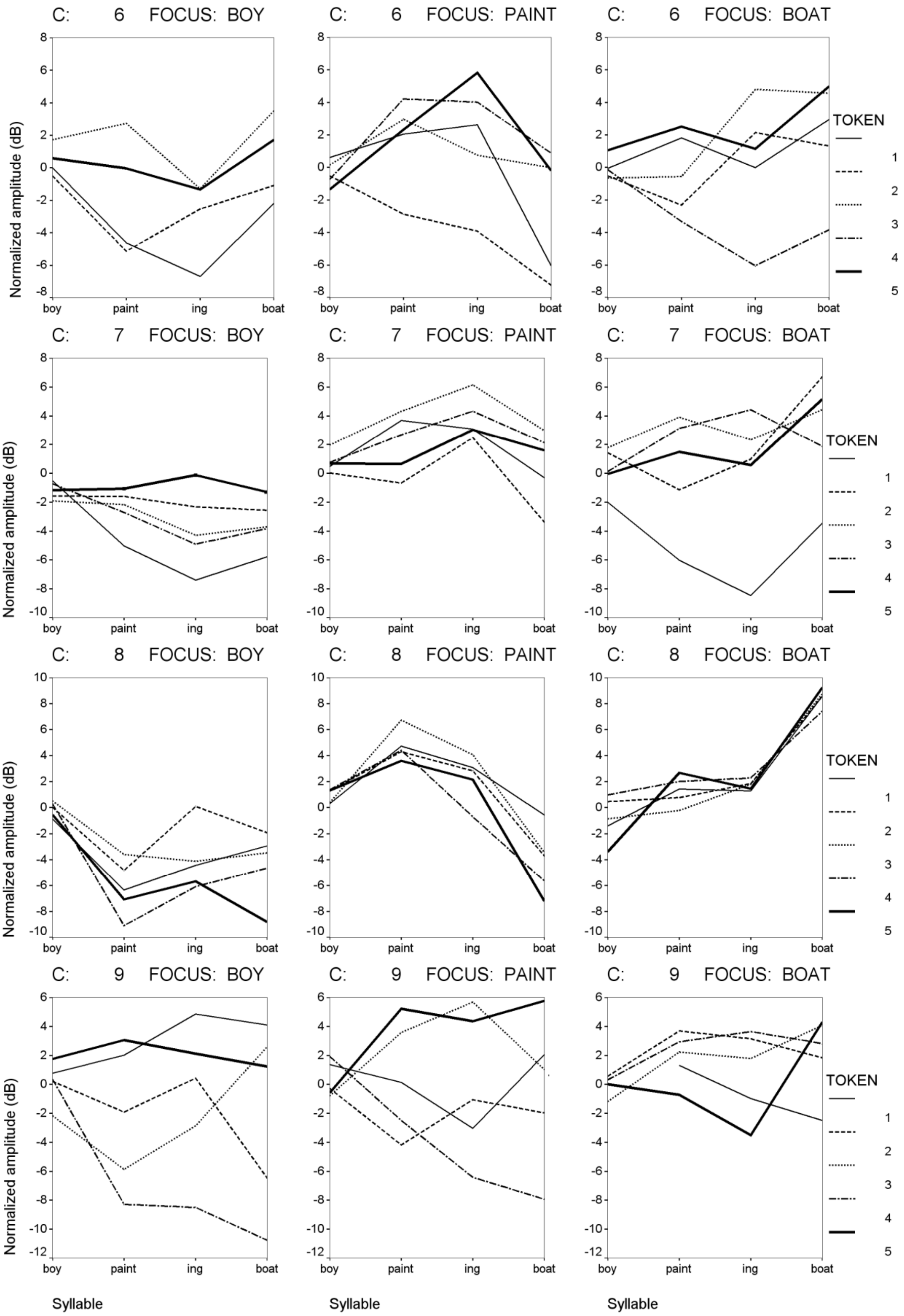


Figure 4.12 *(Continued)*

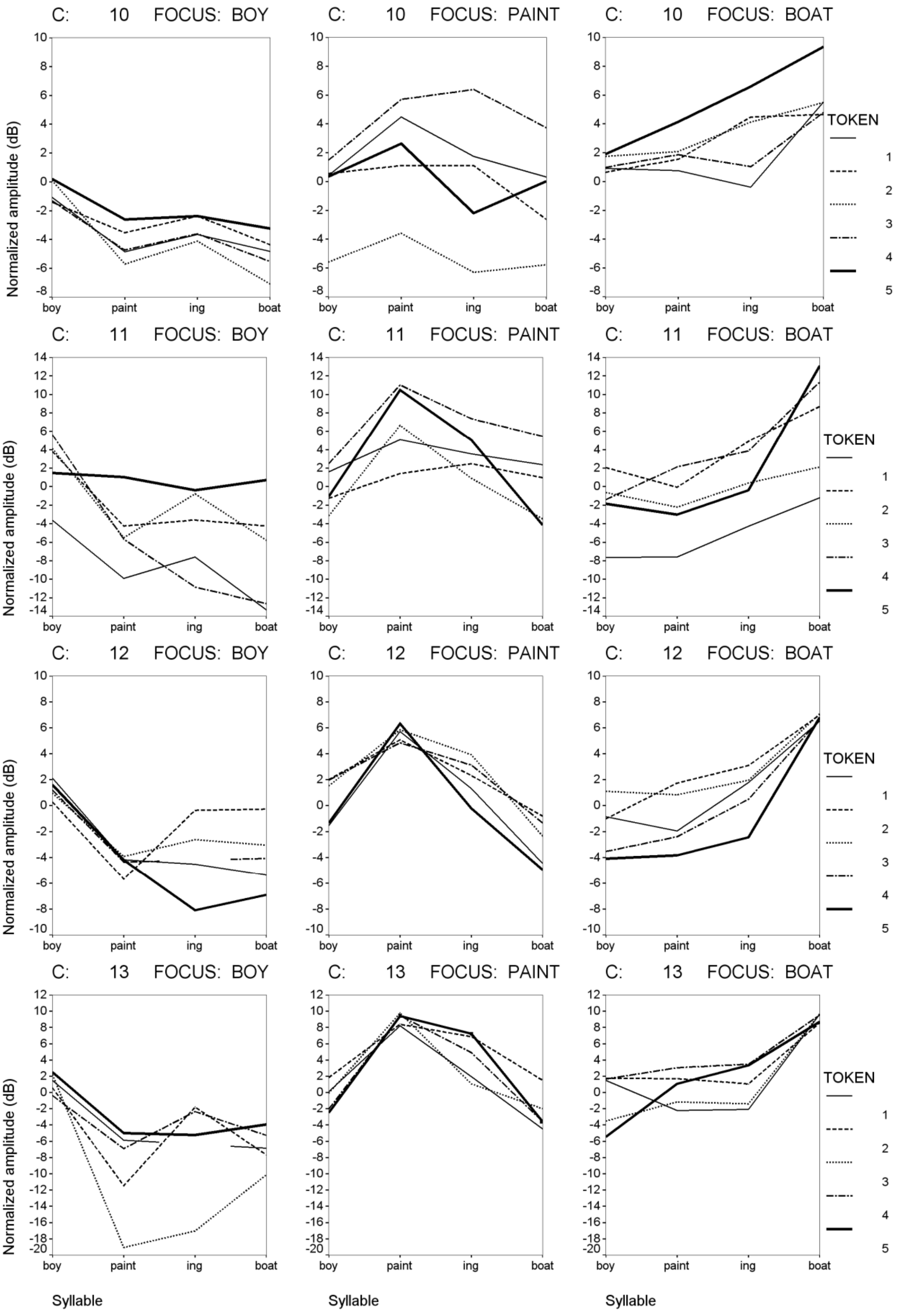


Figure 4.12 *(Continued)*

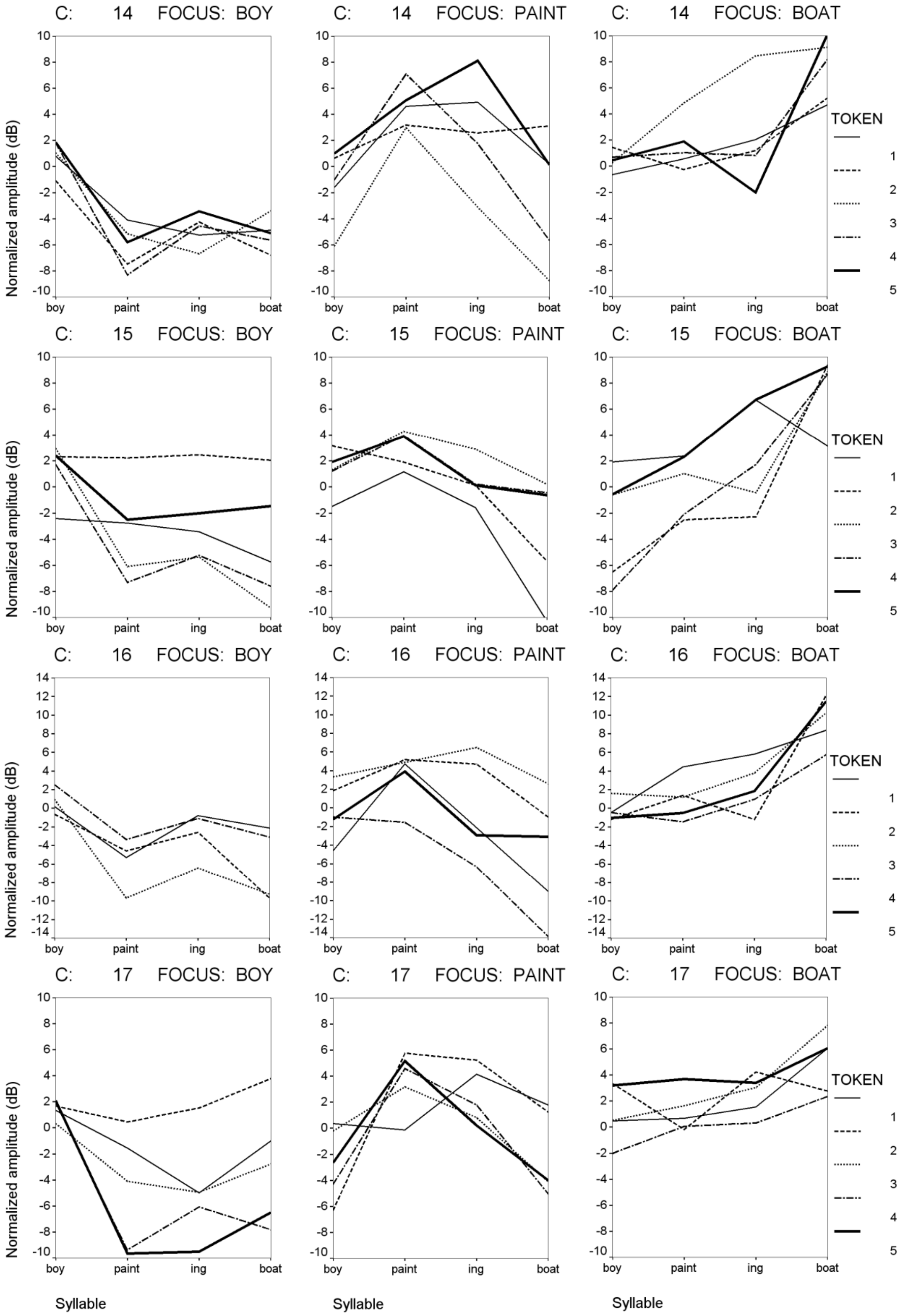


Figure 4.12 *(Continued)*

BOY

Figure 4.12 and Table 4.21 below shows that six talkers increased amplitude of the focus word ***BOY*** (C11, C12, C13, C14, C15, C17) relative to the average (0 dB) for that word in four out of five tokens.

BOY	Above Average Amplitude (0 dB)	Average Amplitude (0 dB)	Below Average Amplitude (0 dB)	Appropriate tokens	Total tokens
1	T2;T3;T4		T1;T4	3	5
2	T1;T2;T5		T3;T4	3	5
3	T3;T1;T2		T5;T4	3	5
4	T2;T3;T5		T1;T4	3	5
*6	T3;T5	T1	T2	2	4
7			T1;T2;T3;T4;T5	0	5
8	T3;T4	T2	T1;T5	2	5
9	T1;T5	T2;T4	T3	2	5
10		T5;T3	T1;T2;T4	0	5
11	T2;T3;T4;T5		T1	**4**	5
12	T1;T3;T4;T5	T2		**4**	5
13	T1;T2;T3;T5	T4		**4**	5
14	T1;T3;T4;T5		T2	**4**	5
15	T2;T3;T4;T5		T1	**4**	5
*16	T4;T3	T1	T2	2	4
17	T1;T2;T5;T4	T3		**4**	5

Table 4.21 *Amplitude details in individual tokens for Focus position1* ***(BOY)*** *sentences in the line graphs for the CI talkers.*

PAINT

Table 4.22 shows that when ***PAINT*** was in focus amplitude was above average for nine CI talkers (C1, C3*, C8, C10*, C11, C12, C13, C14, C15) in all tokens (0 dB) and five talkers (C4, C6, C7, C16, C17) in four out of five tokens. In two tokens* amplitude was unusually low for these talkers and because the rest of the tokens were above average amplitude these tokens were excluded from the discussion of appropriateness. These tokens were excluded from the median amplitudes in Table 4.25 and from the boxplots in Figure 4.13.

PAINT	Above Average Amplitude (dB)	Average Amplitude (dB)	Below Average Amplitude (dB)	Appropriate tokens	Total tokens
1	T1;T2;T3;T4;T5			**5**	5
2	T3;T5		T2;T1;T4	2	5
*3	T2;T3;T4;T5		ignore T1	**4**	4
4	T1;T2;T4;T5	T3		**4**	5
6	T1;T3;T4;T5		T2	**4**	5
7	T1;T3;T4;T5		T2	**4**	5
8	T1;T2;T3;T4;T5			**5**	5
9	T1;T3;T5		T4;T2	3	5
*10	T1;T2;T4;T5		Ignore T3	**4**	4
11	T1;T2;T3;T4;T5			**5**	5
12	T1;T2;T3;T4;T5			**5**	5
13	T1;T2;T3;T4;T5			**5**	5
14	T1;T2;T3;T4;T5			**5**	5
15	T1;T2;T3;T4;T5			**5**	5
16	T1;T2;T3;T5		T4	**4**	5
17	T2;T3;T4;T5		T1	**4**	5

Table 4.22 *Amplitude details in individual tokens for Focus position 2* ***(PAINT)*** *sentences in the line graphs for the CI talkers.*

BOAT

Table 4.23 shows that when ***BOAT*** was in focus amplitude was greater than average in all tokens (0 dB) for 12 CI talkers (C1, C3, C4, C7*, C8, C10, C12, C13, C14, C15, C16, C17) and in four out of five tokens for three talkers (C6, C9, C11). All talkers had a maximum of 5 tokens except for C7 who had 4.

BOAT	Above Average Amplitude (0 dB)	Average Amplitude (0 dB)	Below Average Amplitude (0 dB)	Appropriate tokens	Total tokens
1	T1;T2;T3;T4;T5			**5**	5
2	T1;T5		T2;T3;T4	2	5
3	T1;T2;T3;T4;T5			**5**	5
4	T1;T2;T3;T4;T5			**5**	5
6	T1;T2;T3;T5		T4	**4**	5
*7	T2;T3;T4;T5			**4**	4
8	T1;T2;T3;T4;T5			**5**	5
9	T2;T3;T4;T5		T1	**4**	5
10	T1;T2;T3;T4;T5			**5**	5
11	T2;T3;T4;T5		T1	**4**	5
12	T1;T2;T3;T4;T5			**5**	5
13	T1;T2;T3;T4;T5			**5**	5
14	T1;T2;T3;T4;T5			**5**	5
15	T1;T2;T3;T4;T5			**5**	5
16	T1;T2;T3;T4;T5			**5**	5
17	T1;T2;T3;T4;T5			**5**	5

Table 4.23 *Amplitude details in individual tokens for Focus position 3 (**BOAT**) sentences in the line graphs for the CI talkers.*

Table 4.24 summarises for each CI talker the number of tokens with appropriate increased amplitude values relative to the average for each of the target focus words ***BOY***, ***PAINT***, and ***BOAT.*** Overall in Focus position 1, Focus position 2 and Focus position 3 sentences eleven CI talkers (C1, C3, C4, C8, C11, C12, C13, C14, C15, C16, C17) had significant increases in amplitude in the target focus words, ***BOY, PAINT*** and ***BOAT***.

Talker	BOY (n = 5)	PAINT (n = 5)	BOAT (n =5)	Total Appropriate	Total tokens	Proportion of total
1	3	5	5	**13**	15	**0.87**
2	3	2	2	7	15	0.47
3	3	*4	5	**12**	14	**0.86**
4	3	4	5	**12**	15	**0.80**
6	*2	4	4	10	14	0.71
7	0	4	*4	8	14	0.57
8	2	5	5	**12**	15	**0.80**
9	2	3	4	9	15	0.60
10	0	*4	5	9	14	0.64
11	4	5	4	**13**	15	**0.87**
12	4	5	5	**14**	15	**0.93**
13	4	5	5	**14**	15	**0.93**
14	4	5	5	**14**	15	**0.93**
15	4	5	5	**14**	15	**0.93**
16	*2	4	5	**11**	14	**0.79**
17	4	4	5	**13**	15	**0.87**
(* n = 4)	***Significant at 0.75 for 12 appropriate out of maximum of 15 and 0.76 for 11 out of 14***					

Table 4.24 *The number of tokens relative to the total for each CI talker with appropriate increase in amplitude in the target focus words **BOY, PAINT**, and **BOAT**.*

Table 4.25 shows that most CI individual talkers had median amplitude values which were greater than average (0 dB) for the target words ***BOY***, ***PAINT*** and ***BOAT*** when they were in focus. There were some exceptions, however, such as C7 and C10 for ***BOY***, and C2 for ***PAINT*** and ***BOAT*** (see underlined entries). Two tokens (C3:T1 and C10:T3) were excluded for ***PAINT*** sentences because of unusually low amplitude (see Table 4.22).

T tests for the group indicate that as a whole the CI talkers had a significant increase in amplitude for ***BOY***, ***PAINT*** and for ***BOAT*** with p<0.005 when these words were in focus. However, Table 4.25 shows that the CI talkers resembled the NH talkers with similar mean amplitude increases on ***BOY*** (less than 1 dB), on ***PAINT*** (4 dB) and on ***BOAT*** (6 dB).

CI Talkers	Focus position 1	Focus position 2	Focus position 3
	BOY	**PAINT**	**BOAT**
	dB	**dB**	**dB**
1	1.28	4.19	10.36
2	0.50	-0.08	-0.71
3	2.37	3.62	7.59
4	0.29	2.79	6.73
6	0.30	2.33	2.96
7	-1.18	2.68	4.43
8	0.04	4.42	8.58
9	0.36	0.14	2.82
10	-1.07	3.56	5.48
11	3.80	6.65	8.70
12	1.19	5.73	6.80
13	1.48	9.41	9.54
14	1.07	4.61	8.18
15	2.33	3.88	8.87
16	0.56	4.72	10.25
17	1.68	4.58	6.06
mean	0.937	3.951	6.665
var	1.600	5.241	9.354
t	2.96	6.91	8.72
df	15	15	15
sig	**0.0048**	**<0.0001**	**<0.0001**

Table 4.25 *Median amplitudes for individual CI talkers for the focus words* ***BOY, PAINT*** *and* ***BOAT.***

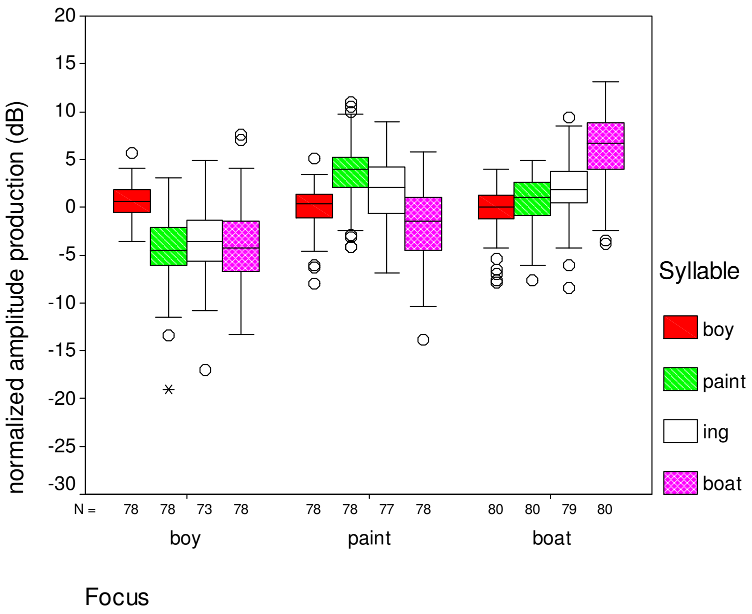


Figure 4.13 *Box and whisker plot of normalised amplitudes for each syllable and focus target for the CI talkers showing smaller median increase for* ***BOY*** *(1 dB) and* ***PAINT*** *(4 dB) than for* ***BOAT*** *(6 dB). (C7:T1 which in* ***BOAT*** *had usually low amplitude was not excluded from Figure 4.13).*

Boxplots in Figure 4.13 for the group of CI talkers (excluding C3: T1 and C10: T3) also show that in general the median amplitudes for the three focus words ***BOY***, ***PAINT*** and ***BOAT*** were greater than the overall average amplitude (0 dB) for these words. However, there were similarities between the NH and CI talkers with the median amplitude increase of 1 dB on ***BOY*** and a subsequent 4-5 dB fall in amplitude on the post-focus syllables which would have been more audible than the smaller amplitude increase on ***BOY***. For both groups the median increases in amplitude on ***PAINT*** (4 dB) and ***BOAT*** (6 dB) were much greater than for ***BOY*** and more likely to be heard in ***PAINT*** and ***BOAT*** than in ***BOY***.

4.3.5.2 Amplitude summary

NH talkers

All the NH talkers increased amplitude of the target focus words except in a few individual tokens in the line graphs in Figure 4.10 for ***BOY*** and ***PAINT***. Individual median amplitude values were greater than the overall average for three NH talkers N1, N2 and N4 but in Table 4.20 the mean increase in amplitude on ***BOY*** (less than 1 dB) might be much less audible than the increases on ***PAINT*** (2 dB) or ***BOAT*** (6 dB). The boxplots for the group of NH talkers in Figure 4.11 also show a smaller increase in median amplitude on ***BOY*** but the subsequent fall in amplitude (2-4 dB) on the post-focus syllables may be more audible. The boxplots show that increases in median amplitude for ***PAINT*** and ***BOAT*** were greater than for ***BOY*** for the group of NH talkers.

CI talkers

Table 4.24 shows that eleven talkers increased the amplitude of the three target focus words ***BOY, PAINT*** and ***BOAT*** in most tokens with a consistency that was significantly above chance in 12 out of 15 tokens (.75) or 11 out of 14 tokens (.76). Tokens which were considered appropriate (see section 4.3.5.1) were above average amplitude (expressed as 0 dB) and tokens with amplitude the same or below average were not considered appropriate.

Individual median amplitude values presented in Table 4.25 show that all except C7 and C10 (***BOY)*** and C2 (***PAINT*** and ***BOAT***) increased the amplitude of the focus words and a t test shows that the group as a whole significantly increased amplitude in ***BOY***, ***PAINT*** and ***BOAT***. However, the mean increase in amplitude for ***BOY*** was very small (1 dB) and was probably less audible than the amplitude increases for ***PAINT*** (4 dB) and ***BOAT*** (6 dB). The boxplots in Figure 4.13 show that the CI group resembled the NH group with a small increase in the median amplitude on ***BOY*** (1 dB) with a subsequent fall of 4-5 dB on the post-focus syllables, and a greater amplitude increase on ***PAINT*** (4 dB) and ***BOAT*** (6 dB).

4.3.6 Correlations between the production of appropriate F_0, duration and amplitude by the CI talkers

In this section the following questions will be discussed in turn*:*

(i) Are there any correlations between the production of F_0, duration and amplitude in Experiment III?

(ii) Overall do CI talkers produce appropriate F_0 contours or increase duration and amplitude of the target focus words, or do they use a combination of cues?

(iii) Are there are any correlations between the production of appropriate F_0, duration and amplitude in Experiment III and rate of stimulation, age at time of production, duration of implant use and age at switch-on?

(iv) Are there F_0 contours WITHIN sentences associated with different focus positions and are they similar or different to patterns produced by the four NH talkers?

(v) If focus *is* heard in individual target words for the CI talkers, which cues are used appropriately?

(vi) What cues are used by CI talkers if focus sounds unambiguous, striking or exaggerated?

(vii) How do CI talkers use F_0, duration and amplitude cues when focus is *not* heard on the target words?

(viii) Are there any differences between NH and CI groups or between CI subjects in the use of F_0, duration and amplitude in the target words ACROSS sentence types in focus and unfocussed position?

(i) Are there correlations between the production of F_0, duration and amplitude in Experiment III?

Correlations are presented for the CI talkers in Table 4.26. A Pearson Correlation test is presented at the top of the table with partial correlations controlling for age at production presented at the bottom of the table. The purpose of the partial correlation test is to test the possibility that the measures correlate simply because of increases in age. Table 4.26 shows that there was a significant correlation with Bonferroni correction between the production of appropriate F_0 and appropriate amplitude, and the production of appropriate duration and amplitude in the target focus words in Experiment III. These correlations remained when age was controlled. However, there was no evidence of a correlation between the production of appropriate F_0 contours vs. duration.

Results for individual subjects presented in the scattergraph in Figure 4.14 show that two talkers (C11 and C14) were significantly above chance in the production of appropriate F_0 and duration (top left), and three talkers (C1, C11, and C14) were significantly greater than chance in the appropriate use of F_0 and amplitude (top right). In the bottom of the figure, eight talkers (C11, C16, C17, C14, C15, C13, C8, C12,) were significantly greater than chance in the appropriate production of amplitude and duration.

		Duration Production	**Amplitude Production**
F0 Production	Pearson Correlation	0.323	0.742
	Sig. (1-tailed)	0.111	**0.001**
	N	16	16
Duration Production	Pearson Correlation		0.659
	Sig. (1-tailed)	.	**0.003**
	N	16	16
Bold type indicates correlations significant at p=0.0167 Bonferroni correct significance level			

Partial Correlations controlling for age at production in Experiment III			
		Duration Productioin	**Amplitude Production**
F0 Production	coefficient	0.3861	0.7377
	df	13	13
	P(1-tailed)	P= .078	**P= .001**
Duration Production	coefficient		0.7453
	df		13
	P(1-tailed)		**P= .001**
Bold type indicates correlation significant at p=0.0167 Bonferroni correct significance level			

Table 4.26 *Pearson correlations (with Bonferroni correction) between* $F_{0,}$ *duration and amplitude production for CI talkers are presented at the top of the table. Partial correlations controlling for age at production are presented at the bottom of the table.*

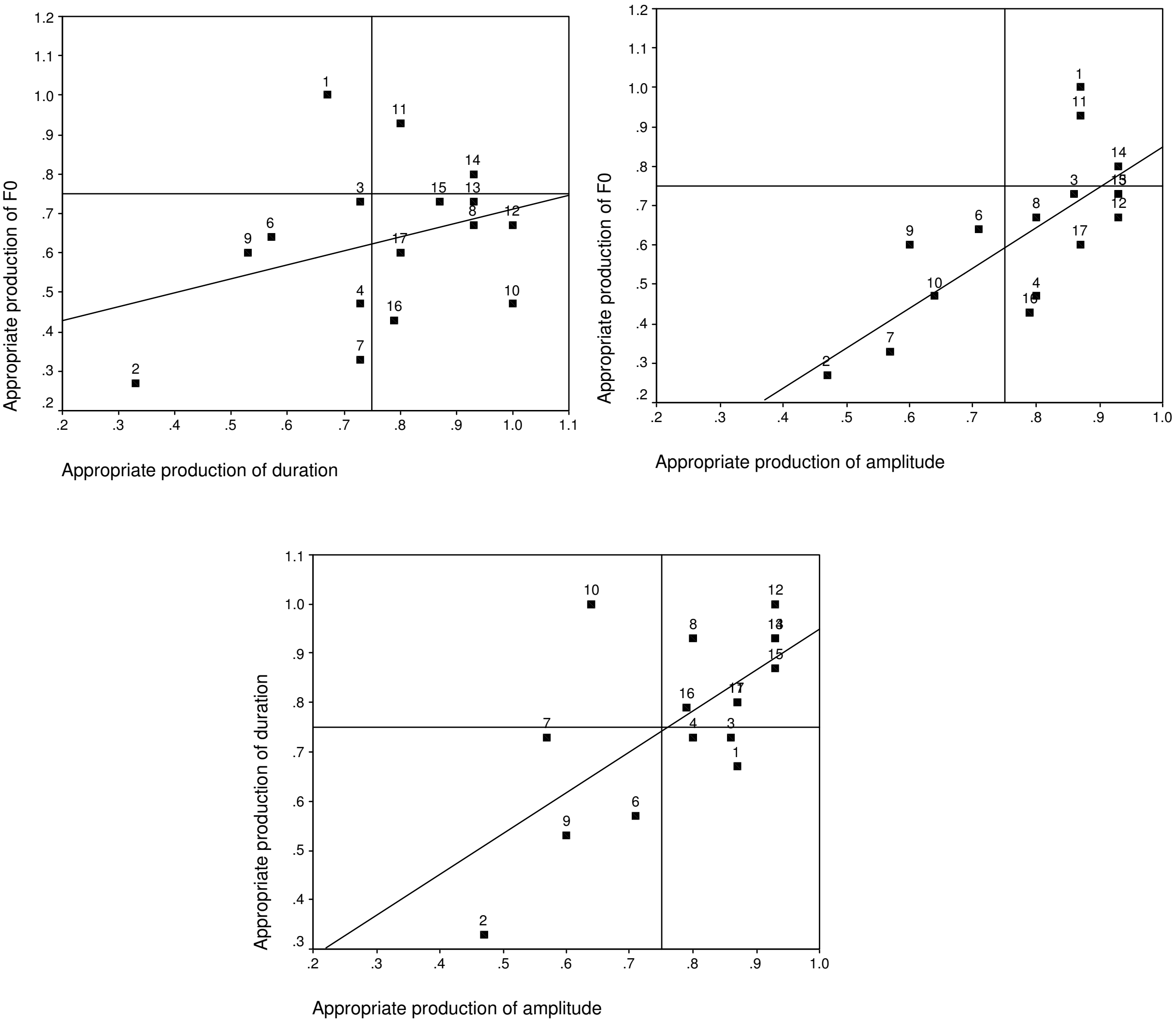


Figure 4.14 *Scattergraphs for the CI talkers showing appropriate production of* F_0 *and duration (top left),* F_0 *and amplitude (top right), and duration and amplitude (bottom). The reference lines at 0.75 on the x and y axes show where the production of appropriate* F_0*, duration and amplitude was significantly above chance.*

(ii) Do CI talkers produce appropriate F_0 contours or increase duration and amplitude of the target focus words, or do they use a combination of cues?

Individual performances in the production of appropriate F_0, duration and amplitude in the three focus words ***BOY, PAINT(ing)*** and ***BOAT*** are presented in Table 4.27 below for individual CI talkers and underlined values indicate the number of appropriate tokens which were significantly above chance level (0.75 or 0.76[4]). Only two talkers (C11 and C14) made significant use of all three cues whereas other talkers used two cues i.e. F_0 and amplitude (C1) or duration and amplitude (C8, C12, C13, C15, C17, C16). There were also a few talkers who only made significant use of one cue i.e. duration (C10) and amplitude (C3 and C4). Four talkers (C2, C6, C7, C9), however, made no significant use of any of the three cues.

	Experiment III Appropriate production		
	F_0	**Duration**	**Amplitude**
1	**1.00**	0.67	**0.87**
2	0.27	0.33	0.47
3	0.73	0.73	**0.87**
4	0.47	0.73	**0.80**
6	0.65	0.57	0.70
7	0.33	0.73	0.60
8	0.67	**0.93**	**0.80**
9	0.60	0.53	0.60
10	0.47	**1.00**	0.67
11	**0.93**	**0.80**	**0.87**
12	0.67	**1.00**	**0.93**
13	0.73	**0.93**	**0.93**
14	**0.80**	**0.93**	**0.93**
15	0.73	**0.87**	**0.93**
16	0.40	**0.78**	**0.77**
17	0.60	**0.80**	**0.87**

Table 4.27 *Appropriate production of F_0, duration and amplitude in individual tokens of the three target focus words for the CI talkers assuming a significant proportion correct at 0.05 level i.e. 12 out of 15 trials (0.75) or 11 out of 14 trials (0.76).*

---

[4] Assuming a sig. proportion correct at 0.05 level i.e.12 of 15 trials (0.75) or 11 of 14 trials (0.76)

(iii) Are there are any correlations between the production of appropriate F_0, duration and amplitude in Experiment III and rate of stimulation, age at time of production, duration of implant use and age at switch-on?

F_0, duration and amplitude production (Experiment III) and rate of stimulation

Pearson correlations with partial correlations controlling for age at time of production in Table 4.28 show that there was no correlation between the production of appropriate F_0, duration and amplitude in Experiment III and stimulation rate.

Results for individual CI subjects in the scattergraphs in Appendix 4.1 show that the majority of talkers who performed significantly greater than chance in the production of appropriate F_0 (two talkers), duration (six talkers), and amplitude (eight talkers) had a stimulation rate of 250 pps. However others who performed significantly above chance in the appropriateness of F_0 (one talker), duration (three talkers) and amplitude (three talkers) were using higher stimulation rates of 900 pps or 600 pps. At the time of Experiment III the number of available talkers within the required age range was limited. In future research is it would be useful to include additional talkers with higher stimulation rates.

		Stimulation Rate	Age at switch-on	Age at production	Duration of CI use at production
F0 Production	Pearson Correlation	-0.296	-0.250	-0.121	0.357
	Sig. (1 - tailed)	0.133	0.175	0.328	0.088
	N	16	16	16	16
Duration Production	Pearson Correlation	0.027	0.374	0.328	-0.270
	Sig. (1 - tailed)	0.460	0.077	0.108	0.156
	N	16	16	16	16
Amplitude Production	Pearson Correlation	-0.356	-0.147	-0.122	0.122
	Sig. (1 - tailed)	0.088	0.293	0.326	0.326
	N	16	16	16	16
Bold type indicates correlation significant at p=0.0042 Bonferroni corrected significance level					

Partial Correlations controlling for age at production in Experiment III			
		Duration of CI use at production	**Stimulation Rate**
F0 production	Coefficient	0.3381	-0.2916
	df	13	13
	P(1 - tailed)	P= .109	P= .146
Duration production	Coefficient	-0.1864	0.0083
	df	13	13
	P(1 - tailed)	P= .253	P= .488
Amplitude production	Coefficient	0.0891	-0.3525
	df	13	13
	P(1 - tailed)	P= .376	P= .099
Bold type indicates correlation significant at p=0.0083 Bonferroni correted significance level			

Table 4.28 *Pearson correlations between* F_0*, duration and amplitude production and stimulation rate, age at time of production, age at switch on, or duration of implant use. Partial correlations controlling for age at time of production are presented at bottom of the table.*

F_0, duration and amplitude production (Experiment III) and age at time of production, duration of implant use, age at switch-on

Table 4.28 shows there was no evidence of any correlations between F_0, duration or amplitude production (Experiment III) and age at time of production, duration of implant use, or age at switch-on.

Overall, individual results in Table 4.27 and in the scattergraphs in Appendices 4.2, 4.3 and 4.4 indicate that CI talkers who were significantly above chance in the production of appropriate F_0 (C1, C11, C14) duration (C8, C10, C11, C12, C13, C14, C15, C16, C17) and amplitude (C1, C3, C4, C8, C11, C12, C13, C14, C15, C16, C17) in Experiment III had a wide age range between 5;6 years to 15;0 years, and were switched on between 2;5 years and 12;7 years. They were using their implants between 1;4 years and 6;7 years. These results show no evidence that production of F_0, duration and amplitude are correlated to age at time of production, duration of implant use, or age at switch-on.

(iv) Are there F_0 contours WITHIN sentences associated with different focus positions and are they similar or different to patterns produced by the four NH talkers?

Only five CI talkers (C1, C11, C12, C14, C15) as shown in Table 4.1 (section 4.3.1) resembled the NH talkers with a fall in F_0 on ***BOY*** in Focus position 1 sentences followed by a levelling or decline in the post-focus syllables.

In Focus 2 sentences seven CI talkers (C1, C3, C6, C8, C11, C13, C15) in Table 4.2 consistently produced a rise-fall in F_0 on ***PAINT*** followed by a decline in F_0 which approximated the NH talkers. Two others (C16 and C12) had a rise-fall or high start F_0 on *boy* and ***PAINT*** followed by a fall in four out of five tokens.

In Focus position 3 sentences in Table 4.3, most CI talkers did not have the terminal rise in F_0 in individual tokens of ***BOAT*** produced by the NH group. However, tokens were considered appropriate if they had lower F_0 peaks on pre-focus syllables, suspended the fall in F_0, or even had a more striking fall in F_0. Four CI talkers (C1, C4, C8, C13) produced F_0 contours which were considered appropriate in all tokens and three talkers (C14, C11, C9) in four out of five tokens.

As discussed in section 1.2 focus may just be a process of boosting or deaccenting acoustic correlates and so might be conveyed by different means such as a striking fall in F_0 or by a terminal rise on ***BOAT***. Although some CI talkers approximated the NH talkers' F_0 contours, it was also pointed out in section 4.3.1.3 that the term *appropriate* did not necessarily mean identical to the NH talkers. Insufficient boosting of F_0, or insufficient deaccenting of pre- or post- focus syllables might have obscured the perception of focus on the target words. This issue is discussed further below.

(v) If focus is heard in individual target words for the CI talkers which cues are used appropriately?

Only four CI talkers (C1, C8 C12, C13) in the present investigator's opinion (see Table 5.1) managed to convey focus in all target focus words ***BOY,***

PAINT(ing) and ***BOAT***. Three of these talkers (C8, C12, C13) made significant use of duration and amplitude, and one talker (C1) used F_0 with amplitude. Other CI subjects who were less consistent in conveying focus to a listener (C11, C14, C15, C16, C17) also varied in their use of different acoustic cues. Two of these subjects (C11, C14) made significant use of all three cues (i.e. F_0, duration and amplitude) which was typical for the NH talkers, three subjects (C15, C16, C17) made significant use of duration and amplitude, and one subject (C10) made significant use of duration only.

Focus was always heard for the NH talkers in the present study and in most cases they increased all three cues in the focus words. However, there were some exceptions, for example, one talker (N3) also had boosted F_0 on *ing* or *boat* following an F_0 peak on the focus word ***BOY*** (T1;T3) and ***PAINT*** (T2;T3) but duration and amplitude adjustments were appropriate and focus was heard on the correct word (see line graphs for F_0, duration and amplitude in Figures 4.1, 4.6 and 4.10). Other talkers had shorter durations in the focus word ***BOAT*** (e.g. N1:T3 and N4:T3) and lower amplitude in ***BOY*** (e.g. N3:T2;T3, N4:T2) and ***PAINT*** (N4:T3) but focus was always heard. Overall the results suggest that the NH talkers generally made use of all three cues whereas there were individual differences for the CI talkers and F_0 did <u>not</u> seem to be a necessary cue to the perception or production of focus (see hypothesis ii)

(vi) ***What cues are used by CI talkers if focus sounds unambiguous, striking or exaggerated?***

In general for some talkers (e.g. C12, C13, C1) the impression of focus was unambiguous and striking on ***BOY*** and ***PAINT*** and even exaggerated at times for others (e.g. C11 and C3). In some tokens for these talkers the fall in F_0 on ***BOY*** or rise-fall on ***PAINT*** and increases in duration and amplitude in the target focus words seen in the line graphs in Figures 4.3, 4.8 and 4.12 are more striking than for other talkers. Table 4.27 summarising production for these talkers shows that overall only one of these talkers (C11) was significantly above chance in the production of appropriate F_0, duration and amplitude, and one (C1) in F_0 and amplitude. Two talkers (C12 and C13) had significant increases in duration and amplitude, and one (C3) in amplitude only. Even though these talkers were capable of producing striking unambiguous or exaggerated focus in some individual tokens only two of the talkers C1 and C11 made significant use of F_0.

(vii) ***How did CI talkers use F_0, duration and amplitude cues when focus was not heard on the target words?***

Table 4.29 shows where focus was not heard in some individual tokens by the present investigator for individual CI talkers, possibly as a result of inappropriate boosting or deaccenting of F_0, duration and amplitude in the pre- or post-focus words.

F_0

The line graphs in Figure 4.3 show that in a few cases (e.g. C11:T1 and C7:T5) an appropriate fall in F_0 occurred on the target focus word ***BOY*** but focus was obscured possibly by inappropriate boosting or insufficient deaccenting of the post-focus syllables. Similarly, in some individual ***PAINT*** sentences focus was not heard despite an appropriate rise-fall in F_0. This may have been due to an insufficient step-up in F_0 (e.g. C4:T2;T3) or insufficient deaccenting after the focus word (e.g.C4:T1), or a terminal rise or striking fall in F_0 on *boat* (e.g. C9:T1;T4). Some individual ***BOAT*** sentences sounded more like neutral declarative sentences and focus was not heard on the target word. For example, C6:T1 and C10:T3 had a gradual decline in F_0 normally associated with neutral sentences, and C16:T1 and C15:T1 had insufficient boosting of F_0 in the terminal rise, and C6:T4 had insufficient deaccenting of pre- focus syllables.

Duration and amplitude

Duration and amplitude were also below average in some of the individual tokens of the three target focus words listed in Table 4.29. Details of individual tokens and talkers are presented in Tables 4.15 - 4.17 for duration and Tables 4.21 - 4.23 for amplitude.

In conclusion the CI talkers may have failed to convey focus either because of insufficient boosting of F_0 in target words or inadequate deaccenting of pre- or post focus words. In some case this may have been combined with inappropriate adjustments in duration and amplitude. However, further investigation is needed and an independent listening test of the production of focus will be carried out in the future with a group of listeners unfamiliar with the data.

	Focus not heard on target words		
	BOY	**PAINT**	**BOAT**
1			
2	T1;T2;T3;T4;T5	T1;T2;T3;T4;T5	
3	T4		T1;T2;T3;T4;T5
4	T1;T2	T1;T2;T3;T5	
6	T5	T2	T1;T3;T4;T5
7	T2;T3;T4;T5	T2;T3;T4	
8			
9	T3;T5	T1;T4	T1;T2;T3;T4
10			T3
11	T1;T5		T3;T5
12			
13			
14			T1
15			T1
16			T1
17	T2		

Table 4.29 *Focus was not heard on individual target focus words* ***BOY, PAINT,*** *and* ***BOAT*** *for some of the talkers.*

(viii) Are there any differences between the CI and NH talkers in the use of F_0, duration and amplitude in the target words ACROSS sentence types in focussed and unfocussed positions?

F_0

As discussed in section 4.3.2 and presented in Tables 4.4 – 4.11, most CI talkers resembled the NH talkers with a rise or fall in median F_0 on the target focus words ***BOY*** and ***PAINT*** which in both cases was significant for the group as a whole. However, only some talkers made a distinction *ACROSS* sentences i.e. between focussed and unfocussed position (i.e. four in the fall from ***BOY***, five in the rise to ***PAINT***, and eight talkers in the fall from ***PAINT***). Only four CI talkers had a terminal rise in F_0 to ***BOAT*** and the rest had a fall (which was significant for the group as a whole) but it was more reduced than when *boat* was not in focus. This would suggest that these talkers were making some distinction between *boat* in focussed and unfocussed positions. Instead of producing a terminal rise like the NH talkers the fall in median F_0 was suspended for the CI talkers when ***BOAT*** was in focus.

Duration

Most CI talkers like all the NH talkers increased the median duration of the target focus words ***BOY***, ***PAINTing***, and ***BOAT*** relative to the average for those words (see Tables 4.14 and 4.19 in section 4.3.4). Exceptions in the CI group were C2 in ***BOY*** and ***PAINTing*** and C9 in ***PAINTing*** but t tests show that ***BOY, PAINT*** and **BOAT** were significantly lengthened for the CI group.

Amplitude

Three of the NH talkers and most CI talkers increased the amplitude of the focus words relative to the averages for those words (see Tables 4.20 and 4.25 in section 4.3.5). Exceptions in the CI group C7 and C10 in ***BOY***, and C2 in ***PAINT*** and ***BOAT***, however t tests show a significant increase for the group as a whole in ***BOY, PAINT*** and ***BOAT***.

4.4 Discussion and conclusion

4.4.1 Acoustic cues to focus used by CI talkers

As mentioned earlier in *The rationale for the analysis of the production data* in section 4.3 the term *appropriate* does not necessarily mean that F_0 contours WITHIN sentences were always identical to the NH talkers so in some cases contours were approaching what was typical for the NH talkers. A conservative chance level of 0.50 was chosen as it was not clear at the outset of this investigation whether the appropriate use of F_0 on the target focus word by CI talkers might be a physiological phenomenon (Cutler and Swinney, 1987) due to tension created by increased interest in the target word or whether the CI subjects had developed an abstract representation of focus or new information even before they acquired concepts such as given vs. new or topic vs. comment (see section 1.3.2.4). However, even if the F_0 or the other acoustic cues (i.e. amplitude and duration) look appropriate in the line graphs, boosting or deaccenting of pre- or post- focus syllables may be insufficient to convey focus on a target word to a listener (see auditory judgement of focus below).

The results of Experiment III as summarized in Table 4.27 show that only some CI talkers made significant use of the three acoustic cues i.e. F_0 (three talkers i.e. C1, C11, C14) duration (nine talkers i.e. C8, C10, C11, C12, C13, C14, C15, C16, C17) and amplitude (eleven talkers i.e. C1, C3, C4, C8, C11, C12, C13, C14, C15, C16, C17) in the line graphs. There were other CI talkers who produced appropriate F_0, duration and amplitude in some individual target focus words which were similar to the NH talkers but they were produced less consistently than the subjects listed above. Table 4.27 shows the CI talkers who were approaching a significant rate of 0.75 for the appropriate production of F_0 (C3, C13, C15) and duration (C3, C4 and C7) which suggests they sometimes use F_0 appropriately but not consistently enough. However, Table 4.27 shows that overall two talkers (C11, and C14) made significant use of all three cues whereas some talkers used a combination of two cues i.e. F_0 and amplitude (C1) or duration and amplitude (C8, C12, C13, C15, C17, C16). There were others who used only one cue i.e. duration (C10) and amplitude (C3 and C4). There were four talkers (C2, C6, C7 and C9) who did not make significant use of any of the cues.

Since only three of the CI subjects (C1, C11, C14) overall made significant use of F_0, the results of Experiment III do not seem to support a physiological theory of F_0 production associated with tension generated by interest in a target focus word. The significant use of amplitude by eleven CI talkers and duration by eight CI talkers seems to lend more support to hypothesis (ii) that F_0 is not a necessary cue to stress and intonation. Judgements of appropriate use of F_0, duration and amplitude are based on visual impressions of the acoustic measurements presented in the line graphs and the auditory impressions of whether focus was conveyed is discussed below in section 4.4.3.

4.4.2 Acoustic cues used by normal hearing children and children with hearing aids

Previous studies of normal hearing children (see sections 1.3.2.2 and 1.11.2) suggest that individual variability in the use of acoustic cues in different prosodic contrasts is not unusual. For example, individual differences in the realization of phonetic exponents used in narrow focus i.e. silence, lengthening, loudness and pitch reset have also been observed by Peppé et al. (2000) in adult speakers of Southern British English. Similarly, Dankovičová et al. (2004) found considerable individual variation and ambiguity in a study of pause duration and final lengthening in a subset of the data for 8;0 year old normal hearing subjects in Wells et al. 2004 (see section 1.3.2.2).

Previous reports of hearing aid users within the same age range as the current study (Rubin-Spitz and McGarr, 1990; Murphy, McGarr and Bell-Berti, 1990) also suggest individual differences in the use of acoustic cues among subjects in the production of syllable stress, but F_0 contours which fell more quickly regardless of the amount were more likely to be perceived as falling. Although listeners sometimes perceived appropriately stressed syllables produced by hearing impaired users the authors conclude that syllable stress might not always be conveyed by the same acoustic correlates. Most (1999) reports that syllable duration in minimal pairs did not play an important role in the perception of correct or incorrect stress production in a study of syllable stress in 10;0 – 13;0 year old Hebrew speakers with hearing aids. F_0 and amplitude were found to be higher in stressed than unstressed syllables for correctly perceived productions. Although individual differences are reported, in most cases where stress was correctly perceived all three parameters were increased.

In a study of contrastive stress (O'Halpin, 1993, 2001) two 8;0 year old subjects did not make appropriate use of F_0 or convey contrastive stress before training and it was anticipated they might have used duration or intensity appropriately. Results show that inappropriate F_0 peaks on normally unstressed syllables obscured appropriate lengthening of target syllables. Following training, however, one talker was able to produce on demand appropriate but often exaggerated F_0, duration and amplitude in target words. These results suggest that variation in the use of acoustic cues is not uncommon in hearing aid users of 8;0 years and older although some make use of all three cues in the production of stress contrasts.

4.4.3 Auditory impression of focus

In the present investigator's opinion, only four CI talkers (C1, C8, C12, C13) managed to convey focus consistently (i.e. in all measurable tokens as presented in the line graphs in Figure 4.3) using a combination of F_0 and amplitude (C1) or duration and amplitude (C8, C12, C13). Because of the limited set of data the CI talkers who consistently conveyed focus are the main concern of the present study. It is also worth mentioning that six other CI talkers (C10, C11, C14, C15, C16, C17) managed to convey focus less consistently i.e. between 11 and 14 out of a total of 15. Table 4.27 and Table 5.1 show that two of these subjects (C11 and C14) used all three cues, three subjects (C15, C16 and C17) used duration and amplitude, and one subject (C10) used duration only. These results, also support the view that F_0 is not a necessary cue to focus (see hypothesis (ii) in sections 1.1.2 and 1.11.4) and indicate that focus, when it is conveyed (either consistently or less consistently), is realized using different combinations of acoustic cues. Six CI subjects in the investigator's opinion (C2, C3, C4, C6, C7, C9) only conveyed focus in 9 or fewer sentences and all except C3 and C4 were older than age 8;0 years at the time of testing. This suggests that the acquisition of the concept of focus might be more delayed for some CI subjects than reported in the literature for some normal hearing children (Cutler and Swinney, 1987; Cruttenden, 1997; Wells et al. 2004).

However, results of these reports vary. For example, Cutler and Swinney concluded in their study that the processing of focus words acquired between 4;0 and 6;0 years whereas Wells et al. found that although focus comprehension lagged behind production some difficult aspects of production of focus (e.g. preference for final focus) and other prosodic contrasts were acquired by 8;0 years. Difficulties reported in the current study for the two CI talkers (C3, C4) who were under 8;0 years may not be altogether unusual in normal hearing children of the same age, but the rest were older which suggests that CI talkers may be more delayed in developing the concept of focus than hearing children. However, Peppé et al., 2000 and Wells et al., 2004 report that ambiguity can be found in normal hearing children up to 13;0 years and even amongst adults (see further discussion of ambiguity in section 4.4.4). In the present study most CI subjects up to 17;0 years failed to convey focus consistently to a listener which indicates that they may not yet have fully acquired this concept, but it is possible that performance might have been affected by the length of experience with the implant. However, the talkers who were least consistent at conveying forms at the bottom of Table 5.1 were using their implants between 1;3 years and 6;2 years (see subject details in Table 2.1) so poor performance does not seem to be linked with years of experience using a cochlear implant.

4.4.4. Ambiguity

Some CI talkers across the age range i.e. (6;0 – 17;0 years) in the present study produced neutral sounding sentences conveying *broad* rather than *narrow* focus and as a result were ambiguous at times. However, ambiguity was observed in medial as well as initial and final position and could be due to insufficient boosting of target focus words or deaccenting of pre- or post focus words. As mentioned above only four CI talkers consistently conveyed focus on the target focus words in the present study and the rest less consistently as indicated in Tables 4.29 and 5.1. Focus was not always heard on the target focus words ***BOY, PAINT*** or ***BOAT*** in such cases and adjustments in F_0, duration or amplitude which looked appropriate in the line graphs may have been obscured by insufficient boosting in one or more of these cues in the focus word or by insufficient deaccenting in pre- or post focus syllables. For example, in some cases focus was not heard by the investigator even though there was an appropriate fall in F_0 in tokens for ***BOY*** (C11: T1, C7: T5) in the line graphs, and this could be because of insufficient deaccenting of post-focus syllables. In other tokens focus on the target focus words could have been obscured by inappropriate boosting of F_0 on other syllables or there may have been an insufficient step-up in F_0 to the target word. There may have been insufficient step – up in F_0 in some tokens of ***PAINT*** (C4:T2; T3) or insufficient deaccenting of the post-focus syllables (C4: T1). Others sounded more neutral e.g. ***BOAT*** (C16:T1 and C15:T1) and the decline of F_0 may have been more typical of a neutral declarative sentence.

However, ambiguity in intonation is not specific to CI children. For example, Wells, et al. (2004) in an investigation of normal hearing children aged 5;0 – 13;0 years report a high instance of ambiguous responses across all age groups (p. 775) especially in utterance final *narrow* focus. It is suggested that this may occur if the final focus word does not have a step-up in pitch or increased duration and amplitude, or if there is more than one strongly accented syllable in the utterance, and they conclude that it may not be developmental as it is also found in adult speech. However, as mentioned earlier, ambiguity for the CI talkers was not just in final position and occurred in initial and medial positions too. Allen and Andorfer (2000) also report that for the hearing aid users in their study (aged between 7;9 and 14;7 years) contrastive use of F_0, duration and intensity in interrogative and declarative sentences was less pronounced (p. 441) than the normal hearing group, and productions were not always correctly categorised by listeners.

4.4.5 Unambiguous and striking focus

In the present investigator's opinion, focus was striking and unambiguous for some tokens for three individual CI talkers (C12, C13, C1) and exaggerated for others (C11 and C3), and in the line graphs F_0, duration and amplitude in some individual tokens looked more striking. Overall, however, only two of these five CI talkers made significant use of F_0 which would support the view that F_0 is not a necessary cue to focus i.e. hypothesis (ii).

4.4.6 NH talkers in the current study

The inconsistency found for some CI subjects in the present study is not unusual in normal hearing children and there are reports of individual variation in children up to 13;0 (Wells et al. 2004). The four NH talkers in the present study managed to convey focus using all three cues with some individual exceptions where duration was shorter and amplitude was lower (see sections 4.3.4.1 and 4.3.5.1) than average as discussed in 4.3.6 (v) or there was inappropriate boosting of F_0 for one talker (N3). The NH talkers (two male aged 16;0 and 20;0 years and two female aged 12;0 and 27;0 years) in the current investigation were used as a small reference group so direct comparisons of the data with the CI talkers could not be made here. Future work, however, will include production data from a group of age matched controls and adults. The line graphs show that in general (see schematic diagrams in Figures 4.2, 4.4, and 4.5) F_0 was increased in individual tokens of the focus words ***BOY*** and ***PAINT*** and was lowered in the post-focus syllables (Xu and Xu, 2005 and see section 1.2) but there were some exceptions as described in section 4.3.1. In Tables 4.4 – 4.11 the measurements of F_0 differences between target focus words and neighbouring words show that four NH talkers had a fall in the median F_0 from ***BOY*** and four talkers had a fall from ***PAINT***. However, three of the NH talkers had a step-up in median F_0 to ***PAINT***, and three talkers had a terminal rise on ***BOAT***. The extent of the rise and fall in median F_0 varied for each talker but in the present investigator's opinion focus was heard on all the target focus words for the NH talkers.

4.4.7 Comparisons between the NH and CI talkers

Similarities and differences between NH and CI talkers were found in the present investigation in the range of median F_0 values used in the rise or fall in F_0 to or from the target focus word to a neighbouring syllable (see Table 4.13 in section 4.3.3). For example, when ***BOY*** was in focus, median F_0 differences between ***BOY*** and *paint* were similar for the NH and CI talkers (1.82 – 6.51 semit. vs. 0.36 – 4.54 semit.) on the other hand when ***PAINT*** was in focus there was a bigger difference in the rise from *boy* to ***PAINT*** (i.e. 0.82 - 3.47 semit. vs. 1.86 - 12.47 semit.) and fall from ***PAINT*** to *boat* (6.94 - 9.19 vs. 0.36 - 13.93 semit.) for the NH and CI talkers respectively. Only four CI talkers had a terminal rise when ***BOAT*** was in focus which was slightly less than for the NH talkers (0.73 – 2.67 vs. 0.61 -1.41 semit.). The rest of the CI talkers had a fall in F_0 on ***BOAT*** which was significant for the group which was more reduced or suspended than when *boat* was not in focus showing differentiation between focussed and non-focussed target words in final position in a different way to the NH talkers.

4.4.8 Difficulty with rising intonation for the CI talkers

Overall, it would appear from acoustic measurements that the median change in F_0 produced by the CI talkers resembled the NH talkers (see Tables 4.4 – 4.11 in section 4.3.2) when the target focus words were in initial and medial position but not for the rise in sentence-final position. CI talkers who did not produce a terminal rise in median F_0 on target focus words in final position had a more reduced fall in median F_0 on ***BOAT*** when it was in focus. The measurements in Table 4.11 shows that twelve CI talkers had a fall in median F_0 on ***BOAT*** which was significant for the group but only four talkers had a terminal rise in F_0 as observed for NH talkers in the line graphs in Figure 4.1 except in a few tokens for subjects N1 and N2 where F_0 remained level or suspended when focus was on ***BOAT***. In medial focus position however, fifteen of the CI talkers were able to produce a non-terminal rise in the median F_0 from *boy* to ***PAINT*** (Table 4.7) and the rise was significant for the group as a whole.

4.4.9 Rising intonation in normal hearing children and hearing aid users

Wells et al. also report difficulties with contrasts such as rising intonation for questioning or a fall-rise in expressing dislikes up to 8;0 years in normal hearing talkers, and Snow (1998, 2001) reports that 4 year-olds had narrow pitch excursions with lengthening in sentence-final rising tones due to motor difficulties. In the current study difficulties with rising intonation were found for CI children across the age range. As mentioned earlier there was a greater range in median F_0 in the rise and fall to and from the target focus word ***PAINT*** for the CI than NH talkers but there was a similar range for both groups in the fall from the target focus word ***BOY*** (Table 4.13). Previous studies (Rubin-Spitz and McGarr, 1990; Most and Frank, 1994) of hearing users indicate rising patterns are more difficult to produce. However, in a longitudinal study of 7;0 - 8;0 year olds Abberton, Fourcin and Hazan (1991) reported that rising intonation began to emerge in the speech of some of their profoundly deaf children after a four year period. More recently, Allen and Andorfer (2000) also managed to elicit rising terminal pitch contours from hearing impaired users aged between 7;9 and 14;7 years. Rubin-Spitz and McGarr (1990) on the other hand report that, unlike their control hearing subject, none of their hearing impaired subjects produced rising contours. Instead they had terminal falling vs. non-falling contours like the CI talkers in the current study on final target words like ***BOAT.*** McGarr et al. also report that in some cases that listeners perceived a fall when duration was short and non-falling when duration was long. Although many of the studies cited above involve hearing aid users and normal hearing talkers there are some similarities in the results of the present study of CI talkers with respect to ambiguity, individual differences in the use of different acoustic cues and the absence of terminal rise in F_0 in final focus position for most CI talkers.

4.4.10 Rising tones in Chinese speaking CI users

Only a few studies have been carried out on children with cochlear implants and their production of F_0 in lexical tones. Peng, Tomblin, Cheung, Lin and Wong (2004), for example, report in a study of Mandarin lexical tones in 30 prelingually deafened children with cochlear implants (aged 6;0 -12;0 years) that production ratings were better for level (T1) and high falling tones (T4) than for mid-high rising (T2) and low dipping (T3) tones. They also found that although the acquisition of tone production was delayed, the order of acquisition was consistent with normal hearing development where level and falling tones are acquired before rising tones. In a different study of Mandarin tone production (Xu, Li, Hao, Chen, Xue and Han, 2004) of four prelingually deafened implanted children (aged 4;0 - 8;75 years), individual variation was found in imitated productions of target tones. The easiest tone to produce was a high falling tone (T4) and some had difficulty with rising tones. A group of normal hearing subjects obtained maximum scores in an intelligibility test whereas the CI subjects ranged from 0.25 to 8.5. In another study Barry and Blamey (2004) measured differences between tones produced by 16 Cantonese speaking implanted children (aged 4;2 to 11;3 years) by plotting onsets (x axis) and offsets (y axis) of F_0 to capture average pitch, direction, extreme endpoint and slope (see section 1.8). It was expected that rising tones would cluster close to the y axis and falling tones cluster to the x axis and level tones would fall midway between the two axes. Where consistent patterns are used for each tone by a speaker the plots are predicted to be well differentiated and should correlate with perceptual judgements. Very little differentiation in the production of falling and rising tones by CI talkers was observed. However, direct comparison with the present study is difficult as it concerns the use of F_0 in the production of focus in English using a different methodology and a wider age range in the subjects. In English there are additional cues to stress and intonation i.e. duration and amplitude which play a more minor role in signalling differences between Mandarin and Cantonese tones.

4.4.11 Correlations between F_0, duration and amplitude production by CI talkers in the current study

Pearson correlations and partial correlations controlling for age at time of production in Table 4.26 for the CI talkers show that there were correlations between the production of appropriate F_0 and amplitude, and between the production of appropriate duration and amplitude but not between the appropriate production of F_0 and duration. This supports the possibility of a trade-off between duration and F_0 (see section 1.4.1) as demonstrated for adult normal hearing speakers in an early study by Isenberg and Gay (1978) and more recently by Kochanski et al. (2005). In other words increased duration may be a better cue to stress and intonation than F_0. The scattergraphs in Figure 4.14 illustrate individual performances and show how nine CI subjects made significant use of duration and amplitude in the target focus words whereas only two subjects made significant use of F_0 and duration and three subjects significant use of F_0 and amplitude. These results seem to support Konchanski et al. (2005) who investigated a large corpus of English (see section 1.4.2 and 1.11.2) and suggested that F_0 plays a minor role and accent and prominence are marked by loudness and duration cues. This according to Kochanski et al. is contrary to the traditional view that F_0 is the main cue to prominence (based mainly on laboratory speech). However, the study did not make a distinction between functional aspects of stress such as focus or lexical stress so results are difficult to compare with the present study where specific contrasts are elicited.

Wells et al. (2004) also suggest that there may be differences between subjects in phonetic exponents of intonational contrasts (silence, lengthening pause, pitch reset) in less controlled social situations compared to laboratory speech. The elicited responses and use of picture prompts in the current study are as close as possible to natural conversational situations without losing control of the linguistic content. The results for the individual CI talkers in the current study support hypothesis (ii) which suggests that F_0 is not a necessary cue to focus for the CI talkers (see section 1.1.2) and that more talkers seem to make significant use of duration and amplitude rather than F_0 in target focus words. However, as mentioned earlier appropriate F_0 contours observed in the line graphs may not always convey focus to a listener for reasons such as ambiguity as discussed above in section 4.4.3 and 4.4.4.

4.4.12 Effects of variables such as age at test, age at implant, duration of implant use and stimulation rate on production of appropriate F_0, duration and amplitude

There was no evidence of any correlations between the production of appropriate F_0, duration and amplitude and stimulation rate, age at time of production, age at switch-on, or duration of implant use. Only some of the CI talkers within the appropriate age range for Experiment III were using higher stimulation rates at the time of production and most were using a slower rate of 250 pps. This was because only a limited number of talkers using higher stimulation rates of 600 pps and 900 pps were available within the required age range at the time of testing. Future work will include additional talkers using higher stimulation rates and differing in age and duration of implant use. The only available data in the literature are drawn from studies of Chinese tones where F_0 is the most important cue to lexical meaning (see section 1.8.2). In a report on the production of Mandarin tones Peng et al. (2004) found no significant difference between faster and slower stimulation rates (CIS and SPEAK), but they state that for the group of children investigated (aged 6;0 – 12;0 years) tone production was better for those implanted at an early age. Barry and

Blamey (2004) found that their implanted children (aged 4;2 – 11;3 years) produced some F_0 contours that could be labelled correct but were not consistent enough to be considered acquired. They suggest that longitudinal studies would be appropriate for measuring tonal development in individual children. Xu et al. (2004) concluded in their study of implanted children (aged 4.0 – 8;75) that limited pitch information delivered through cochlear implants may hinder tonal development and that other variables such as age and stimulation rates needed to be considered.

4.4.13 Summary of Experiment III results

a) Individual CI subjects varied in the appropriate use of acoustic cues (F_0, duration, amplitude) in the production of target focus words.

b) The CI and NH subjects had a similar fall in median F_0 on the target focus word ***BOY*** which was significant for the CI group.

c) Acoustic measurements show that the CI subjects as a group significantly increased the median amplitude and duration of the target focus words in initial (***BOY***), medial (***PAINT***) and final positions (***BOAT***).

d) Ambiguity was observed for many of the CI talkers in the current study but as reported in the literature, this is not uncommon for normal hearing talkers. It is also reported that stress contrasts are not always correctly categorized for hearing aid users.

e) Only four subjects consistently managed to convey focus to a trained listener using different combinations of acoustic cues with and without F_0 which suggests that F_0 is not a necessary cue to stress and intonation i.e. hypothesis (ii).

f) The literature also reports individual variation in the use of different combinations of acoustic cues by normal hearing children and children using hearing aids.

g) Falling intonation, which is normally associated with focus in the literature, occurred on an initial and medial target focus words (i.e. ***BOY*** and ***PAINT***) as shown for individual NH and CI tokens in the line graphs. Falling intonation on these words was observed in the median F_0 measurements for the NH and most CI talkers. The median fall in F_0 was significant for the CI group for ***BOY*** but not for ***PAINT***.

h) Difficulties with rising intonation in final focus position ***(BOAT)*** occurred for most CI talkers and they did not produce a terminal rise in F_0 which was observed for most of the NH talkers. As reported in the literature, this is not unusual for normal hearing children generally and non-terminal falling F_0 contours have been observed in children using hearing aids. In medial focus position, however, most CI talkers managed to produce a rise in the median F_0 to the target focus word ***(PAINT***) which was significant for the group.

i) Instead of a terminal rise in F_0 five CI talkers had a fall in F_0 on the final focus item (***BOAT)***. However, five of these talkers had a more suspended fall than when *boat* was not in focus. Even if the CI talkers did not succeed in conveying focus to a listener (sections 4.4.4 and 4.4.5) it is possible that some were attempting to signal focus in final position by suspending or reducing the more striking decline of F_0 which occurs following focus on earlier words (i.e. ***BOY*** or ***PAINT***).

j) There was variation in performance across the age range of CI subjects but most subjects who were below the chance level in conveying focus to a listener were over 8;0 years. The literature suggest that some normal hearing children also take longer to acquire the concept of focus and some prosodic contrasts may not be acquired by 13;0 years or even into adulthood. However, since only four CI talkers in the present study consistently conveyed focus to the investigator, we can conclude that this contrast not yet been fully acquired by most CI subjects across the age range up to 17;0 years.

k) Studies of Chinese tones with a younger group of CI children suggest that rising tones in lexical tone contrasts, cued mainly by $F_{0,}$ were also not yet acquired. However, it is difficult to compare Chinese tones with English intonation contrasts which can be cued by one or more cues (i.e. F_0 and/or duration and amplitude). In addition the current study also included a wider age range of CI subjects up to 17;0 years.

l) Additional CI data and NH data with age matched controls in future research would facilitate more direct comparison than in the current study. To date there are only a few studies of the production of focus in a normal hearing population and none beyond age 13;0 years, and there are no available normative studies based on a Southern Irish population. However, the NH data in the current study included four Irish subjects aged 12;0 - 27;0 years which provided a useful reference for the analysis of the CI productions in Experiment III.

4.4.14 Issues to be addressed in Chapter Five

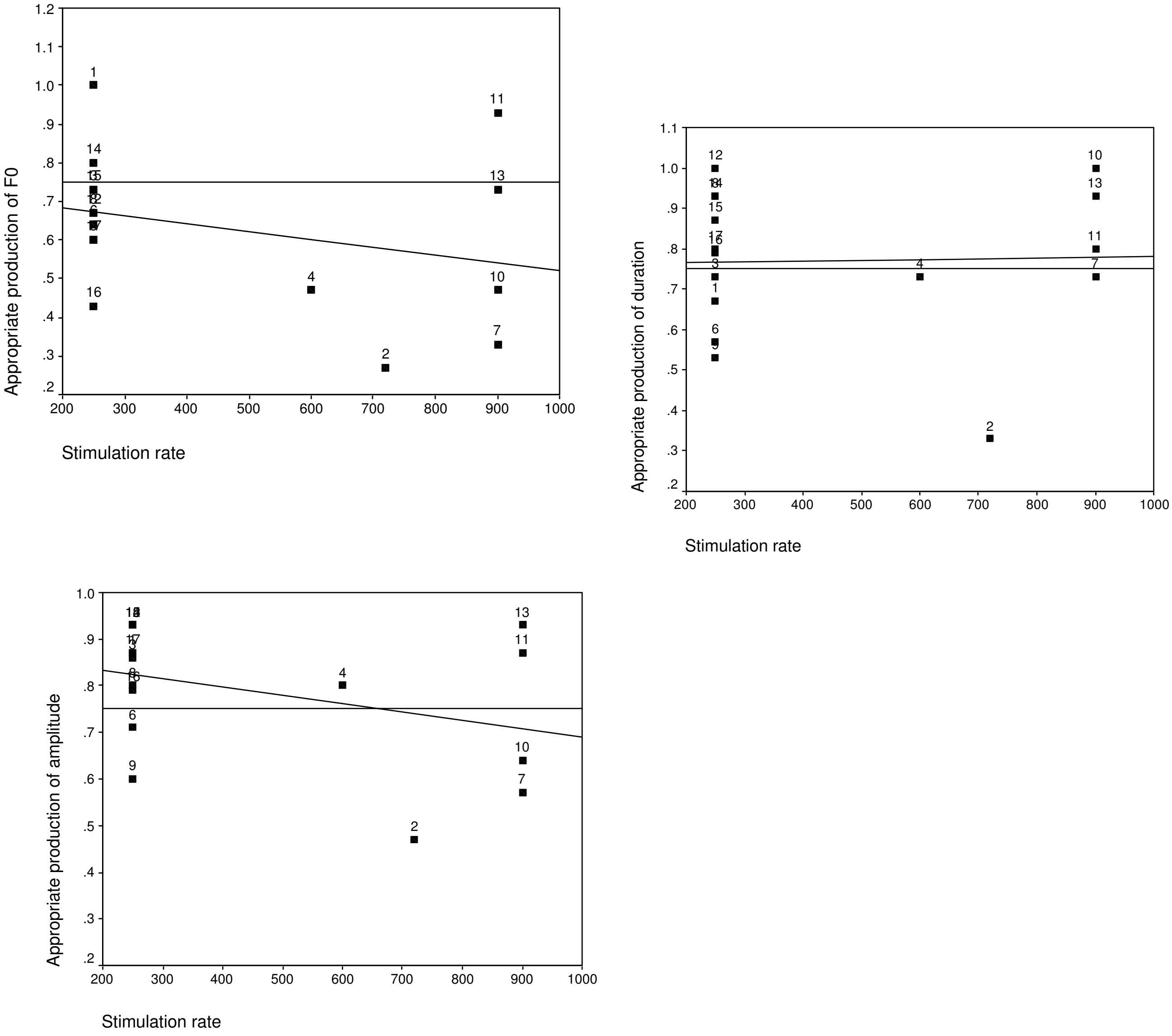
Chapter Five explores in more detail the relationship between perception and production of linguistic focus in Experiments II and III to establish whether it

(i) _is_ directly linked to the implanted children's ability to hear changes in F_0 (with or without duration or amplitude) in Experiment I and whether the development of as linguistic focus depends on their auditory skills, and F_0 is a necessary cue (hypothesis (i) see section 1.1.2).

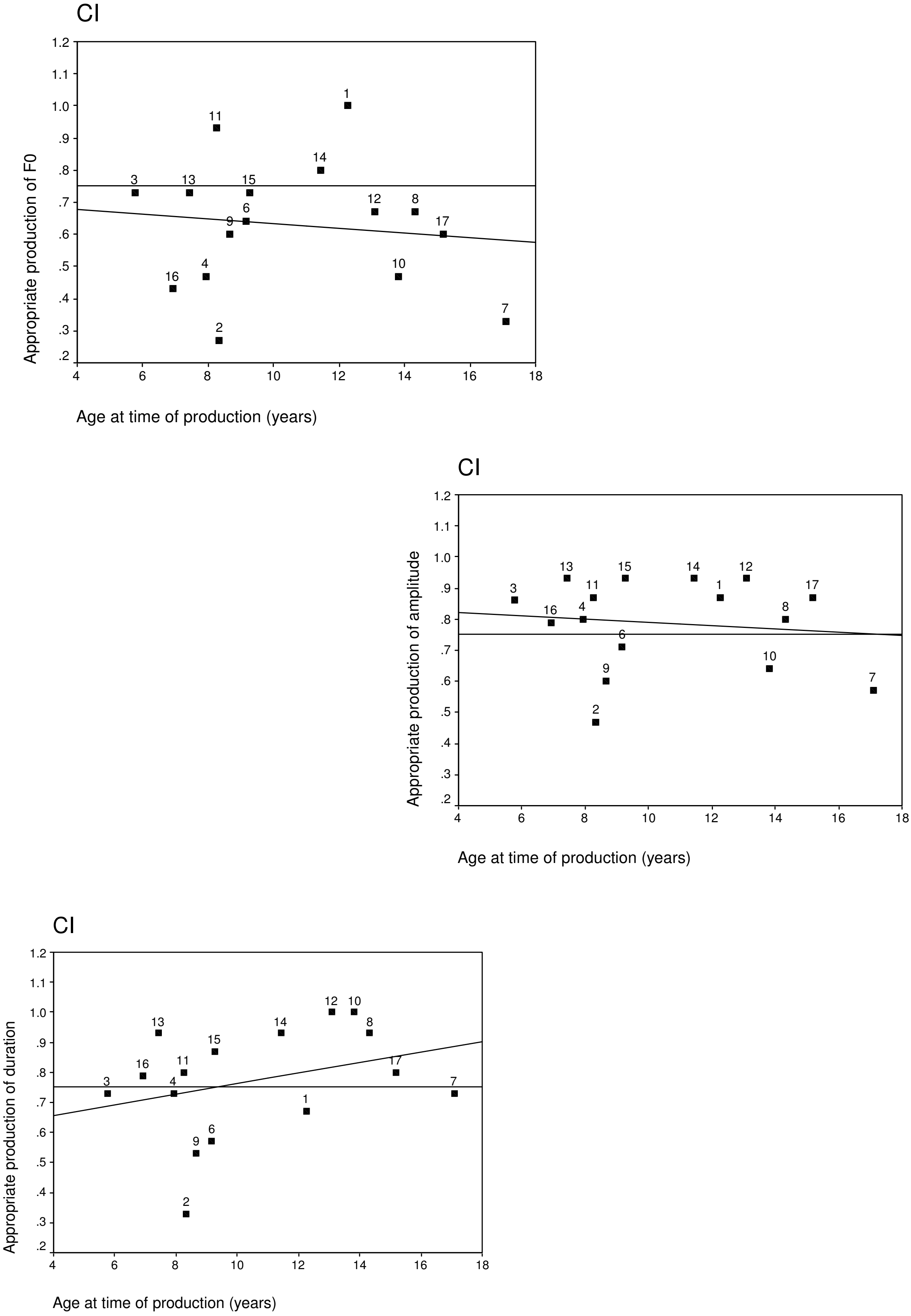
or

is not directly linked to any one cue and the concept of focus develops as an abstract phonological system which is not necessarily perceived and produced by the same cues, and that F_0 is not a necessary cue (hypothesis (ii) section 1.1.2)

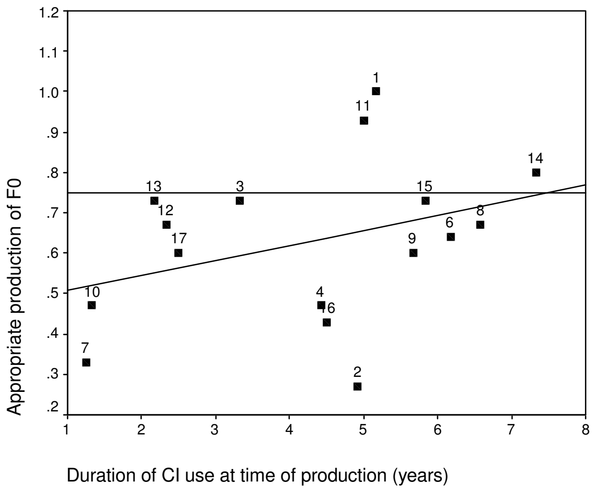
4.5 Appendices

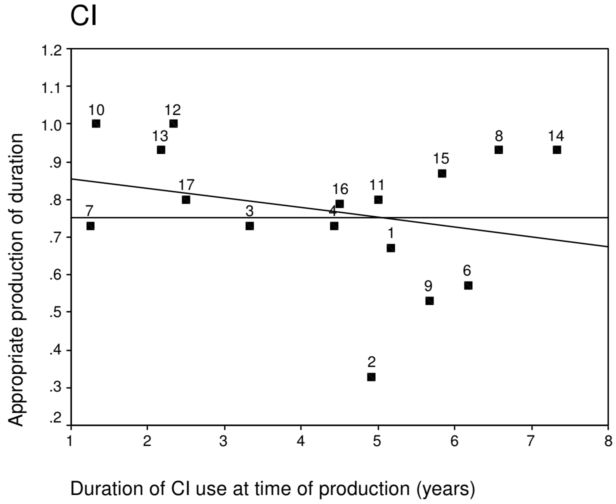


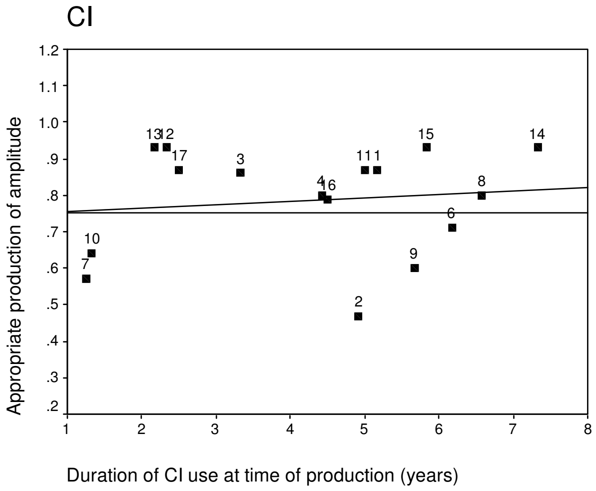
Appendix 4.1 *Scattergraphs for the CI talkers showing production of* F_0, *duration and amplitude and stimulation rates in Experiment III*



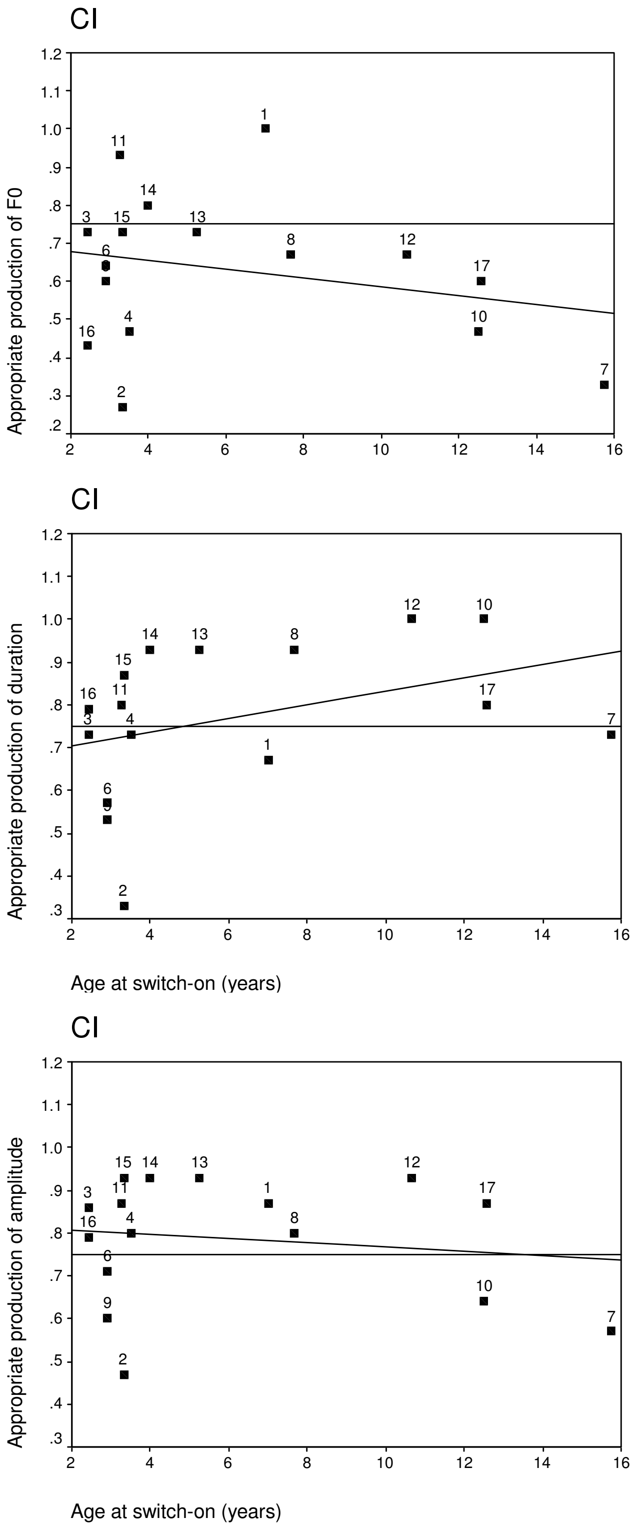
Appendix 4.2 *Scattergraphs showing age at time of production of focus in Experiment III and appropriate* F_0*, duration and amplitude production.*







Appendix 4.3 *Scattergraphs for the CI talkers showing duration of CI use at Experiment III and appropriate* F_0*, duration and amplitude in the production of focus.*



Appendix 4.4 *Scattergraphs for the CI talkers showing age at switch-on and the appropriate production of* F_0*, duration and amplitude in Experiment III.*

CHAPTER FIVE

COMPARISONS BETWEEN PERCEPTION AND PRODUCTION OF F_0, DURATION, AMPLITUDE, AND FOCUS BY CI SUBJECTS

5.1 The relationship between perception and production of focus: implications of Experiments I, II and III results for CI users

5.1.1 Overview of issues raised in Chapter One: Is F_0 a necessary cue to stress and intonation?

As discussed in sections 1.11.4 and 1.3.2.4, there is an accepted view in the literature that perception precedes production in language development but it has also been suggested that this may not be the case for prosodic development (Stackhouse and Wells, 1997; Wells et al. 2004), and that four year old normal hearing children may be able to produce accent and focus in their own speech before they can interpret it in the speech of others. This supports a previous study (Cutler and Swinney, 1987) which suggests that productions of 3;0 and 4;0 year old children may be similar to 5;0 and 6;0 year olds because a semantically interesting word generates tension and excitement. One possible explanation is that a rise in pitch could be due to a physiological reflex rather than prosodic competence in the younger children who may not be able to process contrasts such as given vs. new information yet can produce appropriate accentuation to convey focus or new information. However, it is not yet clear whether CI children who have poor access to F_0 differences less than 0.5 octaves through their implants (see Experiment I in Chapter Two) can produce appropriate changes in F_0 on target words either for the physiological reasons mentioned above or to convey focus to a listener.

The relationship between perception and production of F_0 in Experiments I and III for CI subjects in the current study is addressed in detail in section 5.2 below. Traditionally F_0 has been considered the most important cue to stress and intonation and as discussed in section 1.1.2 and 1.11.4 cochlear implants provide only limited access to F_0. More recently, Xu and Xu (2005) suggest that the location of F_0 peaks in English are determined by lexical stress, metrical structure or information load and are independent of focus, while narrow focus leads to an increase in F0 peak height. While pitch adjustments occur on a focus word such as an increase in the size of the F_0 peak (and increases in duration and amplitude), the pre-focus F_0 peaks remain unchanged, and post-focus F_0 peaks are lower than in neutral conditions (section 1.2). However, Peppé et al. (2000) report differences in the use of phonetic exponents of narrow focus by adult speakers of British English in their study e.g. silence, lengthening, loudness and pitch reset (sections 1.11.2), and they also suggest that in less controlled settings compared to laboratory conditions there may be differences in how intonational contrasts are realised. This view is supported by Kochanski et al. (2005) who reported that syllables perceived as prominent by listeners in their study were marked by loudness and duration cues, and that F_0 played a minor part. However, the results are not conclusive as specific intonational contrasts such as focus were not analysed by Kochanski et al. These issues are addressed in the discussion of the hypotheses in sections 1.1.2 and 1.11.4 and with reference to Experiment I, II and III results in the following section.

Hypothesis (i): F_0 is a necessary cue to stress

The traditional view which suggests F_0 is a necessary cue to stress and intonation as set out in hypothesis (i) in sections 1.11.4 and in 1.1.2 means that implanted children will need good access to pitch (perceptual correlate of F_0) in order to hear these contrasts. In other words perception and production of intonation are directly linked to their ability to hear F_0. If they do not have access to F_0 they will be unable to develop abstract phonological representations of intonation contrasts in the same way as normal hearing children. Since they cannot hear the associated F_0 patterns associated with intonation contrasts they may not have prior knowledge or stored representation of semantic, pragmatic and grammatical contrasts, and might never be able to produce them properly. Previous experiments indicate that children with implants require F_0 differences of almost half an octave (sections 1.8 and 1.11.3) which may be greater than F_0 differences found in everyday speech. Experiment I of the current study (Figure 2.3) provides further evidence that CI listeners have difficulties hearing peak F_0 differences greater that 0.5 octaves, although a few were hearing smaller F_0 differences in the high F_0 range (from a 200 Hz baseline). Median F_0 thresholds for the group of CI subjects were above 0.5 octaves at 77% from a 100 Hz baseline and 57% from a 200 Hz baseline (Figure 2.4). However, despite limited ability to hear F_0 differences, Experiment II results show that perception of linguistic focus (and compound vs. phrase stress) was possible for some of these listeners. Here, scores ranged between 38% and 100% (Figure 3.1) with some individuals scoring above chance levels in each of the three subtests. Furthermore, some of the implanted children who were significantly above chance in the perception of focus or stress pattern were not able to discriminate F_0 differences consistently even at the maximum difference level presented (84%) in Experiment I. F_0 measurements for the four talkers in the focus stimuli Experiment II (Appendix 3.2 and boxplots in Appendix 3.3) make clear that F_0 differences between target focus words and neighbouring words rarely exceeded 0.5 octaves (section 3.5.4.1) and would not have been accessible to most of these CI listeners. Experiment I and II results taken together suggest that F_0 may not be a necessary cue to focus and compound vs. phrase stress.

Hypothesis (ii): F_0 is <u>not</u> a necessary cue to stress

As outlined above, some of the CI subjects were able to hear intonation contrasts in Experiment II at a level significantly greater than chance even though they would not on the basis of Experiment I results be able to hear the F_0 differences cueing focus or stress. This suggests that the perception of these intonational contrasts does not depend on the ability to hear F_0 differences and thus that F_0 is not a necessary cue as set out in hypothesis (ii) in sections 1.11.4 and 1.1.1. It follows that these implanted children must rely on other acoustic cues such as duration and amplitude. If this is the case CI users may not be at a disadvantage during the early stages of prosodic development. It is possible that perception and production of intonation may not be directly linked to any one cue and intonation may develop as an abstract phonological system and that perception and production need not involve the same acoustic cues. However it is also possible that the physiological reasons mentioned above and tension associated with an interesting word might account for appropriate use of F_0 by some CI subjects and other CI subjects who have developed an abstract representation of focus might be able use F_0 appropriately in the production of focus without necessarily being able to hear these F_0 differences (see production of F_0 below).

5.1.2 Is duration a reliable cue to focus for CI subjects?

As discussed in 1.3.1.2 and 1.11.1 prosodic cues such as extra lengthening, longer pauses, differences in loudness, and paralinguistic cues such as eye contact, gesture, jumping up and down can draw attention to certain features such as rhythm or focus and help develop an abstract linguistic system using all available cues. Experiment II results show that most of the CI subjects who scored significantly greater than chance in the perception of linguistic focus were able to hear duration differences less than 60% in Experiment I (see discussion in 3.5.4.2 and Figure 2.6). Although individual duration thresholds varied between 5% and 138%, the median duration threshold for

CI subjects in Experiment I was 35%. Figure 2.6 shows that the CI subjects performed as well as the NH subjects in the simulation condition i.e. CI and NH subjects could hear duration differences less than 60% and the median thresholds were similar at 35%. Duration measurements for the focus stimuli presented in Experiment II (section 3.5.4.2 and Appendices 3.5 and 3.6) indicate that duration differences between target focus words and neighbouring words were generally greater than 35% and should therefore be accessible to most CI subjects. These results suggest that duration may provide a stronger cue to linguistic focus than F_0 for some subjects.

5.1.3 Is amplitude a reliable cue to focus for CI subjects?

Individual amplitude thresholds for CI subjects also varied in Experiment I between 3 dB and 15 dB with most hearing differences of 12 dB or less. The median amplitude threshold for the group was 11 dB (Figure 2.8). Amplitude measurements for the stimuli presented in Experiment II (Appendices 3.7 and 3.8) show a wide variation in amplitude differences and often these differences were too small to be accessible to some CI subjects. However, some CI subjects with large amplitude thresholds were still able to hear focus in Experiments II (section 3.5.4.3 and Figure 3.7), and therefore, prosodic perception could not be entirely due to amplitude cues. These results suggest that duration might be a more reliable perceptual cue than F_0 and amplitude for CI subjects.

Focus perceived		***Experiment I (Perception thresholds)***					***Experiment II (Focus Perception) sig > chance = 45.8%***		***Experiment III Appropriate production of acoustic cues Significance level = 0.75***				***Age at Exp III***
100 %	***N=15 *=14***	***low F_0 range %***	***high F_0 range %***	***Duration %***	***Amplitude dB***	***Age at Exp I***	***Focus 3***	***Age at Exp II***	***F_0***	***duration***	***amplitude***	***Combination of cues sig > chance in production***	
C1	15	27	20	10	5	11;10	89	11;11	**1.00**	0.67	**0.87**	F0 & amplitude	12;3
C8	15	51	27	17	9	14;4	90	14;1	0.67	**0.93**	**0.80**	duration & amplitude	14;4
C12	15	76	21	49	7	12;8	71	12;8	0.67	**1.00**	**0.93**	duration & amplitude	13;1
C13	15	44	25	15	10	7;6	92	7;3	0.73	**0.93**	**0.93**	duration & amplitude	7;5
> 12													
C10	14	80	36	28	11	13;9	81	13;10	0.47	**1.00**	0.67	duration	13;10
C11	11	54	12	15	13	8;7	56	8;1	**0.93**	**0.80**	**0.87**	Fo, duration & amplitude	8;3
C14	14	82	54	43	11	10;11	52	11;00	**0.80**	**0.93**	**0.93**	Fo, duration & amplitude	11;5
C15	14	55	79	58	5	8;9	62	8;0	0.73	**0.87**	**0.93**	duration & amplitude	9;3
C16*	13	81	79	128	11	6;11	31	6;11	0.40	**0.78**	**0.77**	duration & amplitude	6;11
C17	13	53	29	24	3	14;7	90	14;9	0.60	**0.80**	**0.87**	duration & amplitude	15;2
<9													
C2	4	83	82	38	11	8;0	35	8;1	0.27	0.33	0.47	no significant use of cues	8;4
C3	9	59	26	17	10	6;1	56	5;7	0.73	0.73	**0.87**	amplitude	5;9
C4	8	84	83	81	15	7;11	44	7;11	0.47	0.73	**0.80**	amplitude	7;11
C6	8	79	78	108	15	9;0	56	8;10	0.65	0.57	0.70	no significant use of cues	9;2
C7	8	46	58	11	9	17;4	79	16;11	0.33	0.73	0.60	no significant use of cues	17;1
C9	7	81	84	51	11	8;3	44	8;3	0.60	0.53	0.60	no significant use of cues	8;0

Table 5.1 *Individual CI subjects' scores for Experiments I, II and III.*

5.1.4. What acoustic cues are used by CI talkers in the production of focus in Experiment III?

Experiment III results summarized in Table 5.1 show considerable individual variation in the use of acoustic cues in the production of focus, i.e. with three talkers consistently using F_0, nine consistently using duration, and eleven consistently using amplitude. However, only four of the 16 CI subjects (C1, C12, C13, C8) managed to convey focus to a trained listener (the present investigator) and only one of this subset of four (C1) made significant use of F_0 (with amplitude) whereas the other three others used duration with amplitude. Although other CI talkers made significant use of different combinations of cues they did not manage to convey focus consistently to this listener. Sometimes they sounded ambiguous possibly as a result of insufficient boosting of focus words and/or deaccenting of pre- and post focus words (section 4.4.4).

The results of Experiments I, II and III so far seem to support hypothesis (ii) that F_0 is not a necessary cue to intonation contrasts such as lexical stress or focus. Chapter Five explores in more detail the relationship between perception and production of focus and F_0 duration and amplitude for the group of CI subjects as well as individual performances presented in the scattergraphs in Figures 5.1 – 5.9.

The following questions are addressed:

a. Is it necessary to hear differences in acoustic cues (F_0 or duration or amplitude) in order to produce them appropriately in target focus words? (section 5.2)

b. Is it necessary to be able to perceive focus in order to be able to produce it by appropriate use of one or a combination of acoustic cues (i.e. F_0, duration or amplitude) on the target focus words? (section 5.3)

c. Can linguistic focus be perceived by one or a combination of cues and produced by a different set of cues?

5.2 Are there correlations between the production of F_0, duration and amplitude and the perception of F_0, duration and amplitude differences?

Sections 5.2.1, 5.2.3 and 5.2.4 below explore the relationship between perception and production of the acoustic cues F_0, duration and amplitude for the group of CI subjects using Pearson Correlation tests as well as scattergraphs in Figures 5.1 to 5.9. The results have implications for how implanted children might perceive and produce intonation contrasts such as focus, in particular, whether they use one or a combination of cues to perceive focus in Experiment II and the same or a different set of cues to produce it in Experiment III.

5.2.1 F_0 production (Experiment III) and F_0 perception (Experiment I)

The purpose of the Pearson Correlation test was to establish whether there was a statistical link between the ability to produce appropriate changes in F_0 in Experiment III and the ability to perceive F_0 differences in Experiment I. The presence of such a link would be consistent with the view that it is necessary to be able to hear differences in F_0 in order to produce them appropriately (see hypothesis (i) in section 5.1.1). As discussed above, Experiment I results suggest that implanted children needed approximately 0.5 of an octave (i.e. 40%) change in F_0 in the low F_0 range before they could hear a difference. But in the high F_0 range, however, there were some individual subjects who were able to hear smaller F_0 differences. The results will be discussed separately for the high and low F_0 ranges below. For the purpose of these analyses, F_0 production range will be classified as high or low in line with the F_0 range classifications of Experiment I, where F_0 from 100 Hz upwards to 200 Hz was considered “low”, and an F_0 range from 200 Hz upwards was considered “high”.

		Low F_0 threshold	**High F_0 threshold**	**Duration threshold**	**Amplitude threshold**
F_0 production	Pearson Correlation	-0.450	-0.589	-0.318	-0.238
	Sig. (1-tailed)	0.040	**0.008**	0.115	0.187
	N	16	16	16	16
Duration production	Pearson Correlation	-0.124	-0.539	-0.181	-0.257
	Sig. (1-tailed)	0.324	0.016	0.251	0.168
	N	16	16	16	16
Amplitude production	Pearson Correlation	-0.243	-0.504	-0.066	-0.339
	Sig. (1-tailed)	0.182	0.023	0.405	0.099
	N	16	16	16	16
Bold type indicates correlation significant at p=0.0042 Bonferroni corrected significance level					

Table 5.2 *Pearson Correlation tests for CI subjects between appropriate F_0, duration and amplitude production and F_0, duration and amplitude perception. High and low F_0 ranges are combined and presented in a separate table (Mean F_0 thresholds).*

CI subjects		**Low F_0 threshold**	**High F_0 threshold**	**Duration threshold**	**Amplitude threshold**
F_0 production	Coefficient	-0.519	-0.655	-0.396	-0.331
	df	13	13	13	13
	P(1-tailed)	P= .024	**P= .004**	P= .072	P= .114
Duration production	Coefficient	-0.015	-0.488	-0.058	-0.123
	df	13	13	13	13
	P(1-tailed)	P= .479	P= .032	P= .419	P= .331
Amplitude production	Coefficient	-0.302	-0.570	-0.123	-0.456
	df	13	13	13	13
	P(1-tailed)	P= .137	P= .013	P= .332	P= .044
Bold type indicates correlations significant at p=0.0042 Bonferroni corrected significance level					

CI subjects		**Mean F_0 thresholds**	**Duration thresholds**	**Amplitude thresholds**
F_0 production	Coefficient	-0.6617	-0.4058	-0.3415
	df	13	13	13
	P(1-tailed)	**P= .004**	P= .067	P= .106
Duration production	Coefficient	-0.3274	-0.0548	-0.1196
	df	13	13	13
	P(1-tailed)	P= .117	P= .423	P= .336
Amplitude production	Coefficient	-0.5076	-0.1277	-0.4586
	df	13	13	13
	P(1-tailed)	P= .027	P= .325	P= .043
Bold type indicates correlations significant at p=0.0055 Bonferroni corrected significance level				

Table 5.3. *Partial correlations for the CI subjects between appropriate F_0, duration and amplitude production and F_0, duration and amplitude perception. High and low F_0 thresholds are averaged together and presented in a separate table (Mean F_0 thresholds).*

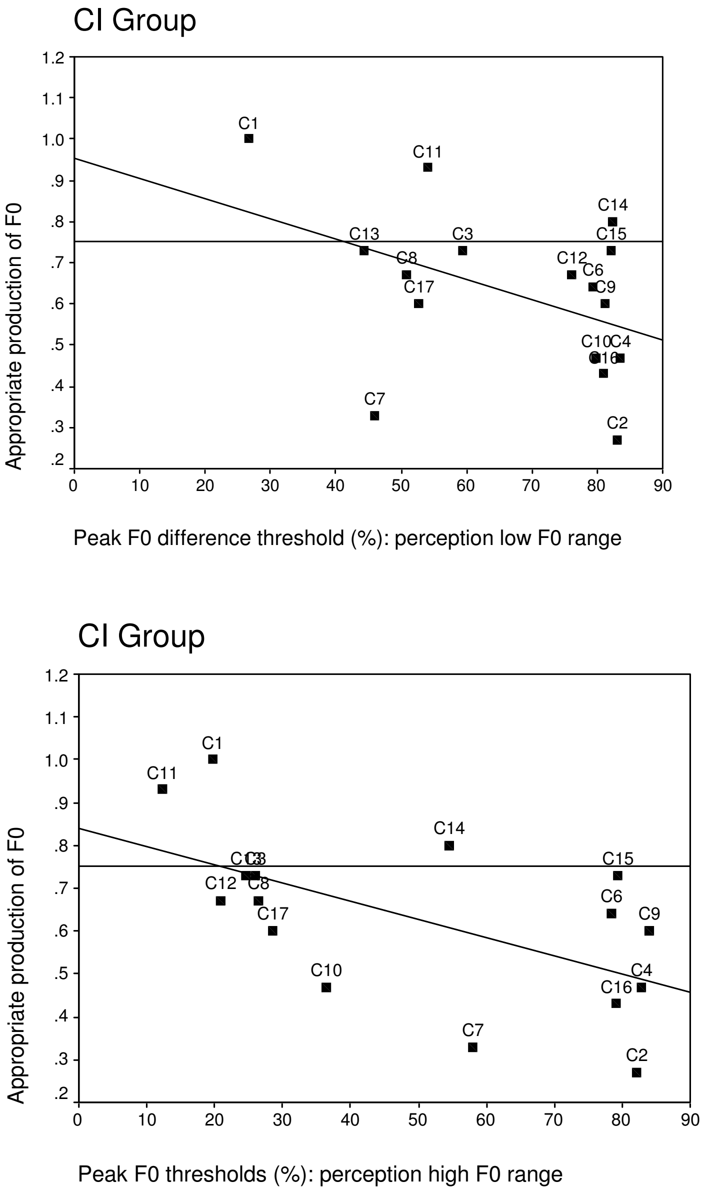


Figure 5.1 *Scattergraphs for individual CI talkers show appropriate production of* F_0 *(Experiment III) and peak* F_0 *difference thresholds (Experiment I) in the low* F_0 *range at the top of the figure and in the high* F_0 *range at the bottom.*

5.2.1.1 Production of F_0 in Experiment III vs. perception in the high F_0 range in Experiment I

Table 5.2 shows the correlation between the frequency of production of appropriate F_0 contours in Experiment III and the ability to perceive smaller peak F_0 differences between synthetic /baba/ bisyllables. When both F_0 ranges are combined (Mean F_0 thresholds) the correlation has a probability of 0.01 which does not reach a Bonferroni-corrected significance level. For high F_0 ranges only, there was a correlation that approached but did not reach a Bonferroni-corrected significance level. However, when age was controlled correlations between F_0 perception and production did reach Bonferroni-corrected significance levels as shown in Table 5.3 both for high and low F_0 range perception thresholds combined and for the high F_0 range thresholds only. Hence, CI talkers who were hearing smaller F_0 differences in the higher F_0 range (around 200 Hz) had more appropriate F_0 contours in the production of focus. Individual performances, however, presented in the scattergraph for the high F_0 perception range at the bottom of Figure 5.1 show that of the three CI talkers (C1, C11, and C14) who were significantly above chance in the production of appropriate F_0 contours, only two (C1, C11) could hear relatively small peak F_0 differences. The third (C14) could only hear F_0 differences greater than 0.5 octaves (i.e. 55%). In contrast with this the scattergraph also shows six other talkers (C3, C10, C13, C12, C8, C17) who were able to hear F_0 differences less than 0.5 octaves in the high F_0 range who did not make significant use of F_0 in the production of the target focus words. Although these six talkers could hear F_0 differences less than 0.5 octaves in the high F_0 range in Experiment I, they could not produce F_0 appropriately and consistently in the target focus words in Experiment III. It would seem from the above results that CI subjects' ability to produce F_0 appropriately is not necessarily linked with sensitivity to F_0 differences indicating once more that F_0 may not be a necessary cue to focus as stated in hypothesis (ii).

F_0 production range: CI talkers				
Talker	**F_0 range**	**Median (Hz)**	**Percentile 95 (Hz)**	**Percentile 05 (Hz)**
2	high	209	239	104
3	high	259	351	121
4	high	238	264	119
6	high	262	331	107
9	high	218	272	125
11	high	266	430	166
12	high	237	271	213
13	high	207	249	140
14	high	217	316	122
15	high	266	306	224
16	high	255	296	80
1	low	165	221	62
7	low	100	122	87
8	low	145	162	61
10	low	194	206	189
17	low	122	159	54

Table 5.4 *F_0 medians and 95th and 5th percentiles produced by the individual CI talkers in the production of Focus 3 sentences in Experiment III. F_0 medians were classified into high and low F_0 ranges in accordance with onset values for the high (i.e. onset 200 Hz) and low (onset 100 Hz) F_0 ranges in Experiment I stimuli*

5.2.1.2 Can CI talkers with a high F_0 production range perceive smaller F_0 differences within the same high F_0 range?

Table 5.4 shows that overall, eleven of the sixteen CI talkers had a high F_0 production range (i.e. median F_0 > 200 Hz corresponded to onset value for high F_0 range in Experiment I stimuli in section 2.2.2). It was considered that they might be able to hear smaller differences within their own F_0 production range. Figure 5.1, however, indicates that six talkers (C2, C4, C6, C9, C15, C16) could not consistently hear F_0 differences at or close to the maximum difference level of 84% in their own high F_0 production range. Production data for these six talkers, as summarized in Tables 5.1 and 5.4, did not show statistical evidence of appropriate F_0 production in the target focus words. However, other talkers (C3, C12, C13) with a high F_0 production range who were hearing smaller F_0 differences (of 25%, 20%, and 25%) did not make significant use of F_0 in production either. This would suggest that good perceptual abilities within their own F_0 production range do not necessarily mean that these talkers can make appropriate use of F_0 in the production of focus. Table 5.1 shows statistical evidence of consistency in appropriate F_0 production for two of the talkers (C11 and C14). While C11 showed a small F_0 difference threshold in the high F_0 range of 15%, C14 has a considerably larger threshold of 54%.

5.2.1.3 Production of F_0 in relation to perception in the low F_0 range

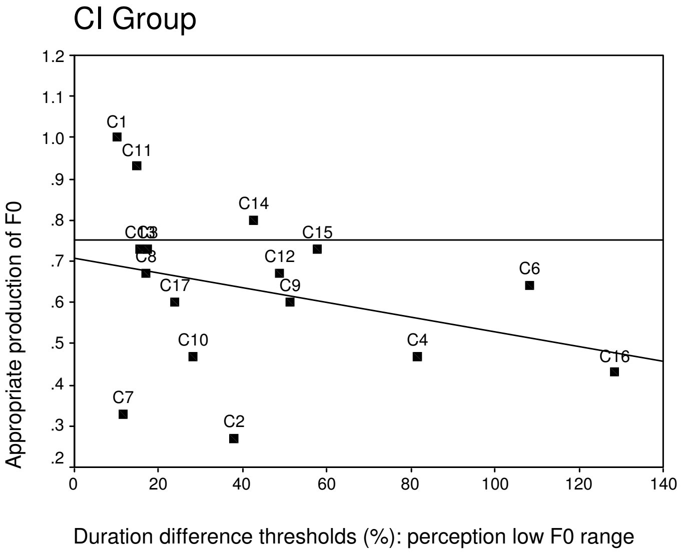
As discussed above Table 5.3 shows that although there was a correlation between appropriateness of F_0 production and the perception of peak F_0 in both F_0 ranges combined when age was controlled, there was no correlation with the low F_0 range when the two F_0 ranges were analysed separately. Consistent with this lack of correlation, the upper panel of top of Figure 5.1 shows that the three CI talkers (C1, C11 and C14) who were significantly above chance in the production of appropriate F_0 varied considerably in their perception of peak F_0 differences (thresholds of 25 %, 55%, and 84% respectively). That C14 shows such a high F_0 threshold suggests that the ability to use F_0 appropriately is not directly linked with perceptual sensitivity to F_0 in the low F_0 range. The scattergraph also shows that the rest of the CI talkers could only hear F_0 differences ranging between 45% and 84% in the low F_0 range and none of them made significant use of F_0 in production.

5.2.1.4 Do CI talkers with a low F_0 production range perceive smaller differences in the low F_0 range?

Five talkers (C1, C7, C8, C10, C17) had a low F_0 production range (i.e. median F_0 > 100 Hz which corresponded to onset value for the low F_0 range in Experiment I stimuli in section 2.2.2). Table 5.4 and the scattergraph at the top of Figure 5.1 shows that four of these talkers (C1, C7, C17, C8) could hear F_0 differences of 50% or less in their own low F_0 production range, and one talker (C10) whose production range was very narrow could not reliably hear differences at the maximum difference level (84%). Only one of these low F_0 production range talkers (C1) was able to hear relatively small F_0 differences (i.e. 25%) in his own low F_0 production range. Although four out of these five low F_0 range talkers were able to hear differences of 0.5 octaves or less within their own range only C1 was making appropriate use of F_0 in production. Although the co-existence of good F_0 perception and appropriate F_0 production in this one talker may suggest a direct linkage of the perception and production as in hypothesis (i), that conclusion cannot be upheld given the findings from the talkers with a higher F_0 production range.

5.2.1.5 What can we infer from the results about the relationship between perception and production of F_0?

Despite significant correlations as discussed above between perception and production of F_0 in the high F_0 range, only two of the three subjects (C1, C11 and C14) who made significant use of F_0 in production showed good F_0 perception. While subjects CI and C11 showed good F_0 perception, subject C14 showed F_0 thresholds of 54% (high range) and 82% (low range), yet he was able to make significant use of F_0 in production. In general, the results show no direct correspondence between the ability to perceive or produce F_0 for most CI subjects. Thus, six other talkers who were able to hear smaller F_0 differences than 0.5 octaves in the high F_0 range did not make significant use of F_0 in production. These results suggest that the ability to make appropriate use of F_0 in production does not necessarily depend on sensitivity to F_0. The relationship between the perception and production of F_0 is not straightforward and results seem to support the view in hypothesis (ii) that F_0 is not a necessary cue to linguistic focus. The other issue addressed above is whether the ability of CI children who perceive smaller differences within their own production range in the controlled experiment in Experiment I places them at an advantage in the production of appropriate F_0. Results so far suggest this is not necessarily the case.



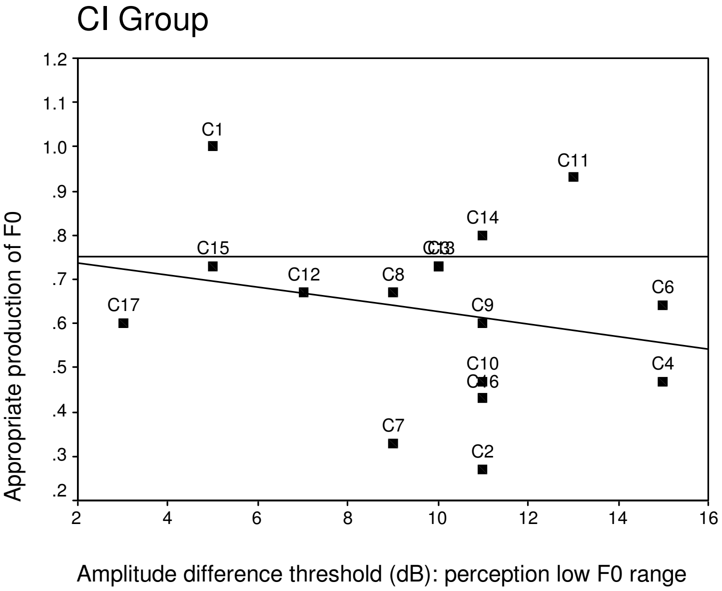


Figure 5.2 *Scattergraphs for individual CI talkers showing appropriate production of F_0 and duration difference thresholds (top panel) and amplitude difference thresholds (bottom panel).*

5.2.2 F_0 production in relation to duration and amplitude perception

As discussed in section 5.1 there may be differences between individuals' use of acoustic cues in the production of focus and F_0 may not be the most important cue to linguistic focus as suggested in hypothesis (ii). It appears from Experiment I that CI listeners might be able to rely on duration and/or amplitude cues in the perception and acquisition of some stress and intonation contrasts. For physiological reasons mentioned earlier for young children generally (i.e. tension associated with an interest in a focus word as discussed in section 1.3.2.4), it is possible that some CI talkers might be able to make significant use of F_0 in target focus words. If they are able to perceive differences in stress using only duration and/or amplitude cues and make appropriate use of F_0 in the production of target focus words a correlation might be expected between the appropriate use of F_0 in production (Experiment III) and duration and/or amplitude perception (Experiment I). Table 5.3, however, shows that there was no evidence of a correlation between the appropriate production of F_0 in Experiment III and duration or amplitude perception thresholds in Experiment I even when age was partialled out. Individual performances are discussed in more detail in the following sections.

5.2.2.1 F_0 production vs. duration perception

The scattergraph at the top of Figure 5.2 shows that eleven talkers could hear duration differences less than 45% but only three of them (C1, C11 and C14) who were able to hear duration differences of 10%, 15% and 42% respectively, made significant use of appropriate F_0 production in Experiment III. Despite perceptual sensitivity to duration differences less than 45% the remaining eight subjects varied in their ability to produce appropriate changes in F_0 with none performing above chance.

5.2.2.2 F_0 production vs. amplitude perception

The scattergraph at the bottom of Figure 5.2 shows that the group of CI subjects generally had a wide range of amplitude thresholds (3 dB - 15 dB). Of the three talkers significantly greater than chance in F_0 production, C1 showed a relatively small threshold of 5 dB, but the other two showed larger thresholds of 13 dB and (C11) and 11 dB (C14).

5.2.2.3 What can we infer from the results in 5.2.2 about the relationship between F_0 production and sensitivity to duration and amplitude differences?

As discussed above when age was partialled out no correlations were found between the production of appropriate F_0 and the perception of duration and amplitude differences. Despite sensitivity to duration differences less than 45% in Experiment I, eight of the CI talkers did not make significant use of F_0 in production. Amplitude thresholds in Experiment I varied for all CI subjects and were unrelated to significant use of F_0 in production. The scattergraphs in Figure 5.2 show that the few individual subjects who made significant use of F_0 in production varied in their ability to hear duration and amplitude differences, so we can conclude that the ability to make appropriate and consistent use of F_0 in the production of focus does not necessarily depend on their sensitivity to duration and amplitude. The results presented in Table 5.1 indicate that there are individual differences between acoustic cues used by CI subjects in the perception and production of focus.

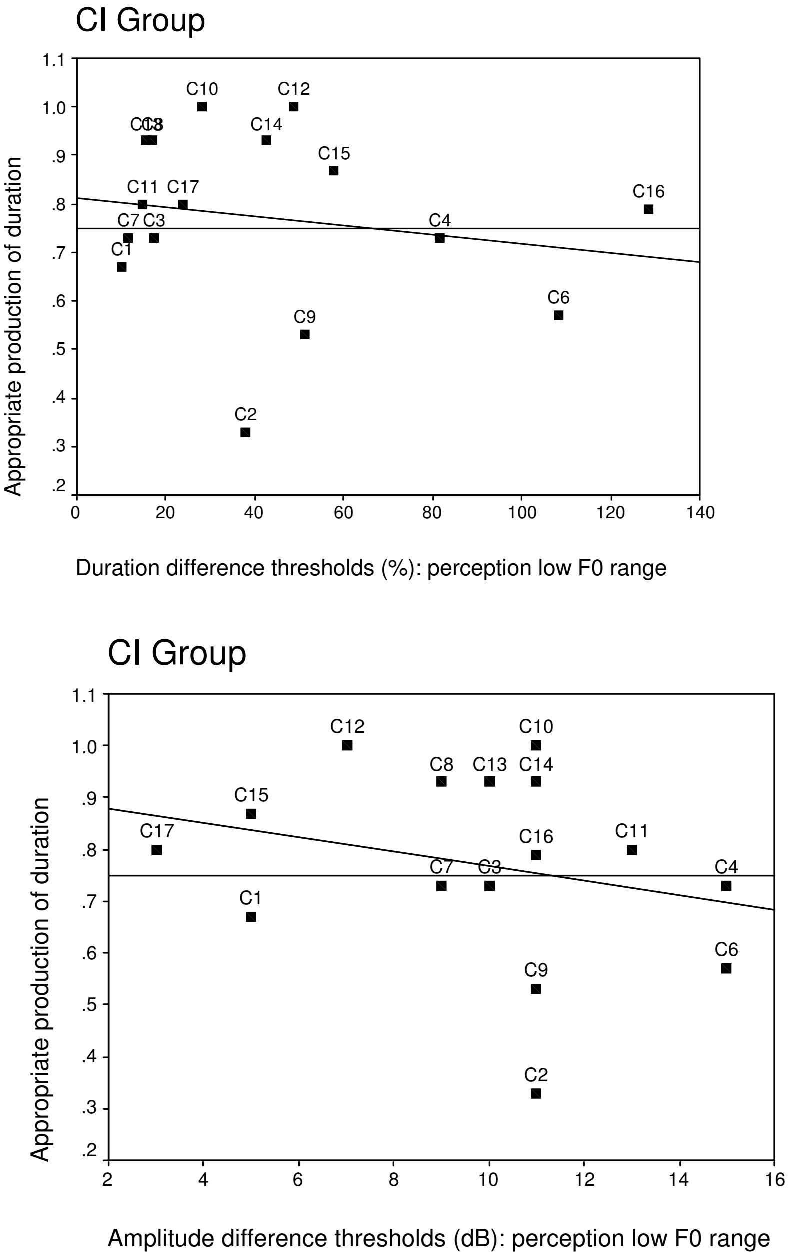


Figure 5.3 *Scattergraphs for individual CI talkers showing appropriate production of duration and duration difference thresholds at the top of the figure, and appropriate production of duration and amplitude difference thresholds at the bottom of the figure.*

5.2.3 Duration production in relation to duration, amplitude and F_0 perception.

The questions addressed in this section are whether

a. it is necessary for CI subjects to be able to hear differences in duration to be able to produce them appropriately

b. CI subjects who can use duration appropriately in production have better sensitivity to other cues such as amplitude or F_0

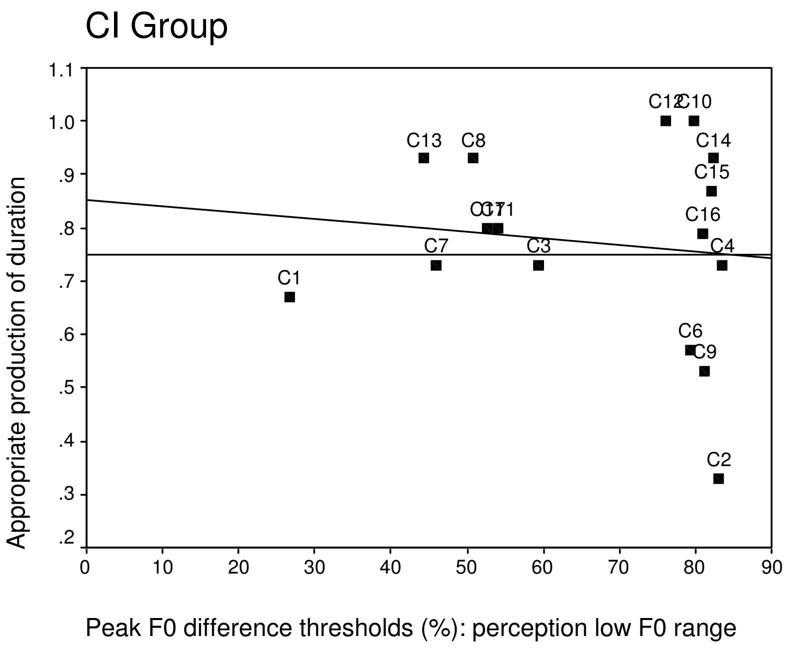
Since durational and amplitude cues might be more accessible than F_0 to implanted children than F_0, a correlation between production of duration and the perception of duration and amplitude cues might be expected. However, even when age was partialled out in Table 5.3 there was no correlation between the appropriateness of duration production in Experiment III and F_0, duration or amplitude perception thresholds in Experiment I.

5.2.3.1 Duration production vs. duration perception

As discussed above ability to hear smaller differences in F_0 in Experiment I by some CI talkers did not necessarily mean they could use F_0 appropriately in production so it is possible that they might make more significant use of a different cue i.e. duration in production. The scattergraph at the top of Figure 5.3 shows that nine CI talkers (C8, C11, C13, C17, C16, C10, C12, C14, C15) performed significantly better than chance in the production of appropriate duration in Experiment III and all except C16 could hear duration differences less than 60%. On the other hand, five other CI subjects who were able to hear duration differences less than 40% did not make a significant proportion of appropriate duration changes in production. It would appear that absence of appropriate durational changes in the production of focus for these other talkers cannot be explained simply by a lack of perceptual sensitivity to duration differences.

5.2.3.2 Duration production vs. amplitude perception

The scattergraph at the bottom of Figure 5.3 shows that nine CI talkers who performed significantly better than chance in duration production could hear amplitude differences ranging from 3 to 13 dB. However seven other talkers who showed no evidence of consistent appropriate duration production also varied in their ability to hear amplitude difference with thresholds ranging from 5 dB to 15 dB. So the absence of appropriate durational changes in production cannot be explained simply by lack of perceptual sensitivity to amplitude.



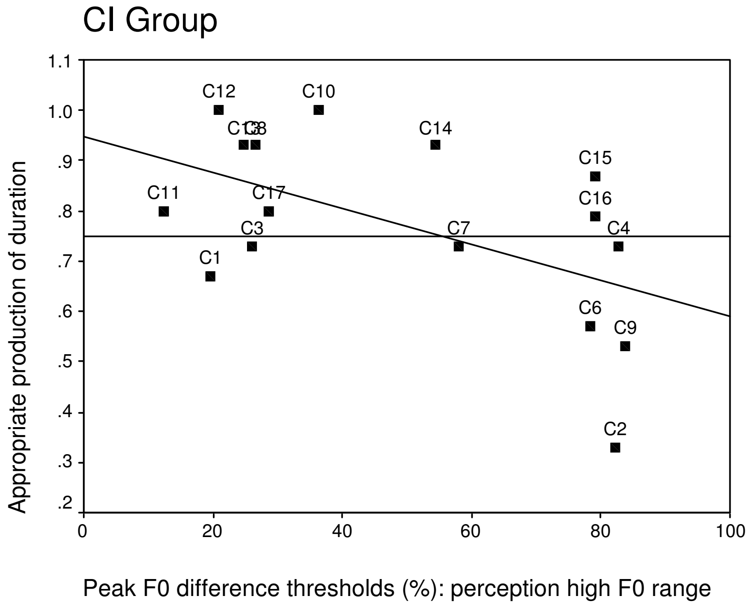


Figure 5.4 *Scattergraphs for individual CI talkers show appropriate production of duration and peak* F_0 *difference thresholds in the low and high* F_0 *ranges.*

5.2.3.3 Duration production vs. F_0 perception

In the scattergraph at the bottom of Figure 5.4 for the high F_0 perception range, four of the nine talkers who performed significantly better than chance in duration production (C8, C11, C12, C13) could hear peak F_0 differences less than 25%. In the low F_0 range (upper panel) four of the nine talkers, (C8, C11, C13, C17) could hear peak F_0 differences ranging between 45% and 55%. Five talkers (C10, C12, C15, C14, C16) in the low F_0 range and two (C15, C16) in the high F_0 range, were only hearing F_0 differences at or close to the maximum peak F_0 difference level of 84%, while these five were significantly better than chance in the production of duration. These results would suggest that significant use of duration in production by CI talkers is not necessarily associated with sensitivity to smaller F_0 differences.

5.2.3.4 What can we infer from the results in 5.2.3 about the appropriate use of duration in target focus word and sensitivity to duration, amplitude and F_0 difference?

Although there was no correlation between the appropriateness of duration production and duration perception thresholds, eight talkers who could hear duration differences of 60% or less (Table 5.1 and Figure 5.3) were able to make significant use of duration in production. However absence of appropriate durational changes in the production of focus for other CI talkers who were hearing differences of 45% or less cannot be explained simply by a lack of perceptual sensitivity to duration differences. No correlations were found between duration production and amplitude or F_0 thresholds even when age was partialled out and the wide range of amplitude thresholds and F_0 thresholds in Experiment I for CI subjects who made significant use of duration in production suggests that the appropriateness of duration production is not necessarily associated with the ability to perceive smaller amplitude (bottom of Figure 5.3) or F_0 differences (Figure 5.4). Overall, the wide variation in perceptual sensitivity to differences in F_0, duration or amplitude amongst individual CI subjects who made significant use of duration in production suggests that there is no direct link between the perception and production of duration. It would also appear from the results that individual subjects who use duration appropriately are not necessarily sensitive to the same perceptual cue(s).

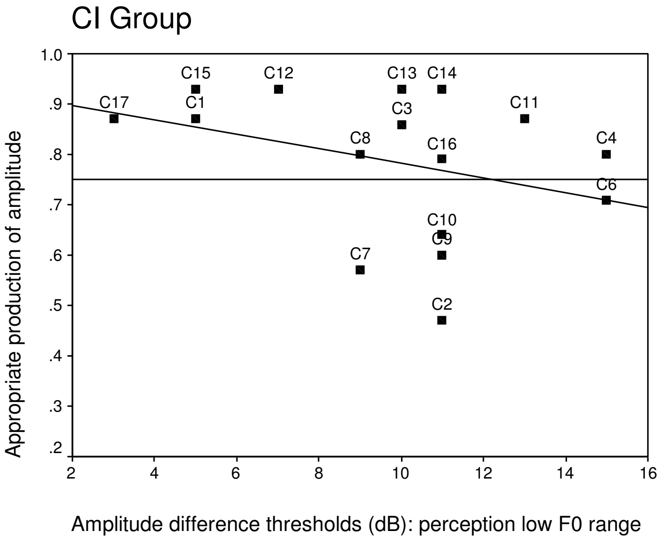
5.2.4 Amplitude production in relation to amplitude, duration and F_0 perception

The questions addressed below are whether

a. it is necessary for CI subjects to be able to hear differences in amplitude to be able to produce them appropriately

b. CI subjects who use amplitude appropriately in production are more sensitive to different cues such as F_0 or duration

The purpose of the correlation tests was to establish if the appropriate use of amplitude in production in Experiment III is linked with sensitivity to amplitude differences and/or duration and F_0 differences in Experiment I. Since results so far suggest that F_0 may not be a necessary cue to focus (see hypothesis (ii) in section 5.1.1) it is possible that CI subjects might respond better to duration or amplitude cues, so we might expect a correlation between amplitude production and duration or amplitude perception. A Pearson correlation test with partial correlations controlling for age at time of production (Tables 5.2 and 5.3) show that there was no correlation between the appropriate production of amplitude in Experiment III and amplitude, duration or F_0 thresholds**.** Individual performances are discussed in more detail below.



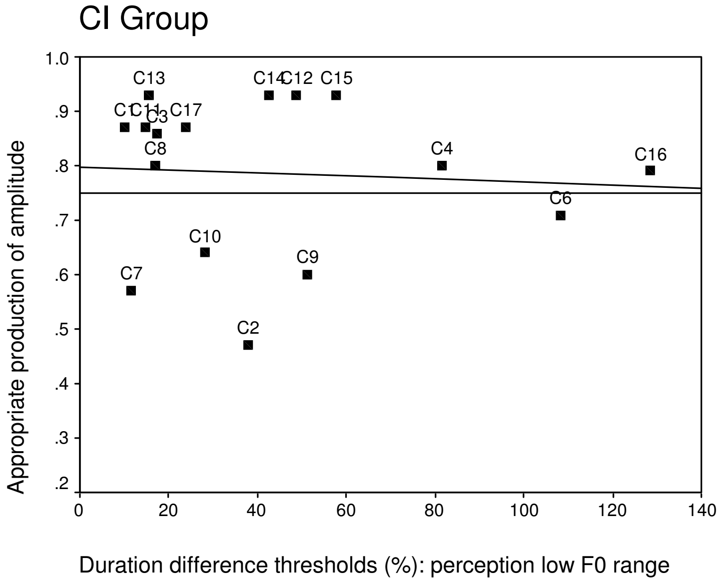


Figure 5.5 *Scattergraphs for individual CI talkers showing appropriate amplitude production and amplitude difference thresholds in the top panel and appropriate production of amplitude with duration difference thresholds in the bottom panel.*

5.2.4.1 Amplitude production vs. amplitude perception

The issue addressed in this section is whether CI subjects need to be able to hear amplitude differences in order to produce them. Individual scores presented in the scattergraph at the bottom of Figure 5.5 also show that the eleven CI talkers (C1, C3, C4, C8, C11, C12, C13, C14, C15, C16, C17) who performed significantly above chance (0.75 or 0.76) in the production of appropriate amplitude, varied in their ability to hear amplitude differences i.e. between 3 dB and 15 dB. However, only three of them (C17, C15, C1) could hear amplitude differences of 5 dB or less and the other six talkers could only hear amplitude differences greater than 7 dB. The limited perception of amplitude differences shown by these subjects suggests that their ability to use amplitude in production is not mediated by direct auditory feedback.

5.2.4.2 Amplitude production vs. duration perception

The scattergraph at the top of Figure 5.5 shows that nine of the eleven CI talkers who performed significantly greater than chance in amplitude production were hearing duration differences less than 60%. This suggests duration might be a more reliable cue than amplitude for these particular talkers.

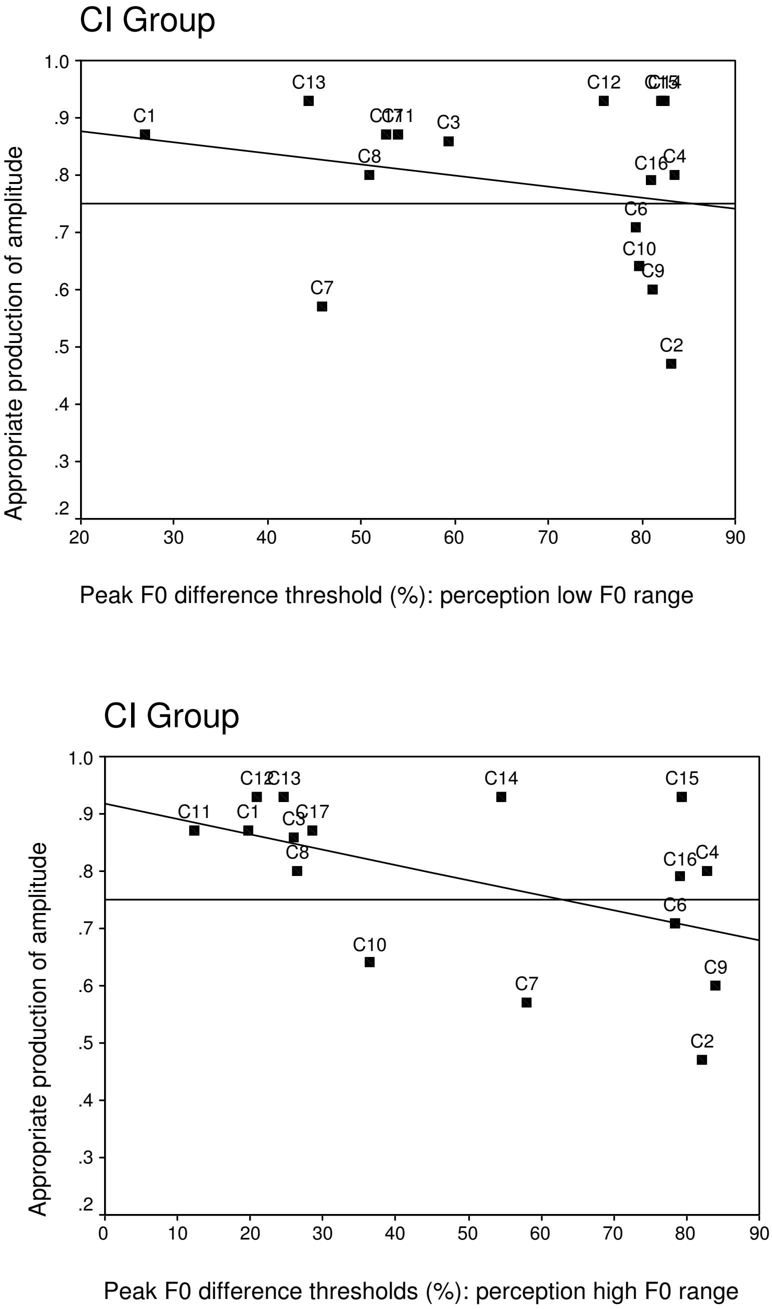


Figure 5.6 *Scattergraphs for individual CI talkers showing the appropriate production of amplitude and peak F_0 difference thresholds in the low and high F_0 ranges.*

5.2.4.3 Amplitude production vs. F_0 perception

Individual scores in the scattergraph on the bottom of Figure 5.6 show that seven (C1, C3, C8, C11, C13, C12, C17) of the eleven talkers who showed above chance rates of appropriate amplitude changes in production could hear F_0 differences in the high F_0 range between 10% and 30%. In the low F_0 range at the top of the figure, only one of the talkers (C1) could hear F_0 differences of less than 45% (a little more than 0.5 of an octave). Figure 5.6 also shows that the remaining five talkers (C4, C12, C14, C15, C16) in the low F_0 range and three (C4, C15, C16) in the high F_0 range who were significantly above chance in the frequency of appropriate production of amplitude could not consistently hear peak F_0 differences at or close to the maximum difference level (84% or almost an octave). Although sensitivity to F_0 changes in the high F_0 range may be linked to the appropriate use of amplitude in production for seven CI talkers in the high F_0 range, it does not appear to be the case for the rest of the subjects.

5.2.4.4 What can we infer from the results about the ability to make appropriate use of amplitude and sensitivity to F_0, duration, and amplitude cues?

The wide range of sensitivity to amplitude differences amongst those who were able to make appropriate use of amplitude in production suggests that ability to use amplitude appropriately does not necessarily depend on sensitivity to amplitude differences. Results show that CI subjects who made appropriate use of amplitude seem to be more sensitive to duration cues and in some cases to F_0 cues in the high F_0 range only. Overall, the results indicate that duration might be a more reliable perceptual cue than amplitude or F_0 for CI subjects who were able to make consistent use of amplitude in production.

5.2.5 Summary

The results in section 5.2 above indicate that CI subjects may be sensitive to one or more cues as presented in controlled synthetic bisyllables in Experiment I but use different cues in production in Experiment III, and they are summarized below.

a. F_0 production vs. sensitivity to differences in F_0, duration and amplitude

When age was partialled out a negative correlation was found between F_0 thresholds in the high F_0 range and appropriate production of F_0 (Table 5.3). As discussed in 5.2.1 individual scores in the scattergraphs (Figure 5.1) are not consistent with a direct relationship between the ability to produce or perceive differences in F_0. Overall, only three subjects (C1, C11, C14) were able to make appropriate use of F_0 in production and these varied in their ability to hear F_0, duration and amplitude differences (Table 5.1 and sections 5.2.1 and 5.2.2). Sensitivity to F_0, duration and amplitude differences seemed to vary regardless of whether CI subjects made significant use of F_0 in production.

b. *Duration production vs. sensitivity to differences in duration, amplitude and* F_0

No correlations were found between appropriate production of duration and the perception of duration, amplitude or F_0 even when age was controlled (Tables 5.2 and 5.3). A wide variation in perceptual sensitivity to F_0, duration or amplitude differences was found for individual CI listeners (section 5.2.3 and scattergraphs in Figure 5.3 and 5.4) regardless of whether they could make appropriate use of duration in production.

c. *Amplitude production vs. sensitivity to differences in amplitude,* F_0*, and duration*

There were no correlations between the production of appropriate amplitude and, the perception of duration, amplitude, or F_0 differences even when age was partialled out (Tables 5.2 and 5.3). The wide range of sensitivity to differences in amplitude and F_0 for those who could produce amplitude appropriately suggests that amplitude production does not necessarily depend on ability to hear smaller differences in amplitude or F_0 (section 5.2.4 and scattergraphs in Figures 5.5 and 5.6). However, since nine of the eleven subjects who could use amplitude appropriately were able to hear duration differences less than 60%, duration might be a more reliable perceptual cue.

The next section explores the relationship between amplitude, duration and F_0 production in Experiment III and the perception of linguistic focus in Experiment II. Acoustic measurements of the Focus 3 stimuli in Experiment II (Appendices 3.2 – 3.9) combined with F_0, duration and amplitude thresholds in Experiment I will indicate whether duration or amplitude or F_0 are reliable cues to linguistic focus for CI subjects.

5.3 Are there correlations between the production of F_0, duration and amplitude and the perception of linguistic focus?

The question set out in section 5.1 above is whether it is necessary to be able to perceive focus in Experiment II in order to use it appropriately and consistently in Experiment III using one or more acoustic cues (F_0, duration, amplitude) on target focus words. To address this question, Pearson correlation tests were carried out to establish for the CI children (aged between 5;7 and 17;1 years) whether there is any statistical link between ability to make appropriate use of F_0, duration or amplitude in target focus words in Experiment III and the ability to perceive focus in the same target words in Experiment II. Although the acoustic cues are not controlled in the linguistic focus stimuli in Experiment II, measurements of the differences in F_0, duration and amplitude between target focus words and neighbouring words for the stimuli (Appendices in Chapter Three) can give some indication of which acoustic cues are likely to be accessible to CI listeners in the light of their F_0, duration and amplitude thresholds in Experiment I, and they are taken into consideration in the discussion below.

CI subjects: **Pearson Correlations**		**Focus Perception**
F_0 production	Pearson Correlation	0.342
	Sig. (1-tailed)	0.098
	N	16
Duration production	Pearson Correlation	0.526
	Sig. (1-tailed)	*0.018*
	N	16
Amplitude production	Pearson Correlation	0.323
	Sig. (1-tailed)	0.111
	N	16
Bold type indicates correlations significant at p=0.0167 Bonferroni corrected significance level		

CI subjects: **Partial Correlation Coefficients controlling for age at Experiment II**		**Focus Perception**
F_0 production	Coefficient	0.535
	df	-13.000
	P (1-tailed)	*P= .020*
Duration production	Coefficient	0.448
	df	-13.000
	P (1-tailed)	P= .047
Amplitude production	Coefficient	0.523
	df	-13.000
	P (1-tailed)	*P= .023*
Bold type indicates correlations significant at p=0.0166 Bonferroni corrected signifcance level		

Table 5.5 *Pearson correlations for production measures compared to focus perception for CI subjects. Partial correlations controlling for age are presented at the bottom of the table.*

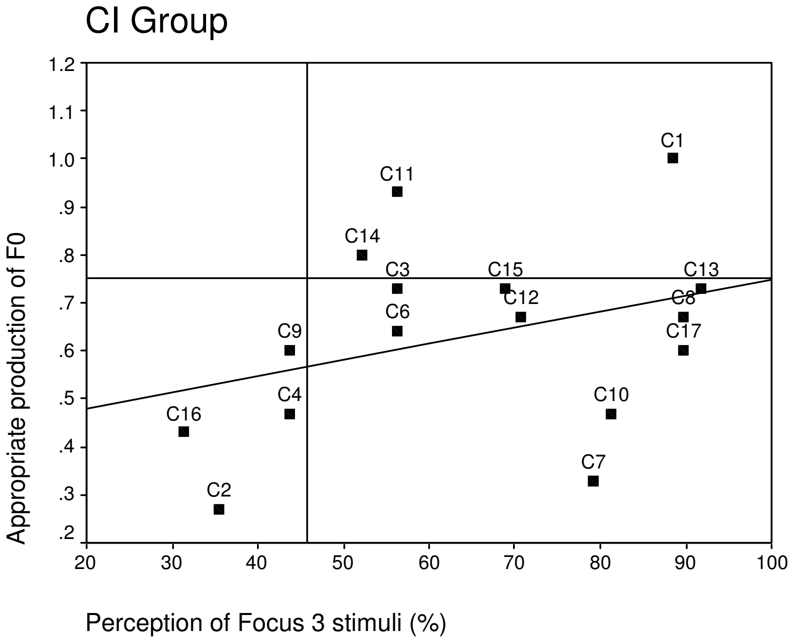


Figure 5.7 *Scattergraph for individual CI talkers showing the appropriate production of* F_0 *in Experiment III and focus perception scores in Experiment II.*

5.3.1 F_0 production in relation to the perception of focus

When age was controlled in the partial correlations (See Table 5.5) the correlation between the production of appropriate F_0 contours in Experiment III and the perception of linguistic focus in Experiment II had a p value of 0.02 which was approaching significance compared to a Bonferroni–corrected significance level of p = 0170. The scattergraph in Figure 5.7 and individual scores in Table 5.1 indicate, however, that only three talkers (C1, C11 and C14) showed statistical evidence of appropriate F_0 production. Although they were significantly better than chance in the perception of focus, individual performances in Experiment II for these subjects varied (89%, 56% and 52% respectively*)*. Figure 5.7 and Tables 5.1 and 5.6 also show that nine other individual talkers (C3, C6, C15, C12, C13, C8, C17, C10 and C7) who did not make significant use of appropriate F_0 in production also performed significantly above chance in the perception of focus with scores ranging from 45% up to 90%. In other words, these nine subjects could hear focus on the appropriate target word more often than expected by chance but did not make significant use of F_0 in the production of focus, although three of these nine (C3, C13, C15) showed rates of appropriate F_0 production that were very close to the adopted significance level of 0.75 (see underlined in Table 5.6)

As discussed in section 3.5.4.1 F_0 differences between target focus words and neighbouring words, which rarely exceeded 0.5 octaves in the perception stimuli, would have been inaccessible to the nine listeners. It is possible that they were relying on other acoustic cues which suggests that F_0 is not a necessary cue to focus as stated in hypothesis (ii). The wide variation in sensitivity to duration and amplitude regardless of ability to make appropriate use of F_0 (Table 5.6) in Experiment III indicates that duration and amplitude changes in the focus stimuli might have been inaccessible to some listeners.

	Experiment III	***Experiment I***				***Experiment II***
	Appropriate F_0 production	***Amplitude Thresholds (dB)***	***Duration Thresholds (%)***	***High F_0 Range (%)***	***Low F_0 range (%)***	***Focus 3 Perception (%)***
CI subjects	***At or below chance or approaching significance level (0.75)***					***Significance level = 45.8%***
C3	0.73	10	17	26	59	***56***
C6	0.65	15	108	78	79	***56***
C7	0.33	9	11	58	46	***79***
C8	0.67	9	17	27	51	***90***
C10	0.47	11	28	36	80	***81***
C12	0.67	7	49	21	76	***71***
C13	*0.73*	10	15	25	44	***92***
C15	0.73	5	58	79	55	***62***
C17	0.60	3	24	29	53	***90***
	Significantly greater than chance (0.75)					
C1	***1.00***	5	10	20	27	***89***
C11	***0.93***	13	15	12	54	***56***
C14	***0.80***	11	43	54	82	***52***

Table 5.6 *Summary of CI talkers' appropriate production of F_0 (Experiment III), F_0, duration and amplitude thresholds (Experiment I), and the perception of focus (Experiment II).*

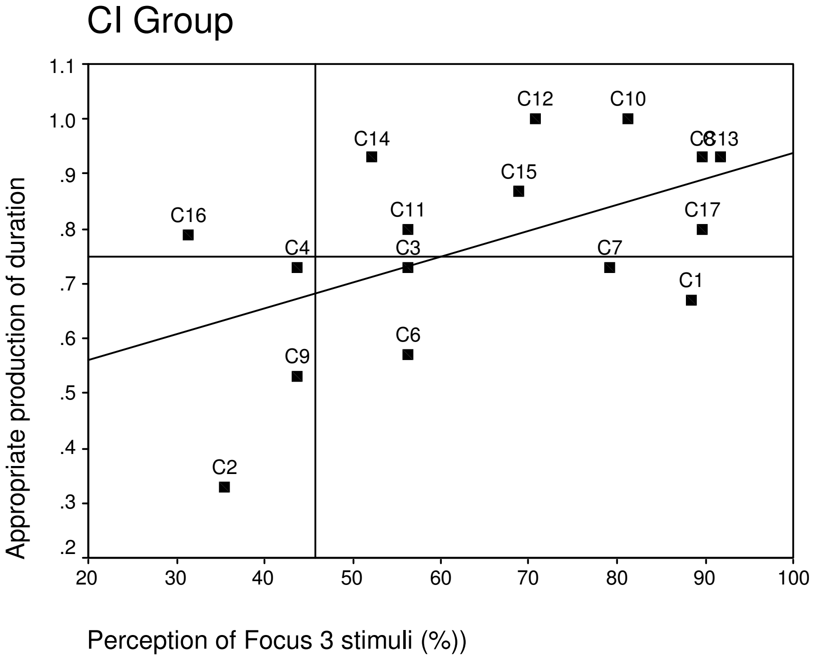


Figure 5.8 *Scattergraph for CI talkers showing appropriate production of duration in Experiment III and the perception of Focus 3 stimuli in Experiment II.*

5.3.2 Duration production in relation to the perception of Focus

As indicated in Table 5.5 when age was controlled a correlation which was approaching significance disappeared between the production of duration in Experiment III and the perception of linguistic focus in Experiment II. This would suggest the perception of linguistic focus and the appropriate production of duration improve together with increasing age. The scattergraph in Figure 5.8 and individual subjects' scores presented in Table 5.1 and Table 5.7 show that eight of the nine CI talkers (C8, C10, C11, C13, C12, C14, C15, C16, C17) who showed statistical evidence of appropriate duration production in Experiment III were significantly above chance (45.8%) in the perception of focus in Experiment II. However, performance for these subjects varied ranging between 52% and 90%. Despite the ability to use duration appropriately in production one of these subjects (C16) performed below chance (31%) in the perception of focus. On the other hand there were four other talkers (C1, C3, C6, C7) who did not consistently produce appropriate durational changes in production yet performed above chance in the focus perception test (89%, 56%, 56%, 79% respectively). As mentioned above C16 performed poorly in the perception of focus but was able to make consistent use of duration in production. This subject could only hear very big durational differences (128%) in Experiment I and some of the duration differences mentioned above between target and neighbouring words (***BOY*** and ***DOG*** above) which were less than 128% would have been inaccessible to C16 in the focus words. Table 5.1 shows that this subject was at an additional disadvantage in the perception of focus in Experiment II as he could only hear large F_0 differences (81% and 79% in the low and high F_0 ranges respectively) and amplitude differences of 11 dB in Experiment I so may not have been sensitive to any cues. As discussed in section 3.5.4.2 the target words in the perception stimuli, which were longer when in focus in three of the four sentences (i.e. 75% - 140%), should have been accessible to the other listeners since the median duration for the CI group was 35%.

Table 5.7 shows that for the nine CI subjects who made appropriate use of duration in production in Experiment III and performed significantly greater than chance in the perception of focus in Experiment II, there was a wide range of sensitivity to amplitude and F_0 differences in both F_0 ranges in Experiment I. It would appear that those subjects who make appropriate use of duration in production and perform well in the perception of linguistic focus seem to have better sensitivity to durational cues than amplitude or F_0. These results support the view that F_0 is not a necessary cue to focus in hypothesis (ii) in sections 1.1.2 and 1.11.4.

	Experiment III	***Experiment I***				***Experiment II***
	Duration Production	***Duration Thresholds (%)***	***Amplitude Thresholds (dB)***	***High F_0 range (%)***	***Low F_0 range (%)***	***Focus 3 Perception (%)***
CI subjects	***Significantly greater than chance (0.75)***					***Significantly greater than chance (45.8%)***
C8	***0.93***	17	9	20	27	***90***
C10	***1.00***	28	11	36	80	***81***
C11	***0.80***	15	13	12	54	***56***
C12	***1.00***	49	7	21	76	***71***
C13	***0.93***	15	10	25	44	***92***
C14	***0.93***	43	11	54	82	***52***
C15	***0.87***	58	5	79	55	***62***
C16	***0.78***	128	11	79	81	*31*
C17	***0.80***	24	3	29	53	***90***
	Just above chance or approaching significance (0.75)					
C1	0.67	10	5	20	27	***89***
C3	0.73	17	10	26	59	***56***
C6	0.57	108	15	78	79	***56***
C7	0.73	11	9	58	46	***79***

Table 5.7 *Summary of CI talkers' appropriate production of duration (Experiment III), duration, amplitude and F_0 thresholds (Experiment I), and the perception of focus (Experiment II)*

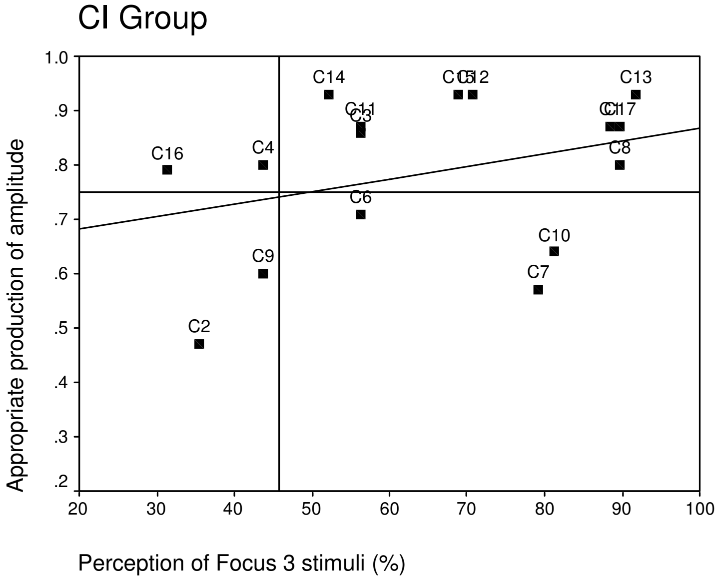


Figure 5.9 *Scattergraph for CI talkers showing amplitude production and the perception of Focus 3 stimuli.*

5.3.3 Amplitude production in relation to the perception of focus in Experiment II

Pearson Correlations tests were carried out to establish if there is a relationship between the appropriate use of amplitude in the production of focus and the ability to hear linguistic focus on target words. As Table 5.5 shows when age was partialled out the correlation was approaching Bonferroni-corrected significance with $p = 0.023$. Table 5.1 and the scattergraph in Figure 5.9 show that nine of the eleven talkers (C1, C3, C8, C11, C12, C13, C14, C15, C17) who were significantly above chance in the production of appropriate amplitude in Experiment III were significantly above chance (45.8%) in the perception of linguistic focus in Experiment II. These talkers had good focus perception skills and made significant use of amplitude in the production of focus, but as presented in Table 5.8 below scores varied widely.

The scattergraph in Figure 5.9 and Table 5.8 also show that there were three other talkers (C6, C7, C10) who were significantly above chance in the perception of focus yet did not show statistical evidence of appropriate amplitude production. Some of the subjects who had high amplitude difference thresholds performed well in the perception of focus (C 8, C13, C11, C7 in Table 5.1). Amplitude differences in the perception stimuli varying from <1 up to 10dB (section 3.5.4.3) might be less accessible to them than duration cues. Table 5.8 shows that nine CI subjects who were making appropriate use of amplitude could hear duration differences of less than 60% so it is likely that increased duration for target focus words (75% - 140%) in three of the four stimulus sentences in Experiment II (mdc, bpb, deb) were more accessible to these talkers in Experiment II. Table 5.8 above shows that seven of the nine talkers significantly above chance in amplitude production could hear F_0 differences in the high F_0 range between 10% and 30%. However, since there were other subjects making appropriate use of amplitude who could only hear F_0 differences in both F_0 ranges at or close to the maximum difference level (84%) in both F_0 ranges, F_0 may not always be a reliable cue for these listeners.

Overall, Experiment I results seem to suggest that CI listeners who could make appropriate use of amplitude in Experiment III and scored significantly above chance in the perception of focus in Experiment II were able to rely more on duration rather than amplitude or F_0 cues. These results would also support the view that F_0 is not a necessary cue to stress and intonation (see hypothesis (ii)).

	Experiment III	***Experiment I***				***Experiment II***
	Amplitude Production	***Amplitude thresholds (dB)***	***Duration Thresholds (%)***	***F_0 High Range (%)***	***Low F_0 range (%)***	***Focus 3 Perception (%)***
CI subjects	***significantly greater than chance (0.75)***					***Significantly greater than chance = 45.8%***
C1	***0.87***	5	10	20	27	***89***
C3	***0.87***	10	*17*	26	59	***56***
C8	***0.80***	9	*17*	27	51	***90***
C11	***0.87***	13	*15*	12	54	***56***
C12	***0.93***	7	*49*	21	76	***71***
C13	***0.93***	10	*15*	25	44	***92***
C14	***0.93***	11	*43*	54	82	***52***
C15	***0.93***	5	*58*	79	55	***62***
C17	***0.77***	3	*24*	29	53	***90***
C6	0.70	15	*108*	78	79	***56***
C7	0.60	9	*11*	58	46	***79***
C10	0.67	11	*28*	36	80	***81***

Table 5.8 *Summary of CI talkers' appropriate production of amplitude (Experiment III), amplitude, duration and F_0 thresholds (Experiment I), and the perception of focus (Experiment II).*

CHAPTER SIX

CONCLUDING CHAPTER:

DISCUSSION AND CONCLUSIONS

6.1 Discussion and conclusions

6.1.1 The relationship between the skills tested in Experiments I and II and III:

6.1.1.1 Is F_0 discrimination related to perception of linguistic focus and phrase/compound contrasts?

A significant correlation was found when age was partialled out between F_0 thresholds in both the high and low F_0 ranges in Experiment I and scores in the Focus 2 and Focus 3 tests both individually and combined together (MFocus) in Experiment II (section 3.5.4.1). This suggests that perception of linguistic focus depends on the ability to hear smaller differences in F_0. However, more detailed analysis of the results in Table 5.1 shows that the majority of CI subjects were able to perceive linguistic focus in the Focus 3 test at a level which was significantly greater than chance despite the fact that most of them were unable to hear F_0 differences less than 0.5 octaves (2.3). Some subjects could not consistently hear differences in F_0 even at the maximum difference level of 84 % yet performed well in the perception of focus which suggests they may be relying on other cues such as duration and amplitude. As discussed in section 3.5.4.1, the measurements for the Focus 3 perception stimuli in Experiment II show that the median semitone differences between target focus words and neighbouring words were generally less than 0.5 octaves and so these differences would not be accessible to most CI subjects. Although performance varied between individual subjects the perception of focus might not necessarily depend on the ability to hear F_0 differences. Rather the results seem to support the view that F_0 is not a necessary cue to focus and implant users might be more sensitive to other cues such as duration and/or amplitude (hypothesis (ii)).

Although in the literature F_0 has frequently been regarded as the most important perceptual cue to stress and intonation, the present results do not fit that view in common with some other recent studies of normal hearing subjects. For example, Kochanski et al. (2005) found that F_0 played a more minor part than loudness and duration in their study of prominence in young adults although they did not make a distinction between contrasts such as lexical stress or focus in their analysis. Peppé et al. (2000), however, do make this distinction and report in their study of adults that pitch movement or pitch reset might not be as reliable as loudness and duration at signalling compounds vs. phrase stress. They also suggest that there may be differences in the use of acoustic cues by adult speakers in the realisation of intonational contrasts in less controlled settings compared to laboratory conditions.

In the present study the linguistic stimuli for the perception tests in Experiment II (and also the production data in Experiment III) were not laboratory controlled and were elicited in as natural a context as possible in order to obtain consistent measurable responses using a set of questions based on a picture. If F_0 only plays a minor role in the perception of stress and intonation, as suggested by Kochanski et al. and Peppé et al., or if individual subjects vary in their use of acoustic cues, it is possible that CI children are not at a disadvantage due to poor pitch perception in the early acquisition stages of prosodic development. The detailed analysis undertaken in the current study has not been carried out previously for English speaking children with cochlear implants and further investigations need to be carried out in the future for different regional variations. However, some studies of studies of children using hearing aids (Rubin Spitz and McGarr, 1990; Murphy, McGarr and Bell-Berti, 1990; Most, 1999) also suggest that correctly perceived stress and intonation patterns may be produced using different acoustic correlates or that there may be conflicting cues such as duration or intensity which might affect listeners' perception of F_0 (section 1.11.2). It is difficult to draw comparisons between CI users and hearing aid users because of device limitations (section 1.7), and since limited F_0 information is delivered via the speech processor implant users are more likely to be reliant on duration and amplitude cues.

6.1.1.2 Is F_0 discrimination related to appropriate production of F_0 in target focus words?

Of the four of the sixteen CI subjects who consistently managed to convey focus to a trained listener (the investigator) only one made appropriate use of F_0 in the production of target focus words (Table 5.1). As discussed in section 4.4.4., CI subjects sometimes sounded ambiguous as a result of insufficient boosting of F_0 (or insufficient increases in duration or amplitude) on the target words. However, according to Wells et al. (2004) ambiguity is not uncommon in normal hearing children and adult speakers of English (section 1.11.1 and 4.4.4*)*. This needs to be borne in mind when drawing conclusions from the current and any future investigations of prosodic development of children with implants. Although significant correlations were found (Table 5.3) between the production and perception of F_0, results for individual subjects in Table 5.1 show that the few subjects who could hear smaller differences in F_0 in controlled conditions in Experiment I did not necessarily make appropriate use of F_0 in the production of focus in Experiment III (section 5.2). Therefore, there does not seem to be a direct relationship between perception and production of F_0 in the current results, and ability to hear smaller F_0 differences within a child's own production range was not necessarily an advantage for the CI subjects (sections 5.2.1.2 and 5.2.1.4).

6.1.1.3 Are duration and amplitude discrimination related to the perception of linguistic focus and phrase/compound contrasts?

If pitch adjustments in speech directed at young children (Jusczyk, 1997; Cruttenden, 1994) are not accessible to implanted children, other prosodic cues such as slower articulation, differences in loudness, longer pauses, and paralinguistic cues such as eye contact, gestures, jumping up and down, reaching (Crystal 1986; Snow and Balog, 2002) should help draw attention to certain features such as response required, rhythm or focus (section 1.11.1 and hypothesis (ii) in section 1.1.2).

Duration

Since the median *duration* difference threshold for the group of CI listeners in Experiment I was 35%, duration might provide a more reliable cue than F_0 to linguistic focus and compound vs. phrase stress in Experiment II (section 3.5.4.2). Measured duration measurements for the focus stimuli (Appendices 3.5 and 3.6) ranged from 75% up 140% longer when in focus in most of the stimulus sentences in Experiment II (section 3.5.4.2) so these differences should be accessible to the implanted subjects. Tables 3.4 and 3.6 show that correlations between the ability to hear smaller duration differences remained for Focus 2 test when age was partialled out so performance in this test was linked with ability to hear differences in duration. A correlation between duration and Focus 3 test scores disappeared when age was controlled for suggesting that performance in these tests improve with increasing age.

Amplitude

The median *amplitude* difference thresholds for the CI subjects in Experiment I was 11 dB, so many of the amplitude changes in target focus words in the Focus 3 stimuli in Experiment II (Appendix 3.8) would often not have been accessible to them. These results suggest that good perceptual ability in the focus test is not necessarily accounted for by sensitivity to differences in amplitude. Moreover, good performance in the perception of linguistic focus even amongst subjects with large amplitude thresholds (Figure 3.7) suggests that good prosodic perception ability could not be entirely due to amplitude cues, and duration might provide a more reliable cue (Figure 3.6). A correlation which was approaching significance between amplitude discrimination and performance in Focus 3 test disappeared when age was partialled out which suggests unconnected abilities improve together with increasing age. Although Table 3.3 shows that correlations were found between age at time of testing, age at switch-on and performance in the Focus 3 test, no correlations were found between age and amplitude discrimination (Table 2.6).

6.1.1.4 Is it necessary for CI subjects to be able to hear duration and amplitude in order to produce them appropriately in target focus words?

No correlations were found between perception and production of duration or amplitude (Tables 5.2 and 5.3). For those who made significant use of *duration* in production, variation was found across subjects in perceptual sensitivity to F_0, duration and amplitude differences in Experiment I (section 5.2.3). The absence of appropriate duration changes in production for some talkers cannot be explained simply by lack of perceptual sensitivity to duration differences. The wide range in *amplitude* thresholds for those who produced amplitude appropriately, suggests that the ability to make appropriate use of amplitude in target focus words does not depend on perceptual sensitivity to amplitude (section 5.2.4). Nine of the eleven CI subjects who made significant use of amplitude could hear duration differences of less than 60%, while seven were sensitive to half – octave or smaller differences in the high F_0 range. This suggests that it might be possible for them to perceive focus using one or more cues (e.g. duration or F_0) and make appropriate use of a different cue (i.e. amplitude) in the production of target focus words. The results support the view that F_0 is not a necessary cue to focus (hypothesis (ii) in section 1.1.2) and indicate that CI children should be able to acquire abstract phonological representations of prosodic contrasts such as tonicity and focus using whatever acoustic cues are available to them through the implant.

6.1.2. *The relationship between the perception and production skills tested in Experiment II and Experiment III*

6.1.2.1 Is it necessary to be able to perceive focus in order to realize focus by making appropriate and significant use of one or more acoustic cues?

The consistent use of appropriate F_0 contours or appropriate increases in duration and/or amplitude in the production of target focus words in Experiment III might suggest that CI talkers have developed an abstract awareness of focus, although in some cases the increases or changes in these cues on the target words may be insufficient to convey focus to a listener. This is borne out by the current investigator's impression (section 4.4.4) that some talkers sounded ambiguous and the impression of focus was conveyed consistently by only four out of the sixteen CI talkers (section 4.4.3). For the purpose of the following discussion we can assume that if the CI subjects made significant use of any of these cues (F_0, duration and amplitude) in production, they have probably developed an abstract representation of this concept. The results indicate that subjects who are less consistent but approaching significance level are probably still in the process of acquiring the concept of focus.

6.1.2.2 Individual performances by CI subjects

The question addressed in section 5.3 is whether it is necessary to be able to perceive linguistic focus in order to realize it in production. As discussed in section 1.4.1 an increase in subglottal pressure from the lungs raises amplitude and also partly controls vocal fold vibration (F_0) so when F_0 is increased it is usually accompanied by an increase in amplitude. Duration, on the other hand, seems to be a more independent cue although it is rare for F_0 peaks to be realised on a very short syllable. Experiment III results can tell us whether CI subjects use one or more acoustic cues appropriately on target focus words, and when age was partialled out correlations of the perception of focus with the production of F_0 and amplitude approached significance but that with the production of duration did not (Table 5.5)

In general, most of the CI subjects could perceive linguistic focus, and most used at least one acoustic cue appropriately in production. Better perception of linguistic focus correlated with appropriate use of F_0 and/or amplitude but not with duration (Table 5.5). However, correlation tests do not provide us with a complete picture and some individuals who performed significantly greater than chance in the perception of focus were unable to make appropriate use of acoustic cues in focus production. Conversely, Table 5.1 also shows individuals with poor focus perception who make significant use of more than one acoustic cue in focus production. These results underline the importance of looking at individual performances. For example, two subjects (C16 and C4) surprisingly made significant use of one or two cues (i.e. amplitude only or amplitude with duration) in the production of target focus words. Since they made significant use of one or two cues on appropriate target focus words in production it is possible that these two subjects had developed some abstract awareness of focus possibly through a combination of paralinguistic (e.g. facial expression, body movement, clapping). However, it is also possible that these subjects did not perform well in the perception of focus on the day of testing. In contrast, Table 5.1 also shows that two other subjects (C6 and C7) who were able to hear differences in linguistic focus at a level which was significantly better than chance, did not make appropriate use of any of the acoustic cues in production. Better sensitivity to $F_{0,}$ duration and amplitude difference in Experiment I was not an advantage for the production of these cues. It would appear for these two subjects at least (aged 9;2 and 17;1 at the time of testing), that the ability to hear linguistic focus does not necessarily mean they can consistently make appropriate use of F_0, duration or amplitude in an attempt to convey focus on target focus words.

Overall, only four (C1, C8, C12, C13) of the sixteen subjects managed to convey focus successfully to a trained listener, and the summary of individual scores in Table 5.1 shows that one of these four subjects (C1) managed to make significant use of F_0. These four subjects were among the eleven subjects who made consistent use of amplitude in the production of focus, while three of these subjects (not C1) were among the nine subjects who made consistent use of duration. These results provide some evidence that F_0 is not a necessary cue to focus (see hypothesis (ii)).

The results of Experiments II and III summarized in Table 5.1 show that

(i) appropriate production of one or more of the acoustic cues (i.e. F_0, duration or amplitude) by CI talkers as indicated in the line graphs does not necessarily mean that focus is conveyed to a trained listener.

(ii) although twelve of the sixteen CI subjects could perceive focus and make significant use of at least one acoustic cue in production, only four subjects overall (i.e. C1, C8, C12, C13) managed to convey focus consistently to a trained listener.

(iii) six other subjects across the age range investigated (C10, C11, C14, C15, C16, C17) managed to convey focus less consistently to a trained listener which indicates their prosodic skills were still developing.

(iv) Some CI children can perceive focus but they seem unable to make appropriate use of any acoustic cue in the production of focus in Experiment III (e.g. C6, C7).

(v) Two subjects (C4, C16) who performed poorly in the perception of focus were able to make appropriate use of one or two cues (amplitude with or without duration). However, the consistent and appropriate use of duration and/or amplitude cues in production suggests they may have developed abstract awareness of the concept of focus, and perhaps they did not perform well on the day of testing.

(vi) The relationship between perception and production is not straightforward and CI users may make use of one or a combination of acoustic cues for perception of a prosodic contrast such as focus and use a different set of cues for production.

(vii) Results provide some evidence that F_0 is not a necessary cue to focus (see hypothesis (ii) in sections 1.1.2 and 1.11.4).

6.1.2.3 Higher order developmental implications of the results of Experiments II and III: Do CI children follow the same developmental trajectory as NH children?

Although limited, these results suggest that prosodic concepts such as focus might be acquired if CI children have access to other physical cues (sections 1.3 and 1.11.1) even in the absence of sufficient acoustic information. But it may be the case that the consistent use of one or more acoustic cues on target words and the ability to convey focus successfully to a listener may take longer to stabilize and might not be fully acquired even by age 17;1. A follow up study of the same children or additional long - term CI users as they approach adulthood might give us better insight into the trajectory of acquisition. It has been discussed in section 1.11.4 that in language acquisition it is widely accepted for normal hearing children that perception precedes production. But it is also suggested that prosodic development might differ and by age 4;0 years normal hearing children might be able to produce accent and focus in their own speech before they can interpret them in the speech of others (Stackhouse and Wells, 1997; Cutler and Swinney, 1987). This phenomenon is explained by the physiological reflex associated with tension and excitement arising out of an interesting word and it is reported that children at this age are not yet able to process given vs. new and other contrasts. Although some studies suggest that normal hearing children of 6;10 years should be able to process focus words other studies found that variation, ambiguity and difficulty with intonational meaning can continue up to and beyond age 13;0 years (section 1.11.1).

It is difficult to ascertain whether this occurs for children with implants as here only four out of sixteen CI subjects across the age range (5;9 – 17;1) managed to convey focus consistently. All of those subjects made significant use of amplitude in combination with a different cue i.e. with duration (three subjects) and with F_0 (one subject). Subjects who were making significant use of F_0 (three subjects), duration (eight subjects) and amplitude (nine subjects) according to the acoustic measurements on appropriate target words also performed well in the focus perception test which suggests that these subjects have acquired the concept of focus but are not all yet able to convey it consistently. As discussed earlier these subjects may use one of more of the acoustic cues appropriately but increases may be insufficient to make target focus words stand out to listeners. Although the CI subjects were a lot older than the normal hearing subjects referred to above, their perception skills seemed to be developing ahead of production.

However, is difficult to generalize on the basis of these limited results and a more objective listening experiment should inform us whether any additional subjects managed to convey focus to untrained listeners for comparison with the trained listener's judgement. Although Experiment II results indicate that *perceptual skills* seem more delayed for CI children than the NH subjects i.e. by 13;6 most NH subjects scored 100% whereas by 14;6 years most CI children were significantly above chance levels. There seems to be a gradual improvement in performance across that age range (Figure 3.2) which suggests that despite limitations of the implant CI listeners have developed abstract phonological representations at the perceptual level of prosodic contrasts such as focus or compound vs. phrase stress using whatever cues were accessible to them. There are additional complexities to be taken into account for children with cochlear implants which might be expected to account for individual variation including device limitations, age at implant, duration of deafness and age at time of testing which are discussed in more detail in section 6.1.4 below. Experiment III results varied across the age range and confirmed that unlike perception, the ability to convey focus in *production* does not necessarily improve with age**.**

In a study of a different prosodic feature (i.e. weak syllable processing) Titterington et al. (2006) found that children with cochlear implants had a similar prosodic hierarchy to a group of language age matched hearing children showing a preference for a strong/weak (trochaic) template in their speech production (section 1.3.2.3). The influence of prosodic foot structure had not previously been considered for children with implants and the authors conclude that difficulties associated with perceptual salience cannot fully account for the omission of some weak syllables (e.g. in **ba**nana).

However, it would appear in the current perception findings that children with cochlear implants are more delayed in their ability to perceive prosodic contrasts than hearing children whereas the ability to make significant use in production of acoustic cues and to convey focus seems to be more variable for CI subjects and does not necessarily improve with age. These results are not yet conclusive as there were only a small number of children at each age interval who participated in the experiments and there are very few detailed comparative studies of NH and CI children to draw on especially for prosodic development in different varieties of English. In future experiments a matching group of normal hearing children should be included in the production data for comparison with CI subjects and in general a larger number of hearing and implanted children should be included in any future perception and production experiments.

6.1.2.4 How do the results of the current investigation of English speaking CI children support previous studies of CI children using Cantonese and Mandarin tones?

Barry and Blamey (2004) in a study of Cantonese tones report that their normal hearing subjects (aged 3;0 – 6;0 years) were still acquiring a tonal system and found evidence as here of a mismatch between perception and production. Many of Barry and Blamey's CI subjects produced some appropriate F_0 contours which could be labelled as correct from a visual inspection of acoustic measurements but only a few subjects were judged to be able to produce meaningful tonal differentiation with sufficient frequency for the tonal system to be considered as acquired. In a study of Mandarin tones Peng et al. (2004) found that 6;0 – 12;0 year old CI children who performed well in tone production also performed well in tone identification but not the reverse, and they also found that correlations between tone identification and tone production were not significant when high scoring children were removed. They concluded (section 1.8.3) that tone identification and tone production do not develop in parallel and while perception correlated significantly with duration of implant use, production correlated negatively with age at implant (i.e. better performance by children implanted at a younger age).

Direct comparisons between lexical tones and English intonation and stress patterns are not straightforward for acoustic and methodological reasons. As discussed in section 1.11.3 lexical differences in Cantonese and Mandarin tones are mainly signalled by F_0 with some limited amplitude and duration information in Cantonese and Mandarin respectively, so CI listeners may be more dependent on F_0 for the perception of lexical differences in tone languages rather than an abstract representation of different tones. As discussed in section 1.4.4 falling intonation in declarative sentences occurs in both Southern British English and in Southern Hiberno English but in Belfast English a terminal rise in F_0 is more typical. Given the difficulties CI children have in hearing changes in F_0 generally these dynamic differences in F_0 are unlikely to be perceptible to them. This would suggest that some prosodic contrasts in English expressing emotions and attitudes (e.g. likes vs. dislikes, reservation vs. certainty) might be less accessible to implanted children than others if they are only signalled by rising or falling F_0. It is possible that perception of contrasts other than those investigated in the current study might be more reliant on F_0 cues and

CI listeners may perform better with faster stimulation rates. It may be particularly difficult for CI children to develop abstract phonological representations of prosodic contrasts which are predominantly signalled by F_0 so they will be even more dependent on paralinguistic cues such as facial expression or gesture which might convey an emotion or attitude but not the important changes in F_0.

The clinical and developmental implications of limited access to F_0 for children using cochlear implants are discussed in more detail in section 6.1.5 below. All these issues need to be investigated systematically in future research**.** However, the results of the present experiments suggest that F_0 is not a necessary cue to the contrasts of focus and compound vs. phase stress (hypothesis (ii) in sections 1.1.2 and 1.11.4). The results summarized in Table 5.1 support the findings reported by Peng et al. for Mandarin tones. The four CI subjects in the present study who managed to convey focus to a listener performed well in the perception test but good performance in the perception of focus did not necessarily ensure that the child could convey focus successfully in production. It remains to be seen whether a listening test measuring untrained listeners' ability to identify the intended focus position in the CI children's production would confirm the analysis of acoustic measurements and judgements of the expert listener reported here.

6.1.2.5 Does stimulation rate affect perception performance?

The current study indicated no advantage for faster stimulation rates in the perception of focus and compound vs. phrase tasks (section 3.5.6). There were some individuals using both ACE (600 – 1800 pps) and SPEAK (250 pps) who were performing significantly above chance levels in Focus 2 and Focus 3 tasks (Figure 3.4). These results support Barry et al. who also found there was no significant difference between ACE and SPEAK users. However, studies of Chinese tones (section 1.8) reported better perception performance when one of a pair of tones was a high tone whereas dynamic aspects of pitch such as rising or falling were reported to be less salient (sections 1.11.3 and 1.11.5*)*. Listeners with a higher pulse rate strategy (ACE) tended to respond better to dynamic changes in pitch than users of the lower pulse rate SPEAK strategy, but the difference was not significant. When comparing current results with previous studies it must be taken into account that methodologies and stimuli vary and as discussed above in section 6.1.2.4, there are also differences in the importance of F_0 or other acoustic cues in the perception or production of prosodic contrasts in Chinese tones (sections 1.8 and 1.11.3) and English (sections 1.11.1. and 1.11.2)

6.1.3 Experimental design considerations in the present study

Since Experiments I, II and III in the current study were measuring different skills, the differences in the experimental design are discussed below.

6.1.3.1 The merits of group vs. single case studies in clinical research

In *group research* statistical analysis of the data is useful if a particular variable e.g. cochlear implantation is predicted to affect all the subjects in a particular way. For example in Experiment I in the present study changes in the acoustic parameters F_0, duration and amplitude are controlled and it is expected that the implant will affect perception performance, and that any significant correlations between the independent variable (e.g. cochlear implant) and the dependent variable (i.e. performance in the perception test) will be assumed for the group (Bullis and Anderson, 1986). However, there can be disadvantages in group studies as there are sometimes confounding factors that can affect the validity of the results. In clinical data such as the present study there are several variables such as age at implant, age at time of testing, duration of implant use and stimulation rate that need to be taken into account.

However, the task in *Experiment I* does not make any linguistic demands and the normal hearing and implanted children do not have to draw on stored knowledge or abstract phonological awareness of prosodic concepts, so chronological age, age at implant or duration of implant should not affect performance once it is established that the subjects understand the nature of task. However, variables such as duration of implant and stimulation rate of the implant might vary between implanted subjects and might have some influence on individual performances, so they need to be incorporated into the data analysis. *Experiments II* and *III* on the other hand concern the perception and production of linguistic contrasts, and developmental issues and variables such as age at time of testing for both NH and CI groups, and age at implant and duration of implant use for the CI group might be expected to have an affect on performance in these tests but they can be factored out in statistical analyses. However, since it is difficult to get equal numbers of subjects in a clinical population with comparable ages, duration of implant use, and similar level of linguistic competence it is inevitable that the subjects with cochlear implants will differ on a number of those variables.

The relationship between all of these variables as well as technical limitations of the implant such as the stimulation rates in the processing strategies are complex and affect the results in different ways for individual subjects. Although comparison of group averages can be useful for comparison between performances by CI subjects and a normal hearing control group, there are limitations and results should be interpreted with caution. For example, details regarding about individual performances can be lost in the averaging process, and it is not clear which of the subjects are performing poorly or and which of them are performing significantly greater than chance. Unquestioning acceptance of the statistical significance of the data can obscure individual performances, and statistical methods (e.g. Bonferroni adjustment for multiple comparisons) of correcting inherent differences between groups do not always provide a perfect solution (p. 345). Another disadvantage of group analysis is that little practical clinical application whereas the advantage of focussing on single cases is that relationships between the variables can be inferred using relevant criteria rather than statistics. Replication of single case studies can be carried out to establish the external validity of research findings which can support or refute a particular theoretical position or hypothesis. In the current study individual results are presented in scattergraphs and line graphs to facilitate discussion of individual performances and this is used in addition to statistical analyses of the NH and CI group results.

6.1.3.2 The use of non-meaningful stimuli in Experiment I

Experiment I involved the perception of controlled changes in the acoustic parameters stress (i.e. F_0, duration and amplitude) in pairs of non-meaningful synthetic /baba/ stimuli as described in section 2.2.2. The advantage of the controlled conditions was that perception thresholds for each acoustic parameters could be tested in isolation across the age range without imposing any linguistic demands on any of the subjects. The results informed us of individual subject's sensitivity to differences in F_0, duration and amplitude and gave some indication of how accessible these cues might be to the same listeners in natural speech. These F_0, duration and amplitude thresholds together with the measurements of these cues produced by the four NH talkers in Experiment II stimuli (see Appendices in Chapter Three) gave some indication of how accessible the F_0, duration and amplitude cues were to individual CI subjects in the focus stimuli. These results provided some explanation as to why focus might or might not have been perceived by individual subjects in Experiments II.

6.1.3.3 The use of meaningful linguistic stimuli in Experiments II and III

Experiments II and III differed from Experiment I in that both experiments were concerned with the perception and production of one or more of these acoustic cues in meaningful target words in natural speech. Experiment II required listeners to use whatever acoustic cue(s) were available to them to perceive differences in linguistic focus and compound vs. phrase stress. Production performance in Experiment III was concerned with the appropriate production of F_0, duration and amplitude cues in target focus words and measurements were presented in the line graphs. However, as discussed earlier the appropriate use of one or more acoustic cues on the target focus word was probably in some cases insufficient to convey focus to a listener. Only one of the three tests used in Experiment II (i.e. Focus 3 test) was analysed in detail in the production data in Experiment III. The decision to analyse acoustic measurements for three target focus words in the Focus 3 test was because there were two pre-final target focus words (section 3.2.2) which were not competing with boundary markers or end of a conversational turn in final focus position (Wells et al., 2004). For normal hearing listeners boundary markers such as the above are signaled by final lengthening or terminal fall in F_0 in Southern British English or Hiberno English, or terminal rising F_0 in other varieties of English such as Belfast English. The two pre-final focus words in Focus 3 tests stimuli would not be affected by these boundary cues, whereas in Focus 2 stimuli there was only one pre-final target word. Other differences between the Focus 2 and Focus 3 sentence types are discussed in more detail below in section 6.1.3.4.

Preparation of the production materials for the acoustic analyses required far more manual intervention that had been expected preventing the analysis of additional data that was recorded (section 4.2.2.1). The limited sample of the production data in Experiment III made it difficult to set up robust statistical tests of the hypotheses. In the future detailed analysis of the Focus 2 stimuli would be useful for comparison with the acoustic measurements for Focus 3 stimuli.

6.1.3.4 Differences between NH and CI results

The NH subjects who participated in Experiment I and II were not identical in the current study although some participated in both experiments. Since we were only concerned with how NH performances within each study compared with the CI group this was not a disadvantage.

Perception of controlled F_0, duration and amplitude differences in Experiment I by NH and CI subjects

In *Experiment I* we were concerned with how the NH subjects with a simulation of CI processing performed and results indicate that the ability to hear smaller F_0 differences was poorer in the high F_0 range (Figure 2.4) than in the low F_0 range. The results for the NH children in the simulation condition exceeded expectations given the limited glide identification reported by Green et al. (2002, 2004) for adults in simulation studies (sections 1.10 and 1.11.5). However, results need to be interpreted with caution (Laneau et al., 2004) as vocoders and filters vary in different simulation experiments with NH subjects, and CI subjects have additional complexities such as duration of deafness, age at implant, neural survival, experience with the implant, and stimulations rate which might affect subjects in different ways. The current study indicates that some NH subjects in a simulation condition were hearing smaller F_0 differences in the low F_0 range than the CI subjects and the difference between the two groups was significant (section 2.3.1 and Figure 2.4). In the high F_0 range there was more variability for the CI subjects than the NH group in a simulation condition but the difference between the two groups was not significant. There was no significant difference in the perception of *duration* by the NH subjects in the simulation condition and the CI subjects (section 2.3.2 and Figure 2.6) where both groups could hear differences of 60% or less. *Amplitude* discrimination, however, was significantly better for the NH group in the simulation condition than the CI group (section 2.3.3 and Figure 2.8).

Perception of linguistic focus in Experiment II by NH and CI subjects

The NH listeners' perception of focus (three target words) and compound vs. phrase stress correlated as for the CI group with age at time of testing (see Tables 3.2 and 3.3), but the correlation was approaching significance in Focus 2 for the NH group but not for the CI group. The scattergraph in Figure 3.2 shows that the NH subjects improved consistently in all three subtests across the age range whereas there was more variability in individual scores for the CI subjects in the Phrase test. Overall they were more delayed and unlike the NH subjects scores never reached ceiling level (see more discussion in section 3.5.3.1). As discussed in section 3.3.1 all of the NH subjects (total = 22) scored significantly higher than chance in the Focus 3 test but there were some individual subjects who were below chance in the Focus 2 test (five subjects) and the Phrase test (five subjects). Performance was more variable for the CI subjects with six subjects in the Phrase and Focus 2 tests and twelve subjects in the Focus 3 test who scored significantly greater than chance.

As discussed in section 3.5 there were additional differences between the focus sub-tests other than the number of target focus items which might have accounted for variation in performance by the CI subjects. A higher chance level of Focus 2 (50%) in the two choice test made it even more challenging for the CI subjects to have a score which was significantly better than chance than in the three choice test in Focus 3 (33.3%). There were also differences in syntactic and prosodic structure (i.e. adjective + noun vs. subject + verb + object) in Focus 2 and Focus 3 respectively with more stressed and unstressed syllables in the latter e.g. *a **BLUE** book* vs. *the **DOG** is eating a bone.* However, the differences in the decline and terminal fall or boosting of F_0 on target focus word in these two sentence types would have only been accessible to the NH subjects and not to the CI subjects who had to rely on amplitude and duration cues (section 3.5.4). Despite the limited access to F_0, good performances by individual CI subjects in all three subtests support hypothesis (ii) which suggests that F_0 may not be a necessary cue to the perception of linguistic focus or compound vs. phrase stress.

Acoustic measurements of the production data in Experiment III for CI subjects and NH talkers

Acoustic measurements of the F_0, duration and amplitude measurements were also carried out for the four NH talkers (aged between 27;0 and 12;0) who produced the Focus 3 stimuli in Experiment II and this formed a reference set for discussion in the analysis of CI subjects' productions in Experiment III. Although it was useful to have these four talkers' productions, a group of age matched NH children would have facilitated more direct comparison with the CI group so this will be included in future production experiments. Due to time constraints and subjects' availability for testing in the current study production data was not included for the NH children who participated in the perception experiments. Performance by CI subjects in the production of focus in Experiment III was judged on ability to make appropriate use of F_0, duration and amplitude as presented in the line graphs in Chapter Four. The results show that changes or increases in these cues might often have been in the appropriate direction but in some cases were insufficient when focus was not conveyed to a trained listener (the present investigator). Similar analyses for an NH group in future experiments would be useful for direct comparison with the CI group.

Regional variations in English

Although there are similarities between Southern British English (SBE) and Southern Hiberno English (SHE) in that both have a falling intonation pattern in neutral declarative sentences (section 1.4.4) it has been reported in studies of adults that individuals may vary in the use of acoustic cues used to signal compound vs. phrase stress and narrow focus such as silence, lengthening loudness and pitch reset or changes in pitch configuration especially in spontaneous speech (Peppé et al., 2000; Xu and Xu, 2005). Due to time constraints in the current study there were no matching NH children in the production experiment and this would have been useful for comparative purposes in the absence of normative data for speakers of Southern Irish English. The predominance of rising intonation in Northern Hiberno (Belfast) English and the use of pause (sections 1.4.4 and 1.2.1) rather than pitch in signaling boundaries in this and other regional variations such as Edinburgh Scottish English (ESE) and the implications for children with cochlear implants with limited access to F_0 also need to be investigated in the future. It has yet to be established for other dialects or varieties of English, such as ESE or Belfast English, whether F_0 is a necessary cue to intonation contrasts such as focus (hypothesis (i)), or whether F_0 is not a necessary cue (hypothesis (ii)) as indicated in the results of the present study of Southern Hiberno English.

Objective listening test in the future for CI and NH production data in Experiment III
In the future a listening test involving all the available production data for CI subjects in Experiment II will be delivered to a group of untrained listeners for comparison with the investigator's impression. The results of the listening test will be used to analyse the relationship between the perception of focus and the ability to convey focus to listeners who are unfamiliar with the Experiment III data. Additional data for CI subjects' production of the other prosodic contrasts (compound vs. phrase stress and focus in two element phrases) will also be analysed and included in future listening tests with data from age matched NH subjects.

6.1.4 Variables affecting CI individual performances in Experiment I, II and III

6.1.4.1 Do factors such as age at implant/switch-on, duration of implant use, age of testing, or stimulation rate account for variability in performance?

As discussed in 1.11.5 the effects of variables such as duration of deafness, age at time of testing, stimulation rate (section 6.1.2.5) are well documented in general outcome studies of speech perception and production skills for English-speaking CI children (Nikolopoulos et al., 1999; Tait and Lutman, 1997;Walzman and Cohen, 2000; Blamey et al, 2001). It is also reported in experimental studies of adult implant users that F_0 discrimination varied according to subject, speech processing strategy and F_0 range (see section 1.9). Overall, in the current investigation there were enough subjects to carry out some statistical analyses for the NH and CI groups. There was also discussion of individual scores presented in scattergraphs for all three experiments which is essential for clinical populations where performances can vary for individual subjects due to different influencing factors.

In the present study variables such as age at implant/switch-on, age at time of testing, duration of implant use, and stimulation rate of the speech processor were considered. As mentioned above the CI subjects were drawn from the cohort of children who were available at the time of testing so there were variations in these factors for individual subjects across the age range (5;7 – 16;11 years). Results show that were no correlations between the appropriate *production* of F_0, duration and amplitude in Experiment III and variables such as age at production, duration of implant use, stimulation rate, or age at switch-on. Previous studies such as Barry and Blamey (2004) report that a Cantonese tonal system was still developing in normal hearing children and children with implants in their study whereas Peng et al. (2004) found that Mandarin tone production was better for those implanted at an earlier age. Xu et al. (2004) concluded that age and other variables should be considered in the future.

As discussed in section 3.5.3.1 *perception* scores in Experiment II improved with age for the NH and CI children and correlations were found between age at time of testing and perception of compound vs. phrase stress and focus (i.e. Phrase and Focus 3 tests). However, Pearson Correlations tests show that high and low F_0 range thresholds correlated significantly with Focus 2 and Focus 3 scores when age was controlled which suggests that performance in these tests was linked with ability to hear differences in F_0 (section 3.5.4.1). When age was controlled a correlation between duration thresholds and Focus 2 remained but a correlation with Focus 3 disappeared. These results indicate an age effect for Focus 3 (section 3.5.4.2) whereas performance in Focus 2 seemed to be linked with ability to hear differences in duration.

The results support previous results by Ciocca et al. (2002) who found that the correlation between tone perception and age at testing and age at implantation was not significant. Barry et al. (2002b) also concluded that the effects of linguistic development and the gradual development of tone needed to be established for NH and CI children. In the future a longitudinal study of English speaking CI children might be useful to monitor the development of prosodic perception and production skills up to adulthood. A similar study of normal hearing children in the same linguistic environment (i.e. Southern Hiberno English) in the same range would be useful for comparison. Although the current results show a gradual acquisition of prosodic competence which supports previous studies (Atkinson-King, 1973; Vogel and Raimy, 2002; Wells et al., 2004) there was a difference in performance between the NH and CI groups. By 13;6 years all the NH children were at or close to 100% whereas most CI children were significantly greater than chance by 14;6 years. However, there was no evidence of a correlation between perception of linguistic contrasts (i.e. compound vs. phrase stress and focus) in Experiment II and duration of implant use whereas reports vary (Ciocca et al. 2002; Peng et al. 2004) in studies of Chinese tones (section 1.8).

6.1.4.2. Additional factors that might contribute to variability: pre-operative hearing loss, pre-operative perceptual skills, number of electrodes, aetiology.

There are other factors presented in Table 2.1 and Table 2.2 which were not considered formally in the analyses of the current data. These might account for the diversity in performance and could be addressed explicitly or controlled in the design of future experiments (Waltzman, 2000). The CI children in the current study were drawn from a cohort of implanted children who could complete the tasks at the time of testing so individual variation in baseline *pre-operative hearing loss* was inevitable. Pre-operative hearing losses for the CI subjects varied considerably and as reported by Dowell, Blamey, and Clark (1995) this is one of five variables along with duration of profound hearing loss, progressive hearing loss, oral/aural education and duration of implant use which account for 37% of the variance in post-operative speech perception results. General speech perception skills at the time of testing were not formally addressed for the CI children in the current investigation. A variety of standard speech discrimination tests were used which reflected a range of general speech perceptual ability across individual subjects of different ages, and in the future *pre-operative language ability* should also be considered. Better *pre-operative speech perception skills* might contribute to better speech discrimination post-operatively and in future investigations it might be worth grouping children with similar pre-operative perceptual skills. Table 2.2 also shows that *onset of deafness* for eleven CI subjects in the current investigation was congenital, but for five subjects onset of deafness was between two weeks and three years and for one subject onset of deafness was unknown. However, the effects of age at implant and duration of implant use were incorporated into the current analysis because the length of auditory deprivation affects plasticity and ultimately performance with an implant (Sharma, Dorman and Spahr, 2002; Sharma and Dorman, 2006).

Table 2.1 also shows that all except one subject attended mainstream school, and although the *aetiology* of deafness was unknown for the majority (ten subjects) there was some variation for the rest of the subjects i.e. meningitis (five subjects), CMV (one subject) and Waardenburg (one subject)) which might have contributed to the variation in results. However, Table 5.1 indicates that although perception performance for the Focus 3 test was significantly greater than chance (33.3%) for most subjects, there was wide variation in scores and even within the group of children who were deaf as a result of meningitis (CI, C2, C4, C6, C16), and age at onset of deafness for these subjects, which ranged between two weeks and three years, might also have accounted for variation of scores. *Ossification of the cochlea* can occur following meningitis and sometimes only a *partial insertion* of the electrode array is possible. Only one of the subjects who were deaf as a result of meningitis (C4) had a partial insertion (i.e. 14 electrodes) and the rest had a full electrode array inserted.

Individual thresholds can increase or decrease over time and might affect performance but this can be managed by regular tuning of the speech processor. In the future, advancements in *implant design* and *speech processing* might change the relationship between different known and unknown variables and help improve individual perception and production performances of CI subjects. Studies of adult implant users (section 1.11.5) report some improvement with modified speech processing strategies but it remains to be seen whether this makes a difference for children with implants. There may be other factors beyond the scope of the present study such as differences in the *placement of the electrode array* in the cochlea or individual variation in *neural survival* which may account for differences in perceptual skills and are also worth considering in the future. The interaction between all the variables is not yet known but the wide variation in performance among implanted children does not seem to be solely due to the implant (Waltzman, 2000).

6.1.5 Clinical implications: practical relevance of the results

6.1.5.1 Acquisition issues: how can young implanted children acquire stress and intonation skills at home or in clinical and educational settings in the absence of F_0 *(pitch) information?*

The results of Experiment I and II suggest that F_0 is not a necessary cue to lexical stress and focus (hypothesis (ii) in sections 1.1.2 and 1.11.4) and that in normal conversational speech most CI subjects would have difficulty hearing most of the changes or increases in F_0 in prosodic contrasts such as focus which are less than half an octave (Chapter Three). This suggests that CI listeners have to rely on other cues such as exaggerated lengthening and loudness in addition to paralinguistic cues, such as facial expression, body movement and rhythmic clapping. The few subjects who could hear smaller differences in the high F_0 range in the present study might be able to hear changes or increases in the speech of women or other children but the results suggest that F_0 changes in natural speech would in general be inaccessible to most implanted children.

These results have important implications for professionals working in different educational settings. For example, playschool and junior class teachers should be made aware of some of these limitations so that stress and intonation contrasts (e.g. compound vs. phrase stress and focus and other contrasts), which are important aspects of language development, can be made more accessible to an implanted child in group activities such as circle time or story time. In this way an implanted child might pay more attention and also gain better access to emotions and feelings expressed by teachers through stress and intonation such as anticipation, surprise, anger, emphasis, disappointment, amusement, excitement while telling stories using large picture books. Young implanted children with delayed language and vocabulary should then be better able to participate and derive some benefit and enjoyment as well as some understanding of what is going in a story which will promote language acquisition. The results underline the importance of clinicians exaggerating cues with young children such as facial expression, rhythmic cues such as clapping or tapping, increased lengthening and loudness without distorting natural rhythmic patterns to highlight key vocabulary and phrases in clinical sessions and make them as accessible as possible to children using implants (section 1.11.1). However, some clinicians have taken the view that auditory training should be carried out by covering the mouth or by sitting alongside the child. This approach may be useful for some testing purposes but for normal interaction and promotion of prosodic development in young implanted children a more natural form of face to face communication allows the child to use any available prosodic cues.

It is important that all of these issues are explained and incorporated into pre- and post implant support offered to teachers and speech therapists by clinicians in cochlear implant teams. Parents can be informed in an accessible way about the limitations of the implant, and modeling by clinicians, which is standard practice, is especially useful for parents who might be a less comfortable using exaggerated intonation or dramatizing body movement and facial expression while telling stories or interacting with their implanted child. These issues have implications for the perception and production of attitudinal and emotional information by CI children during the development of social and interpersonal skills which will ultimately enhance their general language development.

6.1.5.2 How do CI and normal hearing children differ in prosodic development?

As discussed in sections 3.5.3.1 there seemed to be a gradual improvement in the perception of prosodic contrasts (i.e. compound vs. phrase stress and focus) for the CI subjects whereas performance improved more rapidly up to 10;0 years for the NH children and was close to 100% for many subjects thereafter (see Figure 3.2). As presented in Figure 3.2 test scores were at or close to 100% by 13;6 years for the NH subjects whereas the CI subjects scores were significantly greater than chance by 14;6 years. The results of the current study are preliminary and useful information for therapists and teachers but further investigation is needed with more CI and NH children at regular age intervals using different varieties of English. An awareness of individual differences in how prosodic competence develops in CI children should be borne in mind when testing and planning educational and speech programmes. Both cognitive and linguistic factors should be also taken into consideration (Ciocca, 2002 and Barry et al., 2002b) and psychological tests and baseline language assessments might also help account for some variation in performances.

6.1.5.3 Use of visual displays by clinicians to investigate ambiguity or insufficient boosting of one or more acoustic cues in the production of prosodic contrasts such as focus

Experiment III results show some implanted children produced *broad* rather than *narrow* focus by insufficient boosting of one or more acoustic cues on the target focus words (section 4.4.4). These results have useful implications for the assessment of prosodic competence such as the ability to convey focus on a target word. If, for example, focus is not perceptible to a clinician or if a response is ambiguous it might be useful to look at a sentence with a target focus word in a visual display to establish whether there are appropriate but insufficient increases in one or more acoustic cues (F_0, duration or amplitude) for diagnostic purposes. As the results of the present study indicate there may be appropriate adjustment of one or more cues on a target word which might not be sufficient to convey focus, a visual display might help establish whether these talkers are at least attempting to use at least one cue appropriately or trying to convey focus on the target word, and might indicate whether they are developing prosodic competence. In addition, if an implanted child is not producing appropriate F_0 contours yet is managing to convey a prosodic contrast such as focus to a listener, a visual display will tell the clinician if he/she may be making use better of increased lengthening or loudness on target words.

However, visual displays should be used with caution for training or correction purposes (King and Parker, 1980; O'Halpin, 2001) because individual children with implants seem to use different cues to convey or perceive prosodic contrasts as indicated in the results of the present investigation. For example, there might not be a direct correspondence between perception and production of F_0 and just because an implanted child cannot hear differences in F_0 in the linguistic contrast does not mean he/she cannot produce appropriate changes in F_0. It was discussed earlier in sections 1.3.2.4 and 1.11.2 that excitement and tension generated by interest in a focus word by normal hearing children even before they have acquired this contrast may raise $F_{0,}$ and increased amplitude is often associated with a rise in F_0. On the other hand the current study shows that some implanted children may be able to hear smaller changes in F_0 without being able to produce them appropriately. The results of the current investigation show individual CI subjects can vary in the combination of cues they use to convey prosodic contrasts which according to the literature is not altogether unusual in normal hearing adults and children and hearing aid users. Clinicians should be aware of this for planning of appropriate intervention and training programmes as well as for testing and assessment.

6.1.6 Concluding comments

6.1.6.1 Perception issues: main considerations

The results of the current study seem to support the view set out in hypothesis (ii) that F_0 is not a necessary cue to stress and intonation contrasts such as compound vs. phrase stress and focus. It was discussed in sections 1.1 and in 1.11.4 that duration and amplitude adjustments in adult speech such as extra lengthening or changes in loudness help to facilitate prosodic development for normal hearing children in addition to changes in F_0. But it would appear from results of the current study that because of the limited F_0 information available through the implant, duration might be a more reliable cue than F_0 and amplitude for CI children. However, variation in the results of the perception and production experiments suggest that some individual subjects may be able to hear smaller F_0 and amplitude changes than others and that children may perceive an intonation contrast such as focus using one combination of cues and try to produce it with a different set. As discussed earlier, there may be other intonation contrasts besides focus or compound vs. phrase stress where more dynamic aspects of F_0 such as a rising or falling intonation may play a more important role in contrasts such as likes vs. dislikes, reservation vs. certainty. Similar analysis needs to be carried out in future research for these contrasts. Where acoustic cues are inaccessible to implant children they might be able to draw on paralinguistic cues such as eye contact, gestures, jumping up and down and reaching to develop an abstract representation of some prosodic contrasts which is independent of their ability to hear a particular cue.

6.1.6.2 Production issues: main considerations

The ability of 3 to 4 year old normal hearing children to convey focus in their own speech before they can process pragmatic information in the speech of others (section 1.3.2.4) is explained by a *universal physiological mechanism* associated with tension and semantic interest in a word. We might expect a similar phenomenon in children with implants but the current results suggest that only three out of sixteen implanted children (aged 5;9 – 17;1 years) made significant use of F_0 as indicated by the acoustic measurements in Experiment III. Only four of the sixteen children managed to convey focus consistently to a trained listener and only one of these subjects made significant use of F_0.

However, there were other implanted children who were approaching significance in the appropriate use of F_0, and there were also some subjects who conveyed focus to a listener with a consistency that came close to the level adopted here as significant. These results suggest that individual children may be at different stages of the acquisition process regardless of their ability to use F_0 appropriately or convey focus to a listener. The six subjects spanning across the entire age range (Table 5.1*)* who were only able to convey focus at or above chance level, did not make significant use of F_0 or other acoustic cues apart from amplitude by two subjects. They often sounded *ambiguous* as a result of insufficient boosting of target focus words. However, ambiguity is also reported for hearing aid users (Allen and Andorfer, 2000), and for normal hearing children it has been suggested that the acquisition process may continue into adulthood (Wells et al., 2004). It remains to be seen whether those implanted subjects who were at or below chance will develop prosodic competence so that they can consistently convey focus to a listener in the future. Since the current study concerned only a small number of subjects further investigation and longitudinal studies of age matched or language matched normal hearing and implanted children in a Southern Irish population as well as other dialects and regional varieties of English might give us better insight into differences and similarities in prosodic development.

6.1.6.3 Summary of findings arising from the current study

a. Experiment I thresholds indicate that F_0 differences less than 0.5 octaves are not accessible to most CI listeners and that duration seems to be a more reliable cue than amplitude.

b. Experiment II results indicate that most subjects can hear differences in linguistic focus and compound vs. phrase stress even though they though they are unable to hear F_0 differences less than 0.5 octaves.

c. These results seem to suggest that F_0 is not a necessary cue to stress and intonation in focus stimuli (hypothesis (ii) in section 1.1.2).

d. CI users may perceive linguistic focus with one or more acoustic cues and make appropriate use of a different set of cues in production, and a similar pattern has been reported for hearing aid users.

e. Although most of the CI subjects were significantly better than chance in the perception of linguistic focus and most used one or more acoustic cues appropriately in the production of the target focus word, only four out of 16 CI subjects overall managed to convey focus to a trained listener. Many were ambiguous which is not unusual in normal hearing adults and children and hearing aid users.

f. Perception of linguistic focus seems to develop ahead of production skills.

g. Variation in performance across the CI subjects has implications for professionals dealing with children in educational and clinical settings. In the absence of F_0 information they can rely on other acoustic and paralinguistic cues such as facial expression, gesture, amplitude and duration to hear intonation contrasts such as focus in everyday speech.

h. Ability to perceive differences in linguistic focus does not necessarily mean CI subjects can produce them effectively. Those who were less consistent but approaching significance in the appropriate use of F_0, duration and amplitude have not yet stabilized and might still be in the process of acquiring prosodic competence.

6.1.6.4 Future research

a. A listening test will be conducted with a group of untrained listeners who will be required to judge whether focus has been conveyed on different target words in the production data in Experiment III.

b. Results are based on performances of the 17 CI subjects who were available to participate in the experiments at the time of testing. Additional data from more CI subjects will indicate whether the current results can be supported.

c. Since there is no available normative data on prosodic development for a Southern Irish population of normal hearing children a set of age or language matched normal hearing controls should be included in future perception and production experiments for direct comparison.

d. the current investigation only concerns two linguistic contrasts (focus and compound vs. phrase stress), and we need to examine other prosodic contrasts such as attitudes and emotions to establish whether F_0 is a necessary cue for the expression of likes vs. dislikes, certainty vs. reservation (hypothesis (i)). In future experiments more CI subjects could be grouped according to age (i.e. under three years, under five years, over five years), onset of hearing loss (i.e. children with progressive hearing loss, acquired hearing loss) and aetiology.

e. Variables not controlled for in the current study should be considered in the future such as pre-operative hearing, pre-operative perceptual ability, different stimulation rates, pre-operative language and speech skills. There may also be other factors such as neural survival, placement of the electrodes in the cochlea and as yet unknown factors which might account for individual variation in performance.

REFERENCES

Abberton, E., Fourcin A. and Hazan, V. (1991). Fundamental frequency range and the development of intonation in a group of profoundly deaf children. *Proceedings of the XIIth International Congress of Phonetic Sciences, Aix-en-provence*, Volume 5.

Abberton, E. (1972). Visual feedback and intonation learning. In A. Ringauet and A. Charboneau (Eds.) *Proceedings of the 7th International Congress of Phonetic Sciences*.

Atkinson-King, K. (1973). Childrens's acquisition of phonological stress contrasts. *Working Papers in Phonetics*, 25, UCLA.

Allen, G.D. and Andorfer, P.M. (2000). Production of sentence-final intonation contours by hearing-impaired children. *Journal of Speech, Language and Hearing Research*, 43, 441-455.

Ashby, M. (1992). Metalinguistic awareness of nuclear accent and of compound stress in an 8 year old. European Journal of Disorders of Communication, 27, 2, 175-179.

Barry, J.G. and Blamey, P.J. (2004). Acoustic analysis of tone differentiation as a means of assessing tone production in speakers of Cantonese. *Journal of the Acoustical Society of America,* 116, 3, 1739-1748.

Barry, J.G., Blamey, P.J., Martin, L.F.A., Lees, K.Y.-S, Tang, T., Ming, Y.Y. and van Hasselt, C.A. (2002a).Tone discrimination in Cantonese-speaking children using a cochlear implant. *Clinical Linguistics and Phonetics*, 16, 2, 79-99.

Barry, J.G., Blamey, P.J. and Martin, L.F.A. (2002b). A multidimensional scaling analysis of tone discrimination ability in Cantonese-speaking children using a cochlear implant. *Clinical Linguistics and Phonetics*, 16, 2, 101-113.

Beckman, M.E. and Pierrehumbert, J.B. (1986). The Intonational structure in English and Japanese. In C. Ewen and J. Anderson (Eds.) *Phonology Yearbook 3,* Cambridge: CUP.

Bishop, D. (1989). *Test for Reception of Grammar* (second edition). Manchester: Department of Psychology, University of Manchester.

Blamey, P.J. Sarant, J.Z., Praatch, L.E., Barry, J.G., Bow, C.P., Wales, R.J., Wright, M., Psarros, C., Rattigan, K. and Tooher, R. (2001). Relationships among speech perception, production, language, hearing loss, and age in children with impaired hearing. *Journal of Speech, Language, and Hearing Research*, 44, 264-285.

Boersma, P. and Weenink, D. (2005). *PRAAT: Doing phonetics by computer*, version 4.3.19, www.praat.org.

Bolinger, D.L. (1958). A theory of pitch accent. *Word*, 14, 109 - 149

Bolinger, D.L. (1983). Intonation and gesture. *American Speech*. 58, 2, 156-174.

Borden, G.J., Harris, K.S. and Raphael, L.J. (1994). *Speech Science Primer: Physiology, Acoustics, and Perception of Speech*. Baltimore, Maryland: Williams and Wilkins.

Boothroyd, A. (1973). Some experiments on the control of voice in the profoundly deaf using a pitch extractor and storage oscilloscope display. *ILEE Transactions on Audio and Electroacoustics*. Vol AU-21, 3, June

Brazil, D., Coulthard, M. and Johns, C. (1980). *Discourse Intonation and Language Teaching*. London: Longman.

Brown, G., Currie, K. and Kenworthy, J. (1980). *Questions of Intonation*. The Hague: Mouton.

Bullis, M. and Anderson, G. (1986). Single-subject research methodology: an under utilized tool in the field of deafness. *American Annals of the Deaf,* 131, 344 -348.

Chafe, W. (1974). Language and conciousness. *Language*, 50, 1, 111-133.

Ciocca, V., Francis, A.L., Aisha, R. and Wong, L (2002). The perception of Cantonese tones by early-deafened cochlear implantees, *Journal of the Acoustical Society of America,* 111, 5, Pt.1, 2250-2256.

Cleary, M. Pisoni, D.B. and Kirk, K.I. (2005). Influence of voice similarity on talker discrimination in children with normal hearing and children with cochlear implants. *Journal of Speech, Language, and Hearing Research*, 48, 204-223.

Clement, C.J., den Os and E.A. and Koopmans-van Beinum, F.J. (1996). The development of vocalizations of deaf and normally hearing infants. In T.W. Powell (Ed.) *Pathologies of Speech and Language: Contributions of Clinical Phonetics and Linguistics*. A publication of the International Clinical Phonetics and Linguistics Association.

Cooper, W. E. and Sorensen, J.M. (1981). *Fundamental Frequency in Sentence Production*. New York: Springer-Verlag.

Couper-Kuhlen, E. (1986). *An Introduction to English Prosody*. London: Edward Arnold.

Cruttenden, A. (1994). Phonetic and prosodic aspects of baby talk. In C. Gallaway and B. Richards (Eds), *Input and Interaction in Language Acquisition*. Cambridge: CUP.

Cruttenden, A. (1997). *Intonation*. Cambridge: CUP.

Crystal, D. (1969). *Prosodic Systems and Intonation in English*. Cambridge: CUP.

Crystal, D. (1986). Prosodic development. In Fletcher and Garman (Eds.) *Language Acquisition: Studies in First Language development*, 2nd edition. Cambridge: CUP

Crystal. D. (1987). *Clinical Linguistics*, 2nd edition, Cambridge: CUP

Cutler, A. and Carter, D.M. (1987). The predominance of strong initial syllables in the English vocabulary. *Computer Speech and Langu*age, 2, 133 - 142

Cutler A. and Swinney A. (1987). Prosody and the development of comprehension. *Journal of Child Language*, 14, 148-167.

Cutler, A. and Norris, D. (1988). The role of strong syllables in segmentation for lexical access. *Journal of Experimental Psychology: Human perception and performance,* 14, 1, 113-121.

Cutler, A. and Ladd, D. R. (Eds.), (1983). *Prosody: Models and Measurements*. Berlin: Springer-Verlag.

Dalton, M. and Ní Chasaide, A. (2003). Modelling intonation in three Irish dialects. *Proceedings of the 15th International Congress of Phonetic Sciences*. Barcelona.

Dalton, M. and Ní Chasaide, A. (2007). Nuclear accents in four Irish (Gaelic) dialects. Proceedings of the International Congress of Phonetic Sciences. Saarbrücken.

Dankovičová, J., Pigott, K., Wells, B. and Peppé, S. (2004). Temporal markers of prosodic boundaries in children's speech production. *Journal of the International Phonetic Association*, 34 ,1, 17-36.

Dawson, P.W., Nott, P.E., Clark, G.M. and Cowan, R.S.C. (1998). A modification of play audiometry to assess speech discrimination ability in severe-profoundly deaf 2-4 year old children. *Ear and Hearing,* 19, 5, 371-384.

Denes, P.B. and Pinson, E.N. (1993). *The Speech Chain: The Physics and Biology of Spoken Language.* New York: Freeman and Company.

Doherty, C.P., Fitzsimons, M., Assenbauer and Staunton, H. (1999). Discriminaton of prosody and music by normal children. *European Journal of Neurology*, 6, 1-6.

Dowell, R.C., Blamey, P.J. and Clark, G.M. 1995. Potential and limitations in cochlear implants in children. Proceedings of the International Cochlear Implant, Speech and Hearing Symposium, 24-28 October 1994 (Melbourne: Annals Publishing Company) *Annals of Otology, Rhinology and Laryngology*, 104 (Suppl. 166), 324-327.

Faulkner, A., Rosen, S. and Stanton, D. (2003). Simulations of tonotopically-mapped speech processors for cochlear implant electrodes varying in insertion depth. *Journal of the Acoustical Society of America*, 113, 1073-1080.

Faure, G., Hirst, D.J. and Chafcouloff, M. (1980). Rhythm in English: Isochronism, pitch, and perceived stress. In L.R. Waugh and C.H. Schooneveld (Eds.*) The Melody of Language*. Baltimore, Maryland: University Park Press.

Fujimura, O. and Erikson, D. (1997). Acoustic phonetics. In W.J. Hardcastle and J. Laver (Eds.) *The Handbook of Phonetic Sciences*. Oxford: Blackwell Publishers.

Fry, D.B. (1955). Duration and intensity as physical correlates of linguistic stress. *Journal of the Acoustical Society of America*, 27, 765-768

Fry, D.B. (1958). Experiments in the perception of stress. *Language and Speech*, 1, 126-152

Fry, D.B. (1979). *The Physics of Speech*. Cambridge: CUP

Gay, T. (1978a). Effect of speaking rate on vowel formant movements. *Journal of the Acoustical Society of America*, 63, 223-230

Gay, T. (1978b). Physiological and acoustic correlates of perceived stress. *Language and Speech*, 21 4, 347-353

Geurts, L. and Wouters, J. (2001). Coding of the fundamental frequency in continuous interleaved sampling processors for cochlear implants. *Journal of the Acoustical Society of America*, 109, 2, 713-726.

Grabe, E., Post, B., Nolan, F., and Farrar, K. (2000). Pitch accent realization in four varieties of British English. *Journal of Phonetics*, 28, 161 - 185

Grabe, E. and Post. B. (2002). Intonational variation in English. In B. Bel and I. Marlin (Eds.). *Proceedings of the Speech Prosody 2002 Conference*, Aix-en-Provence, 242-346

Green, T., Faulkner, A. and Rosen, S. (2002). Spectral and temporal cues to pitch in noise-excited vocoder simulations of continuous-interleaved-sampling cochlear implants. *Journal of the Acoustical Society of America*, 112, 5, Pt. 1, 2155 -2164

Green, T., Faulkner, A. and Rosen, S. (2004). Enhancing temporal cues to voice pitch in continuous interleaved sampling cochlear implants. *Journal of the Acoustical Society of America*, 116, 4, Pt.1, 2290-2310.

Gussenhoven, C. (2006). Types of focus in English. In: C. Lee, M. Gordon, and D. Buring (Eds.) *Topic and focus: Crosslinguistic Perspectives on Meaning and Intonation*. Heidelberg, New York, London: Springer.

Huckvale, M. (2004). Speech filing system. Version 4/5 www.phon.ucl.ac.uk/resource/sfs/

Isenberg, D. and Gay, T. (1978). Acoustic correlates of perceived stress in an isolated synthetic disyllable. *Journal of the Acoustical Society of America*, 64, S21 (Abstract).

Juscyzk, P.W, Cutler, A. and Redanz, N.I. (1993). Infants' preference for the predominant stress patterns of English vowels. *Child Development*, 64, 675-687.

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

Juscyzk, P.W., Houston, D.M. and Newsome, M. (1999). The beginnings of word segmentation in English-learning infants. *Cognitive Psychology*, 39, 139-207.

Jusczyk, P.W. (2002). Some critical developments in acquiring native language sound organization during the first year. *Ann Otol Rhinol Laryngol.*, 111, 11-15.

Johnson, K. (1997). *Acoustic and Auditory Phonetics*. Cambridge, MA: Blackwell Publishers.

Katz, W.F., Beach, C.M., Jenouri, K. and Verma, S. (1996). Duration and fundamental frequency correlates of phrase boundaries in productions by children and adults. *Journal of the Acoustical Society of America, 99,5, 3179-3191.*

King, A. and Parker, A. (1980). The relevance of prosodic features to speech work with hearing-impaired children. In F. M. Jones (Ed.), *Language Disability in Children: Assessment and Remediation. Lancaster*: MTP Press.

Klatt, D. H. and Klatt, L. C. (1990). Analysis, synthesis, and perception of voice quality variations among female and male talkers. *Journal of the Acoustical Society of America*, 87, 820-857.

Kochanski, G., Grabe, E., Coleman, J, and Rosner, B. (2005). Loudness predicts prominence: fundamental frequency lends little. *Journal of the Acoustical Society of America*, 118, 2, 1038-1054

Kuhl, P.K. and Meltzoff, A.N. (1982). The bimodal perception of speech in infancy. *Science*, 218, 1138-1141.

La Bruna Murphy, A., McGarr, N.S. and Bell-Berti, F. (1990). Acoustic analysis of stress contrasts produced by hearing-impaired children. *The Volta Review*, 92, 80-91

Ladd, D.R. (1980). *The Structure of Intonational Meaning*. Bloomington, Indiana: Indiana University Press.

Ladd, D.R. (1993). On the theoretical status of 'the baseline' in modeling intonation. *Language and Speech*, 36, 435-451

Ladd, D.R. (1996). *Intonational Phonology*. Cambridge: CUP.

Ladd, D.R. and Shepman, A. (2003). "Sagging transitions" between high pitch accents in English: experimental evidence. *Journal of Phonetics*, 31, 81-112.

Ladefoged, P. (2001). *A Course in Phonetics*. Fort Worth, TX: Harcourt College Publishers.

Laneau, J., Wouters, J. and Moonen, M. (2004). Relative contributions of temporal and place pitch cues to fundamental frequency discrimination in cochlear implants. *Journal of the Acoustical Society of America*, 116, (6), 3606-3619.

Laneau, J., Moonen, M. and Wouters, J. (2006). Factors affecting the use of noise-band vocoders as acoustic models for pitch perception in cochlear implants. *Journal of the Acoustical Society of America*, 119, 1, 491-506.

Lehiste, I. (1970). *Suprasegmentals*. Cambridge MA: The MIT Press

Lehiste, I, (1996). Suprasegmental features of speech. In N. Lass (Ed.) *Principles of Experimental Phonetics. Paul Mosby-Year Bk, Inc.*

Levitt, H. (1971). Transformed up-down methods in psychoacoustics. *Journal of the Acoustical Society of America*, 49, 467-477.

Lieberman P. and Blumstein, S.E. (1988). *Speech Physiology, Speech perception, and Acoustic Phonetics*.Cambridge: CUP.

Lieberman, P. (1986). The acquisition of intonation by infants: physiology and neural control. In E. C. Johns-Lewis (Ed.) *Intonation in Discourse*. London: Croom Helm Ltd.

Lowry, O. (2002). The stylistic variation of nuclear patterns in Belfast English. *Journal of the International Phonetic Association*, 32, 1, 33-42

McGarr, N. S., Head, J., Friedman, M., Behrman, A. M. and Youdelman, K. (1986). The use of visual and tactile sensory aids in speech production training: a preliminary report. *Journal of Rehabilitation, Research and Developmen*t, 23, 101-109.

McGarr, N. S., Youdelman, K. and Head, J. (1989). Remediation of phonation problems in hearing impaired children: Speech training and sensory aids. *The Volta Review*, 91, 717.

McNeilage, P.F. (1997). Acquisition of speech. In Hardcastle, W.J and Laver, J. (Eds.) *The Handbook of Phonetic Sciences (*pp. 301-332).

Mahsie, J. (1995). The use of sensory aids for teaching speech to children who are deaf. In G. Plant and K.E. Spens (Eds.) *Profound Deafness and Speech Communication*. London: Whurr Publishers

McKay, C. M., McDermott, H.J. and Clark, G.M. (1994). Pitch percepts associated with amplitude-modulated current pulse trains by cochlear implantees. *Journal of the Acoustical Society of America*, 96, 2664-2673

Monsen, R.B. (1979). Acoustic qualities of phonation in young hearing-impaired children. *Journal of Speech and Hearing Research*, 22, 270 - 288

Moore, B. (2003). Coding of sounds in the auditory system and its relevance to signal processing and coding in cochlear implants. *Otology and Neurotology*, 24,2, 243-254.

Most, T. and Frank, Y. (1994). The effects of age and hearing loss on tasks of perception and production of intonation. *The Volta Review*, 96, 137-149.

Most, T. (1999). Production and perception of syllable stress by children with normal hearing and children with hearing impairment. *The Volta Review*, 101, 2, 51-70.

Murphy, La B. M., McGarr, N.S. and Bell-Berti, F. (1990). Acoustic analysis of stress contrasts produced by hearing-impaired children. *The Volta Review*, Feb/Mar, 81-91.

Nikolopoulos, T.P., Archbold, S.M. and O'Donoghue, G. (1999). The Development of auditory perception in children following cochlear implantation. *International Journal of Pediatric Otorhinololaryngology*, 49 suppl.1, s189-s191.

Nolan, F. (1997). Speaker recognition and forensic phonetics. In W.J Hardcastle and J. Laver (Eds.), *The Handbook of Phonetic Sciences*. Oxford: Blackwell Publishers Ltd.

O'Connor, J.D. and Arnold, G.F. (1973). *Intonation of Colloquial English*. London: Longman.

O'Halpin, R. (1993). An auditory and acoustic analysis of contrastive stress in profoundly deaf children following training. *Child Language, Teaching and Therapy*, 9, 1.

O'Halpin, R. (1994). Nuclear prominence in Hiberno English: a preliminary investigation. *Teanga*, 14, 1-14.

O'Halpin, R. (1997). Contrastive stress in the speech of profoundly deaf children: a preliminary investigation. *The Volta Review*, 99, 89 - 103

O'Halpin, R. (2001). Intonation issues in the speech of hearing impaired children: analysis, transcription and remediation. *Clinical Linguistics & Phonetics*, 15, 7, 529-550.

Oller, D.K. and Eilers, R.E. (1988). The role of audition in infant babbling. *Child Development*, 59, 441-459.

Osberger, M.J. and McGarr, N.S. (1982). Speech production characteristics of the hearing impaired. In N. Lass (Ed.) *Speech and Language: Advances in Basic Research and Practice.* New York: Academic Press.

Papoussěk, M. and Papoušek, H. (1989). Forms and functions of vocal matching in interactions between mothers and their precanonical infants. *First Language*, 9, 6, 137-158.

Parker, A. 1999. *Phonological Evaluation and Transcription of Audio-Visual Language* (Oxford:Winslow)

Patel, R. and Grigos, M.I. (2006). Acoustic characterization of the question-statement contrast in 4,7 and 11 year-old children. *Speech Communication*, 48, 1308-1318

Peng, S-C., Tomblin, J.B., Cheung, H., Lin, Y-S. and Wang, L-S. (2004). Perception and production of Mandarin tones in prelingually deaf children with cochlear implants, *Ear and Hearing*, 25 (3), 251-264.

Peppé, S.M. Maxim, J. and Wells, B. (2000). Prosodic variation in Southern British English, *Language and Speech*, 43 (3), 309-334.

Peppé, S.M. and McCann, J. (2003). Assessing intonation and prosody in children with atypical language development: the PEPS-C test and the revised version. *Clinical Linguistics and Phonetics*, 27, 345-354.

Pierrehumbert, J.B. (1980). *The Phonology and Phonetics of English Intonation*. PhD Thesis, Massachusetts Institute of Technology, published 1988 by IULC.

Rahilly, J. (1997). Aspects of prosody in Hiberno-English. In J. Kallen (Ed.) *Focus on Ireland*, 109-132, Amsterdam Philadelphia, PA: John Benjamins Publishing Company.

Rahilly, J. (1998). Towards intonation models and typologies. *Journal of the International Phonetic Association*, 28, 73-82.

Rahilly, J. (1991). *Intonation Patterns in Normal Hearing and Postlingually Deafened Adults in Belfast*, unpublished PhD Dissertation, Queens University, Belfast.

Richardson, LM., Busby, P.A., Blamey, P. and Clark, G.M. (1998). Studies of prosody perception by cochlear implant patients. *Audiology*, 37, 231-245.

Roach, P. (1982). On the distinction between stress-timed and syllable-timed languages. In D. Crystal (Ed.) *Linguistic Controversies: Essays in Linguistic Theory and Practice in honour of F.R. Palmer*. London: Edward Arnold.

Rosen S. and Howell, P. (1991). *Signals and Systems for Speech and Hearing*. Academic Press.

Rubin-Spitz, J. and McGarr, N.S. (1990). Perception of terminal fall contours in speech produced by deaf persons. *Journal of Speech and Hearing Research*, 33, 174-180.

Semel, E., Wiig, E. and Secord, W. (1987). *Clinical evaluation of language fundamentals-revised.* London: The Psychological Corporation.

Shannon, R. V., Zeng, F. -G., Kamath, V., Wygonski, J. and Ekelid, M. (1995). Speech recognition with primarily temporal cues. *Science*, 270, 303-304.

Sharma, A. Dorman, M. F. and Spahr, A.J. (2002). A sensitive period for the development of the central auditory system in children with cochlear implants: implications for implantation. *Ear and Hearing*, 23, 6, 532-539.

Sharma, A. and Dorman, M. (2006). Central audition development in children with cochlear implants: Clinical implications. *Adv. Otorhinolaryngol,*, 64, 66-68.

Silverman, K. (1984). What causes vowels to have intrinsic fundamental frequency? *Cambridge Papers in Phonetics and Experimental Linguistics*, 3, 9-15.

Skinner, M.W., Clark, G.M, Whitford, L.A., Seligman, P.A., Staller, S.J., Shipp, D.B., Shallop J.K., Everingham, C., Menapace, C.M., Arndt, P.L., Antogenelli, T., Brimacombe, J.A., Pijl, S., Daniels, P., George, C.R., McDermott, H.J. and Beiter, A.L. (1994). Evaluation of a new spectral peak (SPEAK) coding strategy for the nucleus 22 channel cochlear implant system. *American Journal of Otology*, 15, 15-27, Suppl. 2

Skinner, M.W., Arndt, P.L. and Staller, S.J. (2002). Nucleus (R) 24 advanced encoder conversion study: performance versus preference. *Ear and Hearing*, 23, 1, 2S – 17S.

Snow, D. (1998). Children's imitations of intonation contours: are rising tones more difficult than falling tones? Journal of Speech Language, and Hearing Research, 41, 576-587

Snow, D. (2001). Imitation of intonation contours by children with normal and disordered speech. *Clinical Linguistics and Phonetics*, 15, 7, 567-584

Snow, D. and Balog, H. (2002). Do children produce the melody before the words? A review of developmental intonation research. *Lingua*, 112, 1025-1058.

Spaii, G. W.G., Derksen, E. S., Hermes, D. J. and Kaufholz, P. A. P. (1996). Teaching intonation to young deaf children with the intonation meter. *Folia Phoniatrica,* 41, 22-34.

Stackhouse, J. and Wells, B. (1997). *Children's Speech and Literacy difficulties: A Psycholinguistic Framework.* London: Whurr.

Svirsky, M. A., Teoh, S-W. and Neuburger, H. (2004). Development of language and speech perception in congenitally, profoundly deaf children as a function of age at cochlear implanation. *Audiology and Neuro-Otology*, 9, 4, 224-233.

Tait, M. and Lutman, M.E. (1997). The predictive values of measures of preverbal communicative behaviours in young deaf children with cochlear implants. *Ear and Hearing*, 18, 6, 472-478.

Thorsen, N. (1983). Two issues in the prosody of standard Danish. In A. Cutler and D. R. Ladd (Eds.) *Prosody, Models and Measurements*. Heidelburg: Springer-Verlag.

Titterington, J., Henry, A., Kramer, M. Toner, J.G. and Stevenson, M. (2006). An investigation of weak syllable processing in deaf children with cochlear implants. *Clinical Linguistics and Phonetics*, 20, 4, 249-269.

Vaissiere, J. (1983). Language-independent features. In A Cutler and D.R. Ladd (Eds.) *Prosody: Models and Measurements*. Heidelburg: Springer-Verlag.

Vandali, A.E., Whitford, L.A., Plant, K.L. and Clarke, G.M. (2000). Speech perception as a function of electrical stimulation rate; using the nucleus cochlear implant system. *Ear and Hearing*, 21, 6, 608-624.

Vogel, I. and Raimy, E. (2002). The acquisition of compound vs. phrasal stress: the role of prosodic constituents. *Journal of Child Language*, 29, 2, 225-250.

Waltzman, S.B. and Cohen, N.L. (2000). *Cochlear Implants*. New York: Thieme.

Wells, B., Peppé, S.M. and Goulandris, N. (2004). Intonation development from five to thirteen. Journal *of Child Language*, 31, 749-778.

Wells, B. and Local, J. (1993). The sense of an ending: a case of prosodic delay. *Clinical Linguistics and Phonetics*, 7.1, 59-73.

Wilson, B.S. (1997). The future of cochlear implants. *British Journal of Audiology*, 31, 205-225.

Xu, Y. (1999). Effects of tone and focus on the formation and alignment of F_0 contours. *Journal of Phonetics*, 27, 55-105.

Xu, Y. and Wallace, A. (2004). Multiple effects of consonant manner of articulation and intonation type on F_0 in English. *Journal of the Acoustical Society of America*, 115, part 2, 2397.

Xu, L., Li, Y., Hao, J., Chen, X., Xue, S.A. and Han D. (2004). Tone production in Mandarin-speaking children with cochlear implants: a preliminary study. *Acta Otolaryngol.*, 124, 363-367.

Xu, Y. and Xu, C.X. (2005*)*. Phonetic realization of focus in English declarative intonation*.* *Journal of Phonetics*, 33, 159-157.

Youdelman, K, McEachron, M. and McGarr, N.S. (1989). Using visual and tactile sensory aids to remediate monotone voice in hearing-impaired speakers. *The Volta Review*, 91, 197-207.