Copyright Statement

The digital copy of this thesis is protected by the Copyright Act 1994 (New Zealand).

This thesis may be consulted by you, provided you comply with the provisions of the Act and the following conditions of use:

- Any use you make of these documents or images must be for research or private study purposes only, and you may not make them available to any other person.
- Authors control the copyright of their thesis. You will recognise the author's right to be identified as the author of this thesis, and due acknowledgement will be made to the author where appropriate.
- You will obtain the author's permission before publishing any material from their thesis.

General copyright and disclaimer

In addition to the above conditions, authors give their consent for the digital copy of their work to be used subject to the conditions specified on the Library Thesis Consent Form and Deposit Licence.
PROSODY IN TYPICAL AND CLINICAL POPULATIONS: CHILDREN AND ADULTS WITH HEARING LOSS

ROSE THOMAS KALATHOTTUKAREN

A THESIS SUBMITTED IN FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY IN SPEECH SCIENCE
THE UNIVERSITY OF AUCKLAND 2015
Abstract

Aims: Aims of this doctoral thesis were to: 1) evaluate published tools for assessing prosodic skills in children and adults, 2) investigate age effects on different aspects of prosody perception in typically developing children and report normative performance for New Zealand-English speaking 7-12 year olds, 3) compare prosody perception and production in children with hearing loss and age- and gender-matched children with normal hearing, 4) examine the effects of age, hearing level, and musicality on children’s prosody perception, and 5) investigate prosody perception and musical pitch discrimination in adults using cochlear implants.

Methods: Published tools were identified through searching online databases, bibliographies of relevant articles and contacting authors. Six receptive subtests of Profiling Elements of Prosody in Speech-Communication (PEPS-C) and Child Paralanguage and Adult Paralanguage subtests of Diagnostic Analysis of Non Verbal Accuracy 2 (DANVA 2) were used as prosody perception measures. Musical pitch discrimination was assessed using Contour and Interval subtests of the Montreal Battery of Evaluation of Amusia (MBEA). Prosody productions were rated using the perceptual prosody rating scales.

Results: The literature review identified nine prosody assessment tools available for use with children and adults that were appraised for their intended purpose, target population, domains of prosody assessed, feasibility, and psychometric properties. The review highlighted the need to continue to develop and test tools for effective and comprehensive assessment of prosodic skills. The second study revealed that prosodic competence develops significantly between the ages of 7 and 11 years. Results from the PEPS-C test revealed a differential pattern of acquisition for different aspects of prosody, with 7-8 year olds being significantly poorer than 11-12 year olds on Chunking and Contrastive Stress subtests. Performance on the DANVA 2 test of affective prosody perception differed significantly across emotional...
categories (angry > happy > sad > fearful) and the level of emotion intensity (better scores for high emotion intensity items). The third study showed that children with hearing loss aged 7 to 12 years performed significantly poorer than controls on PEPS-C and DANVA 2 tests. Prosody perception scores were significantly correlated with age, hearing level, and musicality. Prosody production evaluated using perceptual rating scales showed greater variation in perceptual ratings of pitch, pitch variation, and overall prosody in the hearing loss group compared to the control group. Adults using cochlear implants performed significantly poorer than adult normative values reported for PEPS-C and DANVA 2 tests and the majority performed at chance on MBEA tasks.

**Conclusions:** The relatively small number of tools available to evaluate prosody compared to other aspects of language suggests that prosody is often overlooked in terms of formal language assessment. The normative results reported for New Zealand-English speaking children will be useful when assessing prosodic difficulties in children with hearing loss or autism spectrum disorder. Together, the studies on children and adults with hearing loss suggest that clinical assessment and therapy services for people with hearing loss should be expanded to target prosodic difficulties.
I acknowledge and thank “my Lord, my God” for provisions of joy, challenges and grace for growth.

I would like to express my sincere thanks to the following, truly amazing people without whom this thesis would have never been possible or completed. I have been truly blessed to have had two such wonderful supervisors - Professor Suzanne C. Purdy and Dr. Elaine Ballard. You have helped me throughout my PhD study and I am grateful for your support and encouragement. Suzanne, your words of wisdom have helped me make and shape decisions and I am honoured to have received your advice that will serve me a lifetime.

I owe a huge debt of gratitude to the Oticon Foundation Denmark for providing me funding to assist my PhD tuition fees, stipend, and other research expenses. I am grateful to the Deafness Research Foundation NZ, MG Martin Charitable Trust NZ and International Speech Communication Association for their financial aids to support my conference and research travel expenses.

Thanks to the members of the University of Auckland Human Participants Ethics Committee (UAHPEC), Auckland District Health Board (ADHB), Dr. Andrea Kelly and team (Greenlane Audiology), staff members at Hearing House, Ellen Giles (Northern Cochlear Implant Programme, UoA), Elizabeth Fairgray (Listening & Language Clinic, UoA), Barbara Greenhorn (Publisher of Coffee News - Remeura edition) and Jamie Mannion (Director of researchstudies.co.nz) for helping with participant recruitment. This doctoral thesis was formatted by Andrew Lavery at Academic Consulting, according to the University of Auckland Third Party Editing and Proofreading of Theses and Dissertations Guidelines.

I appreciate all the participants and their families for their valuable time and willingness to participate in my PhD project.
Many thanks to all Speech Science and Psychology staff who supported me throughout my PhD. Special thanks to Sue O’Shea, Sheryl Robertson, Michelle Scott, Kamalini Gnaniah, and Bronwyn Holcombe for your support.

Cheers to Abin, Oscar, Joan, Jae, Karen, Lynda, Masoomeh, Areej, Sylvia & Renique - no words can explain how much I enjoyed your companionship, I will miss you all.

Thanks to my family for your constant support and encouragement. Special thanks to Santheep for your love, patience, and support.
Table of Contents

Abstract .................................................................................................................................... ii
Acknowledgements ................................................................................................................. iv
List of Tables ........................................................................................................................... ix
List of Figures .......................................................................................................................... x
Co-Authorship Forms ............................................................................................................ xi

Chapter 1 Introduction ........................................................................................................... 1
Interpretations of prosody ...................................................................................................... 1
Communicative functions of prosody .................................................................................... 3
Approaches in prosody research ............................................................................................ 6
   Instrumental studies ............................................................................................................ 6
   Infant studies ......................................................................................................................... 8
   Electrophysiological and imaging studies ......................................................................... 8
   Psycholinguistic approaches .............................................................................................. 10
Prosody research in clinical populations .............................................................................. 10
Prosody in clinical settings .................................................................................................. 14
Aims of the thesis ................................................................................................................ 15
Overview of the thesis ......................................................................................................... 15

Chapter 2 Behavioural measures to evaluate prosodic skills: A review of
assessment tools for children and adults ............................................................................. 17
Introduction .......................................................................................................................... 17
Methods ................................................................................................................................ 22
   Literature search strategy .................................................................................................. 22
   Assessment tool inclusion/exclusion criteria ................................................................... 23
   Evaluation criteria ............................................................................................................ 23
Results and discussion ........................................................................................................ 23
   Intended purpose and target population ........................................................................... 29
   Domains of prosody assessed ........................................................................................... 30
   Norms, validity and reliability data .................................................................................. 32
   Feasibility .......................................................................................................................... 34
Conclusion ............................................................................................................................. 37

Chapter 3 Prosody perception abilities in typically developing school-aged
children ................................................................................................................................... 40
Introduction .......................................................................................................................... 40
Method ................................................................................................................................... 43
Participants ............................................................................................................................ 43
Materials ............................................................................................................................... 44
Table 1. Phonetic-acoustic and phonological correlates of prosody ............................................... 1
Table 2. The classification of prosodic functions adopted in this doctoral thesis ........................... 6
Table 3. Search terms .................................................................................................................... 22
Table 4. Features of the assessment tools that evaluate two or more aspects of prosody ............. 25
Table 5. Features of the assessment tools that evaluate affective prosody...................................... 27
Table 6. Summary of the features of prosody assessment tools .................................................... 37
Table 7. Participant characteristics ................................................................................................ 44
Table 8. Mean percent correct scores, standard deviations, medians, and ranges of scores for PEPS-C subtests by age group ........................................................................................................... 50
Table 9. Percentage of errors for each age group across the four emotional categories and two emotion intensities (24 items in total, 12 per intensity, 6 per emotion) on DANVA 2 subtest ........................................................................................................................................ 52
Table 10. Means and standard deviations (error scores) for DANVA 2 Child Paralanguage subtest by emotion intensity (low and high) and emotion categories .............................................................. 53
Table 11. Mean error scores and standard deviations on four emotional categories at two levels of emotion intensity for DANVA 2 ........................................................................................................... 53
Table 12. Emotion confusion matrix for the entire group of participants (N = 45) on DANVA 2 subtest (% confusions) ........................................................................................................................................... 54
Table 13. Gender wise comparisons using PEPS-C scores and DANVA 2 total scores ............... 54
Table 14. Characteristics of children with hearing loss ................................................................ 66
Table 15. Description of test materials .......................................................................................... 68
Table 16. Perceptual prosody rating scale ..................................................................................... 70
Table 17. Mean (SD) and Wilcoxon signed-rank test statistic for PEPS-C and DANVA 2 tests in control and hearing loss groups ................................................................................................. 72
Table 18. Correlations between age, hearing, musicality, and prosody perception (N = 32) ....... 74
Table 19. Stepwise regression analyses of prosody perception measures as a function of age (in years) and hearing level (in dB HL) ............................................................................................. 75
Table 20. PEPS-C total and DANVA 2 total scores of children with unilateral hearing loss ....... 76
Table 21. Inter-rater agreement between raters on perceptual prosody rating data ...................... 76
Table 22. Participant characteristics .............................................................................................. 88
Table 23. Individual test score profiles for CI participants ................................................................ 92
Table 24. Means, standard deviations, and ranges for scores on PEPS-C subtests and comparison of mean scores between CI users and adults with NH ................................................................. 95
Table 25. Individual error profiles of CI participants on the DANVA 2 Adult Paralanguage subtest ........................................................................................................................................... 97
List of Figures

Figure 1. Response alternatives (happy, sad, angry, and fearful faces respectively) for DANVA 2 Child Paralanguage subtest ........................................................................................................................................................................... 47

Figure 2. Means and 95% confidence intervals for PEPS-C subtests by age group. ................................. 49

Figure 3. Box plots representing percent correct scores obtained by two age groups on PEPS-C subtests. .................................................................................................................................................. 51

Figure 4. Mean percent correct scores and standard deviations on PEPS-C subtests for control vs. hearing loss groups. ........................................................................................................................................ 72

Figure 5. Mean scores on four emotion categories using DANVA 2 Child Paralanguage subtest for control vs. hearing loss groups. ........................................................................................................ 73

Figure 6. Mean scores on low and high emotion intensity stimuli of the DANVA 2 Child Paralanguage subtest for control vs. hearing loss groups. .................................................................................. 73

Figure 7. Mean ratings on pitch, pitch variations, and overall impression for control and hearing loss groups. ................................................................................................................................. 77

Figure 8. Percent correct scores obtained by CI users and adults with NH on four subtests of PEPS-C.................................................................................................................................................. 94

Figure 9. Mean and standard deviation of the DANVA 2 Adult Paralanguage subtest scores for listeners with NH (N = 12) (Saindon, 2010) (grey bar) and individual scores for CI users from present study (N = 12) (black bars) .................................................................................................................. 96

Figure 10. Mean scores and standard deviations for MBEA Contour and Interval subtests obtained by NH group (Peretz et al., 2003) and CI group from the current study ......................... 98
Co-Authorship Form

This form is to accompany the submission of any PhD that contains research reported in published or unpublished co-authored work. Please include one copy of this form for each co-authored work. Completed forms should be included in all copies of your thesis submitted for examination and library deposit (including digital deposit), following your thesis Acknowledgements.

Please indicate the chapter/section/pages of this thesis that are extracted from a co-authored work and give the title and publication details or details of submission of the co-authored work.

- **Chapter 2:**

<table>
<thead>
<tr>
<th>Nature of contribution by PhD candidate</th>
<th>Completed literature search, analyzed the literature, wrote article text</th>
</tr>
</thead>
<tbody>
<tr>
<td>Extent of contribution by PhD candidate (%)</td>
<td>90</td>
</tr>
</tbody>
</table>

**CO-AUTHORS**

<table>
<thead>
<tr>
<th>Name</th>
<th>Nature of Contribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Suzanne C. Purdy</td>
<td>Advice on literature analysis, edited article text</td>
</tr>
<tr>
<td>Elaine Ballard</td>
<td>Advice on article content, edited article text</td>
</tr>
</tbody>
</table>

**Certification by Co-Authors**

The undersigned hereby certify that:
- the above statement correctly reflects the nature and extent of the PhD candidate’s contribution to this work, and the nature of the contribution of each of the co-authors; and
- in cases where the PhD candidate was the lead author of the work that the candidate wrote the text.

<table>
<thead>
<tr>
<th>Name</th>
<th>Signature</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Suzanne C. Purdy</td>
<td></td>
<td>20/04/2015</td>
</tr>
<tr>
<td>Elaine Ballard</td>
<td>Click here</td>
<td>20/04/2015</td>
</tr>
<tr>
<td></td>
<td>Click here</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Click here</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Click here</td>
<td></td>
</tr>
</tbody>
</table>
This form is to accompany the submission of any PhD that contains research reported in published or unpublished co-authored work. Please include one copy of this form for each co-authored work. Completed forms should be included in all copies of your thesis submitted for examination and library deposit (including digital deposit), following your thesis Acknowledgements.

Please indicate the chapter/section/pages of this thesis that are extracted from a co-authored work and give the title and publication details or details of submission of the co-authored work.

Chapter 3:

<table>
<thead>
<tr>
<th>Nature of contribution by PhD candidate</th>
<th>Completed data collection, statistical analyses, wrote article text</th>
</tr>
</thead>
<tbody>
<tr>
<td>Extent of contribution by PhD candidate (%)</td>
<td>90</td>
</tr>
</tbody>
</table>

**CO-AUTHORS**

<table>
<thead>
<tr>
<th>Name</th>
<th>Nature of Contribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Suzanne C. Purdy</td>
<td>Advice on data collection and statistical analyses, edited article text</td>
</tr>
<tr>
<td>Elaine Ballard</td>
<td>Advice on data collection, edited article text</td>
</tr>
</tbody>
</table>

**Certification by Co-Authors**

The undersigned hereby certify that:
- the above statement correctly reflects the nature and extent of the PhD candidate’s contribution to this work, and the nature of the contribution of each of the co-authors; and
- in cases where the PhD candidate was the lead author of the work that the candidate wrote the text.

<table>
<thead>
<tr>
<th>Name</th>
<th>Signature</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Suzanne C. Purdy</td>
<td>[Signature]</td>
<td>20/04/2015</td>
</tr>
<tr>
<td>Elaine Ballard</td>
<td>[Signature]</td>
<td>20/04/2015</td>
</tr>
</tbody>
</table>

Lost updated: 25 March 2013
This form is to accompany the submission of any PhD that contains research reported in published or unpublished co-authored work. **Please include one copy of this form for each co-authored work.** Completed forms should be included in all copies of your thesis submitted for examination and library deposit (including digital deposit), following your thesis Acknowledgements.

Please indicate the chapter/section/pages of this thesis that are extracted from a co-authored work and give the title and publication details or details of submission of the co-authored work.

**Chapter 4:**

| Nature of contribution by PhD candidate | Completed data collection, statistical analyses, wrote article text |
| Extent of contribution by PhD candidate (%) | 85 |

**CO-AUTHORS**

<table>
<thead>
<tr>
<th>Name</th>
<th>Nature of Contribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Suzanne C. Purdy</td>
<td>Advice on data collection, peer-checked analysis of data, edited article text</td>
</tr>
<tr>
<td>Elaine Ballard</td>
<td>Advice on data collection, edited article text</td>
</tr>
</tbody>
</table>

**Certification by Co-Authors**

The undersigned hereby certify that:

- the above statement correctly reflects the nature and extent of the PhD candidate’s contribution to this work, and the nature of the contribution of each of the co-authors; and
- in cases where the PhD candidate was the lead author of the work that the candidate wrote the text.

<table>
<thead>
<tr>
<th>Name</th>
<th>Signature</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Suzanne C. Purdy</td>
<td>Signature</td>
<td>20/04/2015</td>
</tr>
<tr>
<td>Elaine Ballard</td>
<td>Signature</td>
<td>20/04/2015</td>
</tr>
</tbody>
</table>

Last updated: 25 March 2013
This form is to accompany the submission of any PhD that contains research reported in published or unpublished co-authored work. Please include one copy of this form for each co-authored work. Completed forms should be included in all copies of your thesis submitted for examination and library deposit (including digital deposit), following your thesis Acknowledgements.

Please indicate the chapter/section/pages of this thesis that are extracted from a co-authored work and give the title and publication details or details of submission of the co-authored work.

Chapter 5:

<table>
<thead>
<tr>
<th>Nature of contribution by PhD candidate</th>
<th>Collected data, statistical analysis, wrote article text</th>
</tr>
</thead>
<tbody>
<tr>
<td>Extent of contribution by PhD candidate (%)</td>
<td>90</td>
</tr>
</tbody>
</table>

**CO-AUTHORS**

<table>
<thead>
<tr>
<th>Name</th>
<th>Nature of Contribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Suzanne C. Purdy</td>
<td>Advised on data collection, peer-checked analysis of data, edited article text</td>
</tr>
<tr>
<td>Elaine Ballard</td>
<td>Advised on data collection, edited article text</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Certification by Co-Authors**

The undersigned hereby certify that:
- the above statement correctly reflects the nature and extent of the PhD candidate’s contribution to this work, and the nature of the contribution of each of the co-authors; and
- in cases where the PhD candidate was the lead author of the work that the candidate wrote the text.

<table>
<thead>
<tr>
<th>Name</th>
<th>Signature</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Suzanne C. Purdy</td>
<td></td>
<td>20/04/2015</td>
</tr>
<tr>
<td>Elaine Ballard</td>
<td>Click here</td>
<td>20/04/2015</td>
</tr>
<tr>
<td></td>
<td>Click here</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Click here</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Click here</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Click here</td>
<td></td>
</tr>
</tbody>
</table>

Last updated: 25 March 2013
Chapter 1
Introduction

Monrad-Krohn (1947) referred to prosody as a faculty of speech that constitutes the correct placement of stress upon syllables and words, natural rhythm, pauses and rate of speaking, and natural shifting of pitch from syllable to syllable and from word to word. In other words, prosody is the melody and rhythm of spoken language. Researchers working with different theories and frameworks have described prosody as a term that includes a wide variety of facts, concepts and phenomena. From a phonetic point of view, prosody is characterised by pitch, loudness, and length, also known as perceptual correlates of prosody. The acoustic correlates of prosody include fundamental frequency, intensity, and duration. From a phonological point of view, prosody conveys subtle changes in the meaning of spoken messages independent of words and grammatical order (Roach, 2000).

Interpretations of prosody
Researchers have approached prosody from differing points of view, including a phonetic and phonological perspective (Chun, 2002; Ladd, 2008). In general, the term ‘prosody’ includes the: 1) acoustic patterns of F0, duration, and loudness and 2) use of prosodic variations to convey communicative functions (Shattuck-Hufnagel & Turk, 1996). Table 1 show that pitch, loudness, and length are the phonetic aspects of prosody and fundamental frequency, intensity, and duration are their respective acoustic correlates.

Table 1. Phonetic, acoustic and phonological correlates of prosody

<table>
<thead>
<tr>
<th>Phonetic/perceptual</th>
<th>Acoustic</th>
<th>Phonological</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pitch</td>
<td>Fundamental freq</td>
<td>Intonation, Stress</td>
</tr>
<tr>
<td>Loudness</td>
<td>Intensity</td>
<td>Fluctuations in loudness</td>
</tr>
<tr>
<td>Length</td>
<td>Duration</td>
<td>Rhythm</td>
</tr>
</tbody>
</table>
Phonological correlates of prosody include the variation in pitch, loudness, and length (Table 1) that are produced in speech for conveying subtle changes in the meaning of spoken messages independent of words and grammatical order (Roach, 2000). Stress, rhythm, intonation are used to convey grammatical, emotional, and pragmatic functions of language. In English, contrastive pitch variations are used phonologically to convey discourse meaning, mark phrase boundaries (Gussenhoven, 2004), and prominence (Chun, 2002). Interestingly, the purpose of using contrastive pitch variations differs across languages. In languages such as Thai, Hausa (Nigeria), and Mixtec (Mexico) words are distinguished based on a set of distinctive pitch patterns or heights on each syllable whereas Swedish and Japanese make limited use of pitch to distinguish words (Nolan, 2006).

Phonetically, length refers to the acoustic duration of a sound. In phonological terms however, it refers to the relative duration of sounds and syllables (rhythm) when these are linguistically contrastive. There is no universal duration or range of duration for a particular sound. Depending on the language, both vowels and consonants can be phonologically long or short (Chun, 2002). For example, in American English vowel duration is phonetic in nature, i.e., vowels preceding voiceless consonants are shorter than those preceding voiced consonants. In contrast, in Malayalam (a south Indian language) a difference in meaning between words which otherwise have identical phonetic structure can be signaled using difference in vowel duration (/mala/ “mountain” and /ma : la/ “necklace”) (Jensen & Menon, 1972). In the Māori language (spoken by the indigenous population of New Zealand) the word “Inā” with a long vowel is used for emphasis (Inā te kino o ngā ara o reira - he kirikiri katoa, he kōpikopiko katoa; The roads there are really bad - they’re all gravel and windy) and to point to the reason for something special (E kore e tipu he paku aha i reira, inā te makariri; Nothing will grow there, for it’s too cold), whereas the word “Ina” with a short vowel means “when” (Ina puta au i aku whakamātau tau, kua haere au ki te whare wānanga; When I pass my exams, I’m going to university).
Loudness is a phonetic feature with intensity or sound pressure level as the acoustic correlate, which is the characteristic of sound that relates to the amount of energy present in the production of a sound. The fluctuations in loudness while speaking contribute to the perception of stress (emphasis) and emotion. The extent to which different acoustic cues dominate depends on the kind of prosodic contrast (Chun, 2002). Coutinho and Dibben (2012) reported that loudness, tempo and speech rate, melodic and intonation contour, spectral centroid, and sharpness are the psychoacoustic cues important for making judgements of emotion in speech. Timing and pitch cues are used to convey turn-taking during conversation (Hirschberg, 2002). Acoustic analysis of sentences spoken as questions or statements confirmed that in English the last bisyllabic word in a sentence carries dominant cues (F0, duration, and intensity patterns) for the contrast (Peng, Chatterjee, & Nelson, 2012). Peng et al. (2012) reported that the F0 contour is the primary cue for question-statement contrasts, with intensity and duration changes conveying important but less reliable information. This PhD thesis examines prosody in English.

Communicative functions of prosody

Communicative functions of prosody include conveying information about word boundaries, emphasis, and emotion. Prosodic cues in speech convey linguistic information such as chunking a stream of speech in phrases (McCann & Peppé, 2003), signaling new and contrastive information (Roach, 2000), and disambiguating sentences (Steinhauer, Alter, & Friederici, 1999). Prosody can provide paralinguistic information, including the identity, age, gender, and emotional state of the speaker (Paul, Augustyn, Klin, & Volkamar, 2005; Romero-Trillo & Newell, 2012; Stojanovik, 2010). Accurate production and perception of prosodic cues in spoken language are pertinent in conveying and understanding its communicative functions (Marslen-Wilson, Tyler, Grenier, & Lee, 1992; Plante, Holland, & Schmitthorst, 2006; Steinhauer et al., 1999). These communicative functions can be broadly
classified as grammatical, emotional (affective), and pragmatic, although there is a considerable overlap between these functions (Crystal, 1979; McCann & Peppé, 2003; Merewether & Alpert, 1990; Panagos & Prelock, 1997; Paul et al., 2005; Roach, 2000).

The grammatical functions of prosody are to: 1) determine the boundaries of phrases, clauses, or sentences, for example, fruit, salad and milk versus fruit-salad and milk (Wells & Peppé, 2003), 2) help the listener to distinguish between questions, statements, and commands (Roach, 2000), and 3) differentiate between word-classes, such as noun and adjective (récord versus recórd) (Klieve & Jeanes, 2001). Prosodic features can inform listeners of how an utterance should be parsed and interpreted. Parsing refers to the analysis of syntactic structure of a sentence (Eysenck & Keane, 2005). Snedeker and Trueswell (2003) demonstrated that prosodic cues were used by speakers when context failed to clarify the appropriate meaning of ambiguous sentences. They also reported that listeners utilised prosodic cues at a very early stage of parsing an ambiguous sentence. An event related potential (ERP) study of sentence processing by Steinhauer et al. (1999) reported that syntactic parsing is directly influenced by the prosodic information contained in the spoken sentence. Studies using behavioural and electrophysiological measures have reported that prosodic cues in spoken language were immediately used by listeners to disambiguate meaning when the words of a sentence alone were not sufficient to interpret the sentence (Marslen-Wilson et al., 1992; Plante et al., 2006; Steinhauer et al., 1999).

The emotional or affective function of prosody is the use of prosodic features such as pitch contour, pauses, and word stress to express emotions of speakers and their attitude to what they are saying (Peppé, 2009; Roach, 2000). Different prosodic patterns convey different emotions (Banse & Scherer, 1996; Juslin & Laukka, 2003). An extensive review and meta-analysis of 104 studies of vocal emotion expression conducted by Juslin and Laukka (2003) strongly suggested that there are emotion-specific acoustic patterns, and that discrete emotions in the voice are associated with certain “prototypical” acoustic cues that
vary by type of emotion. For example, mean F0, speech rate, and loudness are typically increased when speakers express happiness, anger, and fear, whereas they decrease when speakers convey sadness (Banse & Scherer, 1996; Juslin & Laukka, 2003; Pell, 2001; Scherer, Johnstone, & Klasmeier, 2003). In addition, expressions of happiness and anger exhibit high F0 variation, whereas fear and sadness tend to have low F0 variation (Banse & Scherer, 1996; Juslin & Laukka, 2003).

The use of prosodic patterns to convey discourse functions is the pragmatic aspect of prosody. Roach (2000) reported that intonation can be used to signal what is to be taken as “new” information and what is already “given”. A variety of prosodic features are used by speakers to mark the end of a conversational turn, indicating that, another person is expected to speak and a particular type of response is required and so on. For example, a rising intonation at the end of the utterance typically indicates that a response is required, and a falling intonation suggests end of conversation (McCann & Peppé, 2003). This is sometimes considered to be a grammatical function, as it identifies sentence-type. Higher, more variable pitch tones and longer pauses are typically seen at higher levels in the discourse hierarchy, for example, at topic shifts and the initial position in a paragraph (Noordman, Dassen, Swerts, & Terken, 1999; Smith, 2004). High pitch tones are used to introduce new topics and low pitch tones are used to indicate that the topical anaphor (evokes and further specifies the same situation) is in short-term memory (Wennerstrom, 2001).

Although prosody can be mainly classified as grammatical, emotional, and pragmatic (see Table 2), different authors have used other systems (often overlapping) for identifying prosodic functions. These include attitudinal, accentual, grammatical, and discourse functions (Roach, 2000) and grammatical, semantic, attitudinal, psychological, and social functions (Crystal, 1979). The approach adopted in the current thesis is summarised in Table 2.
Table 2. The classification of prosodic functions adopted in this doctoral thesis

<table>
<thead>
<tr>
<th>Functions</th>
<th>Description</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grammatical</td>
<td>a) Determine word boundaries</td>
<td>‘chocolate, biscuits, and jam’ vs. ‘chocolate-biscuits and jam’</td>
</tr>
<tr>
<td></td>
<td>b) Differentiate sentence type</td>
<td>question/statement/command/sarcasm</td>
</tr>
<tr>
<td></td>
<td>c) Distinguish word classes</td>
<td>‘greenhouse’ vs. ‘green house’</td>
</tr>
<tr>
<td>Emotional/Affective</td>
<td>a) Identify emotions</td>
<td>rise-fall pitch patterns indicate happiness; fall-rise pitch patterns indicate sadness</td>
</tr>
<tr>
<td>Pragmatic</td>
<td>a) Signal turn-taking</td>
<td>rising tone at the end of conversation indicates ‘your turn now’ and falling tone indicates ‘end’ of conversation</td>
</tr>
<tr>
<td></td>
<td>b) Emphasise relevant information</td>
<td>‘I wanted chocolate and honey’ vs. ‘I wanted chocolate and honey’</td>
</tr>
</tbody>
</table>

Approaches in prosody research

The phonetic and phonological analysis of prosody from the field of linguistics, phonetics, neuroscience, and psycholinguistics have used a wide range of methodological paradigms such as acoustic analysis of speech productions, direct measurement of articulator movements, and judgements and reaction times obtained during identification and discrimination tasks. In addition, measurements of brain activity and attention patterns in babies have been examined in response to prosodic features. Acoustic parameters (such as F0, duration patterns, and intensity or amplitude patterns) and characteristics of prosodic phenomena such as stress prominence or prosodic boundary have been widely investigated (Fry, 1955, 1958).

Instrumental studies

Instrumental approaches such as the PRAAT analysis tool (Boersma & Weenink, 2001) and EMU software (Bombien, Cassidy, Harrington, John, & Palethorpe, 2006) have been used to display and quantify the acoustic correlates of prosody (Shriberg, Kwiatkowski, Rasmussen, Lof, & Miller, 1992). Studies using acoustic measures have contributed in constructing acoustic profiles of dysarthria and apraxia of speech (Bunton, Kent, Kent, & Rosenbek, 2000; Kent & Rosenbek, 1983; Leuschel & Docherty, 1996; Schlenck, Bettrich, & Willmes, 1993).
Intonational phonology has been on focus of both production and perception research. This aspect of prosody focuses on how phonetic elements encode intonational contrasts. Intonation is described as patterns of F0 over spans of speech (prosodic phrases: intonational phrases and intermediate phrases) greater than a syllable or phoneme; thus intonational phonology does not modify the meaning of individual words but rather modifies the intended meaning of a phrase. The autosegmental metrical framework for intonational analysis (Beckman & Pierrehumbert, 1986; Pierrehumbert, 1980) and the related transcription systems such as Tone and Break Index (ToBI) and Intonational Variation in English (IViE) (Beckman & Elam, 1997; Grabe, Nolan, & Farrar, 1998) are used for prosodic labeling. One important innovation in intonational research, ToBI combines impressionistic transcription (using general phonetic categories in researcher’s mind which is language-independent and subject to revision) with acoustic analysis (Beckman, Hirschberg, & Shattuck-Hufnagel, 2004). Initially spectrograms with superimposed autocorrelation pitch tracks (i.e. estimates of the F0) are obtained and then the different kinds of tones are identified impressionistically based on the pitch tracks. Intonation is analysed in terms of two phenomena; pitch accents which are prominent pitch movements and edge tones which are the contours occurring at the end of ‘intonational phrases’ and ‘intermediate phrases’. IViE is based on the ToBI system and allows transcription of variations in rhythm, pitch accents, and tune structure. Unlike the original ToBI, IViE allows for directly comparable transcriptions of several varieties of English in a single labelling system. The ToBI system can explain the ‘pattern of intonation’ for example LH (low-high) pattern or HL (high-low) and why the patient’s speech patterns sound abnormal from a phonological perspective. However, Kent and Kim (2003) suggested that the ToBI system is not easy to apply clinically as it is time-consuming and requires lot of training.

Another important line of investigation in prosody research is the articulatory phonology framework that uses the kinematic data of articulator gestures obtained with a
magnetometer to study the intraoral dynamics (movements of the face, jaw, tongue, soft palate, and chest wall) during prosody production (Bryd & Saltzman, 1998). Several studies have investigated tonal-oral constriction alignment patterns for different languages (Italian, German, and Catalan) and found that the temporal coordination between pitch movements and articulatory gestures is stronger than that between acoustic events and F0 (D’Imperio et al., 2007; Mücke et al., 2006; Preito et al., 2007). Studies using facial and gestural articulatory analysis have been used as part of visual prosody research. Cavé et al. (1996) found an association between eyebrow movements and tonal rises in French and reported that eyebrow movements play a role in turn-taking during conversation.

**Infant studies**

Studies focusing on the attention patterns in infants have shown that babies are sensitive to prosodic cues from birth. Sandner (1981) reported that three months old infants were able to imitate pitch and duration patterns of maternal utterances. Newborns and young infants are able to discriminate different prosodic patterns at phrase (Christophe, Guasti, Nespor, Dupoux, & Ooyen, 1997) and word level (Nazzi, Floccia, & Bertoncini, 1998). Eye tracking experiments (Dehaene-Lambertz & Houston, 1998) and head-turning paradigms (Jusczyk, Cutler, & Redanz, 1993) have demonstrated a preference for native language in two and nine month old infants respectively. Both of these studies reported that this preference for native language was not found when the prosodic information was distorted. Subsequent studies of child language acquisition showed that infants use prosodic cues to segment a speech stream, and differentiate word categories and word order in their native language (Cristiá & Seidl, 2011; Morgan & Demuth, 1996; Newman, Ratner, Jusczyk, Jusczyk, & Dow, 2006).

**Electrophysiological and imaging studies**

A number of neurophysiological and neurobehavioural methods such as electroencephalography (EEG) and brain imaging techniques (e.g. fMRI) have been used to study various aspects of prosody processing. Studies have investigated the neural correlates
of intonational phrase boundaries using event-related brain potentials (Li & Yang, 2009) and have reported a right hemisphere dominance during the processing of intonation contrasts (Friederici & Alter, 2004; Gandour et al., 2003; Meyer, Steinhauer, Alter, Friederici, & Von Cramon, 2004). Neuroimaging studies have investigated the link between prosody and music (Buchanan et al., 2000; George et al., 1996; Koelsch et al., 2004; Pell & Leonard, 2003; Wildgruber et al., 2004; Zatorre, Evans, & Meyer, 1994) and showed that regions in the cortex involved in music comprehension are also involved in processing certain auditory components of prosody, such as intonation. Dramatic neural changes in auditory, cognitive, and linguistic processing in musicians relative to nonmusicians have been reported. Long-term musical training modulates representation of the human auditory cortex, which is reflected in auditory evoked potentials (AEP) and evoked fields of magnetoencephalography (MEG) (Pantev et al., 1998; Pantev, Roberts, Schulz, Engelien, & Ross, 2001). Pantev et al. (1998) compared auditory cortical representations to piano tones and pure tones in two groups of musicians (absolute pitch and relative pitch) and nonmusicians matched in age. They found that for musicians with absolute and relative pitch, the N1 amplitude was 25% significantly larger for piano tones than for pure tones whereas no significant difference was found for nonmusicians. The increase in N1 amplitude was correlated with the age at which the musicians began to practice; younger the subjects started playing their instrument; the larger was the cortical representation for piano tones compared to the pure tones. Structural brain plasticity induced by musical practice have been reported in cortical and subcortical structures using longitudinal and cross-sectional studies (Draganski & May, 2008; Hyde et al., 2009). Regarding the auditory cortex, Bermudez and Zatorre (2005) found greater grey matter volume in the right superior temporal gyrus of adult musicians compared with nonmusicians.
Psycholinguistic approaches

A psycholinguistic approach to prosody research was introduced recently (Preito, 2012; Wells, Peppé, & Goulandris, 2004). This approach focuses on using behavioural experiments (identification and discrimination tasks). An identification task requires listeners to categorise stimuli taken from a continuum whereas a discrimination task involves listeners to judge pairs of stimuli as being either same or different. Turk (2009) suggested that prosody evaluation should consider the various functions of prosody and assess if individuals can use and understand them effectively. For example, assessment of contrastive focus should inform clinicians whether or not the client is able to highlight the most important information in an utterance and get the meaning across to listeners, abstracting away from the specific pattern of F0, intensity, and duration parameters being used (Martens et al., 2011). Peppé (1998) developed the Profiling Elements of Prosody in Speech-Communication (PEPS-C) test based on the psycholinguistic framework. This test involves receptive and expressive prosody tasks to assess perception and production of communicative functions of prosody (Wells et al., 2004). The PEPS-C test provides options to assess perception and production of the grammatical, emotional, and pragmatic aspects of prosody using analogous receptive and expressive subtests. This doctoral thesis uses the assessments developed by Peppé (1998) and Nowicki and Duke (1994) to conduct a comprehensive assessment of prosody perception (involving grammatical, emotional, and pragmatic aspects) in typically developing children and children and adults with hearing loss using a psycholinguistic framework.

Prosody research in clinical populations

There are well-established terminologies and theoretical frameworks to explain prosody. Despite this, information about assessment and intervention for atypical prosody is scarce (Crystal, 2009; Diehl & Paul, 2009; Peppé, 2009). Studies have reported expressive and receptive prosodic difficulties in people with communication disorders such as hearing loss.
(Nakata, Trehub, & Kanda, 2012; Peng, Tomblin, & Turner, 2008; Torppa, Faulkner, Järvikivi, & Vainio, 2010, Torppa et al., 2014), autism spectrum disorders (ASD) (Mafezsky & Oswald, 2007; Peppé, Cleland, Gibbon, O’Hare, & Castilla, 2011), specific language impairment (Marshall, Harcourt-Brown, Ramus, & Van Der Lely, 2009; Samuelsson, 2009), Parkinson’s disease (Martens et al., 2011; Monetta, Cheang, & Pell, 2008), apraxia (Odell & Shriberg, 2001), and aphasia (Danly & Shapiro, 1982; Karow, Marquardt, & Marshall, 2001). Nevertheless, prosodic evaluations are not routinely performed in clinical settings (Crystal, 2009). Early studies used perceptual descriptions (i.e. feature classification) by speech-language therapists to describe expressive prosody characteristics of children with ASD and reported poor inflection, excessive or misassigned stress, slow syllable-timed speech, and a fast rate of speech or an adopted accent different from that of peers (Baron-Cohen & Staunton, 1994; Hargrove, 1997).

Several studies have used acoustic measures to create experimental stimuli in studies of prosody comprehension in ASD (Chevallier, Noveck, Happé, & Wilson, 2008; Järvinen-Pasley, Paisley, & Heaton, 2008) but only a handful of peer-reviewed published studies have acoustically examined prosody production in this population (Green & Tobin, 2009; Paul, Bianchi, Augustyn, Klin, & Volkmar, 2008). Paul et al. (2008) acoustically analysed imitated speech in 44 youth and adults with ASD and found that the ASD group showed significantly less difference in duration between stressed and unstressed syllable and a trend for larger pitch range than control group. Similar studies using acoustic analyses have reported wide range of prosodic contours (Green & Tobin, 2009), increased fundamental frequency variation (Diehl, Watson, Bennetto, McDonough, & Gunlogson, 2009), and longer utterance duration (Diehl & Paul, 2012, 2013) in children with ASD compared to controls. Gebauer, Skewes, Hørlyck, and Vuust (2014) investigated emotion recognition and neural processing of emotional prosody in high-functioning adults with ASD and reported that the ASD group...
showed a trend towards increased activation of the right caudate during processing of emotional prosody and rated the emotional intensity lower than control group.

Peppé, McCann, Gibbon, O’Hare, and Rutherford (2006) reported that children with high-functioning autism (HFA) performed significantly poorer than controls (6 to 13 years) on receptive and expressive prosodic tasks from the PEPS-C test. In contrast, Grossman, Bemis, Skwerer, and Tager-Flusberg (2010) investigated perception and production of lexical stress and processing of affective prosody in children and adolescents with HFA (7 to 18 years) and found that there were no group differences in the accuracy levels of perception of affective prosody and lexical stress. The HFA group also produced appropriately differentiated lexical stress patterns but demonstrated atypically long productions indicating reduced ability in natural prosody production. These findings are inconsistent with Peppé et al. (2006) study. This could be because participants in Grossman et al. (2010) study were older children and had better receptive vocabulary and verbal IQ skills compared to participants in Peppe et al. (2006) study. Shriberg et al. (2001) reported that the children with ASD used less appropriate prosodic phrasing (including misplaced lexical stress), slowed phrasing, and less appropriate resonance qualities than controls. Chevallier et al. (2008) assessed comprehension of aspects of grammatical prosody such as interpretation of word stress, grammatical pauses, and discrimination of question contours in adolescents with Asperger Syndrome (AS) and reported that AS group performed as well as typically developing controls. Peppé, McCann, Gibbon, O’Hare, and Rutherford (2007) reported that receptive prosody impairment in children with ASD has implications not only for understanding the various communicative functions of prosody but also for general language comprehension.

Most of the studies focusing on prosody in Parkinson’s disease (De Letter et al., 2007; Goberman, Coelho, & Robb, 2005; Skodda, Rinsche, & Schlegel, 2009) and apraxia of speech (Kent & Rosenbek, 1983; Skinder, Strand, Mignerey, 1999; Strand & McNeil, 1996)
have evaluated prosody in terms of its acoustic dimensions. In these studies, inappropriate word stress and monotonous speech were associated with apraxia of speech whereas monopitch, reduced stress, and variable rate was reported in Parkinson’s disease. Several studies have also investigated affective prosody as a language-related function of right hemisphere and affective prosody deficits have been reported in adults with acquired brain injury (Bowers, Blonder, Slomine, & Heilman, 1996; Ross, 1981, 1993; Ross, Thompson, & Yenkosky, 1997). Buxton, MacDonald, and Tippett (2013) reported impaired recognition of emotions from prosody and facial expressions in individuals with Parkinson’s disease. Impaired recognition of emotional facial expressions has been described in neuropsychiatric disorders (Kohler, Bilker, Hagendoorn, Gur, & Gur, 2000; see Kornreich & Philippot, 2006 for a review).

Prosody in people with hearing loss has been under-studied although speech and language impairments in individuals with hearing loss have been a topic of research for many years. There have been no comprehensive investigations so far into prosody production and perception in people with hearing loss (Chin, Bergeson, & Phan, 2012). In contrast, other aspects of language such as phonological development (Stelmachowicz, Pittman, Hover, Lewis, & Moeller, 2004), vocabulary (Moeller, 2000), and expressive language skills (Blamey et al., 2001; Nicholas & Geers, 2007; Ramkalawan & Davis, 1992; Svirsky, Toeh, & Neuburger, 2004; Yoshinago-Itano, Sedey, Coulter, & Mehl, 1998) have been extensively investigated in children with hearing loss. Peppé (2009) suggested that this relative neglect of prosody is a result of the somewhat invisible nature of prosodic errors compared to errors in pronunciation and word choice. Some studies have investigated specific aspects of prosody production such as grammatical characteristics, in particular the distinction between declarative and interrogative sentences using intonation contrasts in children with hearing loss (Loeb & Allen, 1993; Peng et al., 2008). In addition to grammatical aspects, prosody is also employed to convey emotional information and highlight prominence in speech (Roach,
2000; Wells et al., 2004). Nakata, Trehub, and Kanda (2012) reported that Japanese-speaking children with cochlear implants (CI) performed significantly poorer than children with normal hearing on affective prosody imitation and perception tasks. They found a significant correlation between perception and production results in child CI users. To our knowledge no study has examined production of vocal emotion by children with hearing loss. House (1994), Pereira (2000), and Hopyan-Misakyan, Gordon, Dennis, and Papsin (2009) reported difficulties on emotion perception tasks by adult CI users, however. Results from adult CI users cannot be directly translated to children as prelingually deafened children have no exposure to music and pitch information during a period of normal hearing, as do postlingually deafened adults, so they do not have a normal hearing auditory memory. They also lack exposure to high-resolution pitch cues, potentially impeding the development of pitch-related central auditory processing skills and they can be implanted at a young age when cortical plasticity is high, potentially enabling them to adapt differently to electric stimulation and have greater neuroplasticity (Looi, 2014).

**Prosody in clinical settings**

Auditory perception assessments are an integral component of clinical services related to CIs and hearing aids. However, aspects of prosody perception are not included in any of the commonly used speech perception assessments for adults (e.g. Consonant Nucleus Consonant (CNC) test, Hearing in Noise test, City University of New York Sentences, Sentences in Noise test) and children (e.g. Bamford-Kowal-Bench Sentences in Noise test, CNC test, Lexical Neighborhood test, Ling six sound test) with CIs.

Recent research proposes that prosody has influence on multiple areas of language and other related skills, which highlights the importance of early assessment and intervention for prosodic difficulties in children. Mellon (1999) suggested that inadequate recognition of a speaker’s emotion by children with hearing loss was misinterpreted as reflecting social
problems or lack of empathy by others, highlighting the importance of work in this area for children with hearing loss. Schorr, Roth, and Fox (2009) used a self-reported quality of life questionnaire and assessments of speech perception (single words) and emotion identification in 37 children with CIs. They reported that children’s increased quality of life predicted better performance on the emotion identification task. Peppé et al. (2007) reported that expressive prosody disorders add social and communication barriers for young children and these problems can persist even when other areas of language improve. Accurate prosodic sensitivity has also been related to successful reading (Goswami et al., 2002; Goswami, Gerson, & Astruc, 2010; Holliman et al., 2014; Lochrin, Arciuli, & Sharma, 2015; Schwanenflugel, Hamilton, Kuhn, Wisenbaker, & Stahl, 2004; Whalley & Hansen, 2006) and emotional development in children (Denham, 1998; Dunn, 2003). Appropriate use of prosodic features while reading out loud helps to preserve naturalness of voice and proper phrasing (Koriat, Greenberg, & Kreiner, 2002).

Aims of the thesis
This doctoral thesis aims to: 1) summarise and critically appraise the available behavioural prosody assessment tools for use with children and adults, 2) investigate prosody perception in typically developing New Zealand English-speaking children and to report normative performance for ages 7-12 years, 3) compare prosody perception and production in children with hearing loss with age- and gender-matched controls, 4) investigate the effects of age, hearing level, and musicality on children’s prosody perception, and 5) examine prosody perception and musical pitch discrimination in adults using CIs in relation to non-linguistic auditory measures, demographic variables, and speech recognition scores.

Overview of the thesis
A brief overview of this thesis follows. Four studies were conducted as part of this doctoral thesis and these are presented in Chapters 2-5. Chapter 2 is a literature search conducted to
identify and describe the tools developed to evaluate prosodic skills in children and adults. Nine assessment tools were identified and appraised for their intended purpose, target population, domains of prosody assessed, validity, reliability, feasibility, and normative sample data. Chapter 3 examined age effects on prosody perception in typically developing school-aged children. This study is the first of its kind to report normative performance on prosody perception tasks in New Zealand-English speaking children. The study described in Chapter 4 compared prosody perception and production in children with hearing loss and age- and gender-matched control children with normal hearing. It also investigated the effects of age, hearing level, and musicality on children’s prosody perception. Chapter 5 investigated prosody perception and musical pitch discrimination in adults using CIs.
Chapter 2

Behavioural measures to evaluate prosodic skills: A review of assessment tools for children and adults

Introduction

Over the past few decades, there has been considerable progress in identifying descriptive frameworks to explain prosody and its communicative functions. Progress has been made in the following areas: classifications of prosodic disorders, acquisition of prosody, neurological bases of prosodic impairments, significance of prosody in relation to language processing, and relationship between prosody and other aspects of speech (Crystal, 2009; Grigos & Patel, 2007, 2010; Patel & Brayton, 2009; Patel & Grigos, 2006; Peppé, 2009). However, not many studies have focused on the areas of prosody assessment and intervention, and little consideration has been given to prosody evaluation in clinical settings (Crystal, 2009; Peppé, 2009).

Peppé (2009) reported that the lack of clarity regarding prosodic terminologies, problems of identifying prosody as distinct from other aspects of communication, and lack of empirical information on the acquisition of prosody makes assessment difficult in disordered populations. Crystal (2009) pointed out that there is inadequate assessment, diagnosis, and treatment of prosodic impairments and negligence on the part of clinicians in making efforts to assess prosody in client populations. Following Crystal (2009), we found that there are no published papers that have systematically evaluated and compiled the various tools available to assess prosodic skills in children and adults. It is surprising that this lack of literature on assessment tools exists despite the fact that prosodic difficulties extend across a wide range of communication disorders such as ASD (Green & Tobin, 2009; McCann & Peppé, 2003), specific language impairment (Marshall et al., 2009; Stojanovik, Setter, & Ewijk, 2007), Parkinson’s disease (Martens et al., 2011), apraxia (Odell & Shriberg, 2001), aphasia (Danly
Shapiro, 1982), brain injury (Karow, Marquardt, & Marshall, 2001; Moen, 2009; Ross, Edmondson, Seibert, & Homan, 1988) and hearing loss (Nakata et al., 2012; Peng et al., 2008).

Theoretical approaches to explaining prosody (phonetic and phonological perspectives) and the differences in prosody error profile across clinical populations may have contributed to the development of different assessment approaches. Pitch, loudness, and length are the phonetic correlates of prosody. The physical correlates of these features are the speech’s fundamental frequency (F0), intensity, and phoneme and syllable duration, respectively. Phonological correlates of prosody include the variations in pitch, length, and loudness that are produced in speech for conveying subtle changes in the meaning of spoken messages independent of words and grammatical order (Roach, 2000). In other words, stress, and intonation are used to convey grammatical, affective, and pragmatic functions of language.

Methods for assessing disordered prosody can be classified into instrumental approaches, including software applications, and approaches to measure prosodic functions. Instrumental approaches involving acoustic analysis (e.g. PRAAT, Boersma & Weenink, 2001; EMU, Bombien et al., 2006) focus on displaying and quantifying the relevant acoustic correlates of prosody (Shriberg et al., 1992), whereas prosodic function measures are used to assess phonological or communicative aspects of prosody. The autosegmental metrical framework for intonational analysis (Pierrehumbert, 1980) and the related transcription systems such as ToBI and IViE (Grabe et al., 1998) provides options for prosodic labeling.

Earlier studies on communication disorders have evaluated prosody in terms of its acoustic dimensions. More recently, emphasising the role of prosody in communicative efficiency, the assessment of functions of prosody has been advocated rather than quantifying F0, intensity, and duration parameters (e.g. Profiling Elements of Prosody in Speech-Communication [PEPS-C], Peppé & McCann, 2003; Diagnostic Analysis of Nonverbal
Accuracy 2 [DANVA 2], Nowicki & Duke, 1994). However, the choice of instrumental or prosodic function measures to assess prosody should depend on the prosodic difficulties of the clinical population being assessed. Several studies have acoustically analysed echolalia, imitative and spontaneous speech in conversation, and narratives in children, adolescents and adults with ASD and have found important prosodic differences in the variance of F0, duration of syllables, the use of prosodic contours (Diehl & Paul, 2009; Green & Tobin, 2009), and coordination of prosodic cues such as pitch, amplitude, and duration (Van Santen, Prud’hommeaux, Black, & Mitchell, 2010).

Diehl and Paul (2013) conducted acoustic and perceptual measurements of prosody production by children with ASD and reported that differences in acoustic parameters were present in the speech of the ASD group even when the different aspects of prosody were perceived accurately. Inaccurate production of the acoustic features such as excessive or misassigned pitch, slow syllable-timed speech, fast rate of speech and monoloudness have been reported in individuals with apraxia of speech (Barry, 1995), Parkinson’s disease (Penner, Miller, Hertrich, Ackermann, & Schumm, 2001; Ma, Whitehill, & Cheung, 2010), and other communication disorders (Shriberg et al., 2001). Darley, Aronson, and Brown (1969, 1975) reported that prosodic deficits play a significant role in the characterisation of motor speech disorders. Prosodic deficits were also reported as a core feature of childhood apraxia of speech by the American Speech-Language-Hearing Association (ASHA, 2007). In general, inaccurate prosody productions affect an individual’s speech intelligibility (Chin, Bergeson, & Phan, 2012; Klopfenstein, 2009; Mayo, Aubanel, & Cooke, 2012).

This review focuses on assessment tools that are used to probe for the communicative aspects of prosody. The functions of prosody are identified at the indexical, grammatical, affective, and pragmatic levels of communication. Prosody provides indexical information, i.e. information related to the age, identity, and gender of the speaker (Paul et al., 2005; Romero-Trillo & Newell, 2012; Stojanovik, 2010). The grammatical functions of prosody
include: 1) determining the boundaries of phrases, clauses or sentences particularly when there is ambiguity, e.g. /FRUIT, SALAD and MILK/ versus /FRUIT-SALAD and MILK/ (Wells & Peppé, 2003) and 2) differentiating between word classes when there are homonyms, e.g. a noun and verb that sound the same but are differentiated by prosody (greenhouse versus green house; Klieve & Jeanes, 2001). The affective function of prosody is the use of prosodic features such as pitch contour, pauses, and word stress to express speaker’s emotions and attitudes (Peppé, 2009; Roach, 2000). Prosodic patterns convey different emotions (Banse & Scherer, 1996; Juslin & Laukka, 2003). Happiness, for example, is characterised by fast speaking rate, rising pitch, high variability, and fast voice onsets, and sadness is nearly the opposite (Hirschberg, 2002). The use of prosodic patterns to convey discourse functions is the pragmatic aspect of prosody. Prosody is used to signal to the listener what is to be taken as “new” information and what is already “given”. Prosody also helps the listener to distinguish between questions, statements, and commands (Roach, 2000). A variety of prosodic features are used by speakers to indicate to others that they have finished speaking, that another person is expected to speak, that a particular type of response is required, and so on. For example, a rising tone at the end of the utterance typically indicates that a response is required, and a falling tone suggests the end of conversation (McCann & Peppé, 2003).

It is important to note that researchers have reported the overlapping functions of these different aspects of prosody. For example, the use of prosody to distinguish between question and statement was classified as a grammatical function by Paul et al. (2005), whereas Wells and Peppé (2003) described it as the pragmatic aspect of prosody. Diehl and Paul (2009) reported that the use of prosody to distinguish between question and statement can fit into both categories as it conveys the sentence type and also signals a type of mental state or intent of discourse (either the end of a conversation or a particular response is required). Although there are differences in classification, it is widely recognised that the
accurate perception and production of the ranging aspects of prosody are a significant component of successful social communication (Aziz-Zadeh, Sheng, & Gheytnchi, 2010). Difficulties in perceiving the communicative aspects of prosody as well as deficits in prosody productions have been reported in children with hearing loss by several researchers (Meister, Landwehr, Pyshnoy, Walger, & von Wedel, 2009; Nakata et al., 2012; Peng et al., 2008). Hence, the combined use of instrumental techniques and prosodic function measures to assess prosodic skills is advised for this population.

Certain aspects of prosody are more relevant than others in specific client groups. For example, prosody evaluation in adults with neurological and psychiatric impairments has mainly focused on investigating the perception and production of affective prosodic skills (Moen, 2009; Wildgruber, Ethofer, Grandjean, & Kreifelts, 2009), whereas evaluation of grammatical, affective, and pragmatic aspects of prosody is relevant in individuals with hearing loss, ASD, and specific language impairment. The important communicative functions of prosody and evidence for prosodic deficits in various speech-language and hearing disorders makes it important to assess prosodic skills in typical and atypical populations.

Diehl and Paul (2009) compared three prosody assessment measures - the Prosody Profile (PROP; Crystal, 1982), Prosody-Voice Screening Profile (PVSP; Shriberg, Kwiatkowski, & Rasmussen, 1990), and PEPS-C to the other well-established methods of assessing language, such as the Peabody Picture Vocabulary Test-Fourth Edition (PPVT-1V; Dunn & Dunn, 2007) and the Clinical Evaluation of Language Fundamentals-Fourth Edition (CELF-1V; Semel, Wiig, & Secord, 2003). Diehl and Paul (2009) reported that there are no analogous data on the typical developmental sequence of prosody acquisition and adequate psychometrics derived from studies of spontaneous language use. Diehl and Paul also reported that there is a need for a prosody assessment tool that 1) has a representative normative sample and strong psychometric properties, 2) is based on empirical information
regarding the typical sequence of prosodic acquisition and is developmentally sensitive, 3) assesses various domains of prosody, 4) uses tasks that have high ecological validity, and 5) has established clinical utility.

We decided to review the available published assessment tools that are used to assess prosodic skills in children and adults, explaining the different domains of prosody that can be assessed using these tools, and providing information on the normative data and psychometric properties of each tool and its clinical utility. This review paper provides clinicians and researchers information for selection of appropriate tools for assessment of prosodic abilities.

**Methods**

*Literature search strategy*

A review of literature was conducted to identify clinical and research tools that have been developed so far to assess prosodic skills in children and adults. The goal was to identify tools with established reliability, validity, and normative data. The search terms shown in Table 3 were entered in different combinations into four online databases: ScienceDirect, SCOPUS, Web of Science, and PsycINFO. Manual searches of the bibliographies of published reviews and articles were conducted and test developers were contacted to gain additional information regarding the tools. Interlibrary loan facilities at the University of Auckland were used to access unpublished dissertation work.

**Table 3. Search terms**

<table>
<thead>
<tr>
<th>Concepts</th>
<th>Prosody</th>
<th>Assess</th>
<th>Impairment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Search words</td>
<td>prosody</td>
<td>assessments</td>
<td>impairment</td>
</tr>
<tr>
<td></td>
<td>intonation</td>
<td>measures</td>
<td>delay</td>
</tr>
<tr>
<td></td>
<td>suprasegmental</td>
<td>tests</td>
<td>disorder</td>
</tr>
<tr>
<td></td>
<td></td>
<td>evaluation</td>
<td>difficulty</td>
</tr>
<tr>
<td></td>
<td></td>
<td>tools</td>
<td></td>
</tr>
</tbody>
</table>
**Assessment tool inclusion/exclusion criteria**

We included assessment tools that were published in English and that were developed for use with individuals with acquired neurological disorders and psychiatric disorders, even though these tend to be components of a larger test battery and thus are less comprehensive. Tools that were translated from English to other languages or were developed in languages other than English were excluded from this review (e.g. Foley, Gibbon, & Peppè, 2011; Ladani et al., 2012; Martínez-Castilla & Peppè, 2008; Torppa et al., 2014; Van Zyl & Hanekom, 2013). Assessment tools that were out of print, such as the Right Hemisphere Language Battery (Bryan, 1989), which includes a subtest to assess the production of emphatic stress, were not described in this review.

**Evaluation criteria**

We appraised each assessment tool against a range of criteria pertinent to whether a tool might be suitable for clinical use. Clinical utility of each tool was gauged through examining features such as intended purpose, target population, domains of prosody assessed, validity and reliability data, nature of the normative sample, time and ease of administration and scoring, availability in different formats, and administration method.

**Results and discussion**

Nine assessment tools were identified: PROP, PVSP, PEPS-C, Perception of Prosody Assessment Tool (PPAT; Klieve, 1998), Minnesota Tests of Affective Processing (MNTAP; Lai, Hughes, & Shapiro, 1991), DANVA 2, Aprosodia battery (Ross, Thompson, & Yenkoshy, 1997), Florida Affect Battery (FAB; Bowers, Blonder, & Heilman, 1999), and Advanced Clinical Solutions (ACS; Pearson, 2009). We are not aware of any additional batteries that assess prosody.

The assessment tools were divided into two sections based on the number of aspects of prosody that they assess. Table 4 shows tools with subtests that assess two or more aspects of prosody; Table 5 shows tools that involve only affective prosody. Tables 4 and 5 also
summarize each tool’s target population, subtests involved, domains of prosody assessed, normative sample characteristics, and reliability and validity information available either in the literature or from the test developer. The tools differ in terms of target population, normative data, psychometric properties, and the domains of prosody assessed. The feasibility of each assessment tool is discussed based on factors such as time taken to administer, test format, appropriateness of test items, and ease of scoring.
### Table 4. Features of the assessment tools that evaluate two or more aspects of prosody

<table>
<thead>
<tr>
<th>Test</th>
<th>Age</th>
<th>Purpose and population</th>
<th>Subtests</th>
<th>Prosody subtests</th>
<th>Aspects of prosody</th>
<th>Normative sample and Inclusion criteria</th>
<th>Reliability /Validity</th>
<th>Feasibility</th>
</tr>
</thead>
<tbody>
<tr>
<td>PROP</td>
<td>Children &amp; adults</td>
<td>To obtain information about the expressive prosodic patterns encountered in a sample of clinical data</td>
<td><strong>Four prosodic patterns</strong>&lt;br&gt;Receptive: None&lt;br&gt;Expressive: Intonation (nuclear pitch direction), Tempo (phrasing), Stress (phrasal) &amp; strategies for producing stress</td>
<td>Pitch, tempo, stress</td>
<td>No normative, reliability &amp; validity data provided. Age ranges, guidelines &amp; examples of impaired prosody are provided.</td>
<td>Not available.</td>
<td>Time: depends on expertise in transcription&lt;br&gt;Format: manual scoring</td>
<td>Made for: clinical use</td>
</tr>
<tr>
<td>PVSP</td>
<td>3-81</td>
<td>To assess speakers’ prosody and voice in conversational speech.</td>
<td><strong>7 suprasegmentals (3 prosody, 4 voice)</strong>&lt;br&gt;Receptive: None&lt;br&gt;Expressive: Phrasing, Rate, Stress, Loudness, Pitch, Laryngeal quality and Resonance.</td>
<td>Phrasing, rate, pitch, stress, loudness, laryngeal quality and resonance</td>
<td>252 audiotaped exemplars from 3-19 year old children with normal &amp; disordered speech development were selected &amp; coded using Dictaphone 2550 audiocassette playback device. CSpeech by Milenkovic (1991) &amp; VOCAL (Milenkovic, 1989) were used for instrumental validity study.</td>
<td>Perceptual criterion validity for pitch, quality &amp; resonance ranged between 71-84%. Instrumental criterion validity for rate, stress, pitch &amp; quality ranged between 80-100%.</td>
<td>Time: depends on expertise in transcription&lt;br&gt;Format: manual scoring</td>
<td>Made for: clinical use</td>
</tr>
<tr>
<td>PEPS-C</td>
<td>5-14</td>
<td>To assess receptive and expressive prosodic skills in children. Widely used in children with autism spectrum disorders.</td>
<td><strong>12 subtests (6 receptive prosody &amp; 6 expressive prosody)</strong>&lt;br&gt;Receptive: Short item Discrimination, Short item Imitation, Turn-End Reception, Long item Discrimination, Affect Reception, Affect Expression, Long item Discrimination, Long item Imitation, Chunking Reception, Chunking Expression, Contrastive Stress Reception, Contrastive Stress Expression&lt;br&gt;Expressive: Short item Imitation, Turn-End Expression, Affect Expression, Long item Discrimination, Chunking Reception, Chunking Expression, Contrastive Stress Reception</td>
<td>Pitch direction, grammatical, affective (like, dislike), pragmatic</td>
<td>Normative data on 120 students from North London aged 5-14 years, English as first language, no identified speech &amp; language problems or educational problems, &amp; residents of England for at least 3 years. Normative data including mean &amp; SDs for ages 5.0, 8.0, 10.0, and 13.0 are provided for each subtest (Wells et al., 2004). The performance improved between the ages of 5.0 and 14.3</td>
<td>Test-retest reliability for 30 participants with a 6 month interval were not significantly different, the range of variation +2.08 to -1.04. Intra rater reliability was checked by rescoring 18/30 participants (3 month interval); difference in scores was 2.6%. Inter rater reliability was checked by obtaining 2 judgments on 16% of the production task results (N=30). Reliability of 80-98% reported. No significant difference in performance across gender &amp; order of presentation.</td>
<td>Time: ~60 min&lt;br&gt;Format: electronic administration; automatic scoring</td>
<td>Made for: clinical &amp; research use</td>
</tr>
<tr>
<td>Test</td>
<td>Age</td>
<td>Purpose and population</td>
<td>Subtests</td>
<td>Prosody subtests receptive/expressive</td>
<td>Aspects of prosody</td>
<td>Normative sample and Inclusion criteria</td>
<td>Reliability /Validity</td>
<td>Feasibility</td>
</tr>
<tr>
<td>---------</td>
<td>------</td>
<td>----------------------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------</td>
<td>------------------------------------------------------------------------------------------------------</td>
<td>-------------------</td>
<td>----------------------------------------</td>
<td>-----------------------</td>
<td>------------------------</td>
</tr>
<tr>
<td>PPAT</td>
<td>7-12</td>
<td>To evaluate prosodic perception in children using cochlear implants</td>
<td>Six subtests</td>
<td>Receptive: Apart from linguistic context: Pitch, Duration, Intensity Within linguistic context: question forms, statement forms, command forms; Tone &amp; affect, Grammatical class, Stress, Compound &amp; abutting words.</td>
<td>Pitch, duration, intensity, grammatical, affective (happy, sad, angry, sarcastic), pragmatics</td>
<td>6 children aged 7-12 years. Attended oral school for children with HL, used Nucleus 22 multichannel CI with SPEAK strategy &amp; more than one year of experience with the implant. Participants varied in age of onset of HL, etiology of deafness, length of profound deafness pre implant &amp; experience with the device (Klieve &amp; Jeanes, 2001).</td>
<td>CI participants perceived prosodic cues of duration, intensity &amp; pitch apart from a linguistic context above chance level, 70% or above. Performance on perceiving the prosodic cues meant within a linguistic context was at or just below chance level, between 60% and 70%.</td>
<td>Time: ~60 min Format: electronic administration; manual scoring Made for: research use</td>
</tr>
</tbody>
</table>
Table 5. Features of the assessment tools that evaluate affective prosody

<table>
<thead>
<tr>
<th>Test</th>
<th>Age (in years)</th>
<th>Purpose and population</th>
<th>Subtests</th>
<th>Prosody subtests receptive/expressive</th>
<th>Aspects of prosody</th>
<th>Normative sample and Inclusion criteria</th>
<th>Reliability /Validity</th>
<th>Feasibility</th>
</tr>
</thead>
<tbody>
<tr>
<td>MNTAP</td>
<td>6-11</td>
<td>To assess face perception and recognition of affective stimuli as conveyed through</td>
<td>16 subtests (4 auditory, 11 visual)</td>
<td>Receptive:</td>
<td>Affective-</td>
<td>67 children with ADHD and 38 control</td>
<td>MNTAP scores for</td>
<td>Time: ~2-3 hours</td>
</tr>
<tr>
<td>Lai et al.</td>
<td></td>
<td>facial expression, language and prosody</td>
<td>Training tasks, Inverted faces,</td>
<td>Prosody/content preference,</td>
<td>happy, sad,</td>
<td>subjects aged 6-11. Inclusion criteria</td>
<td>both groups were</td>
<td>Format: electronic administration;</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Identity match (1), Identity match (2), Faces teaching, Affect match,</td>
<td>Lexical comprehension, Prosody/</td>
<td>scared, neutral</td>
<td>for ADHD group included 1) a teacher</td>
<td>compared by sex &amp;</td>
<td>manual scoring</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Affect naming, Affect choice, Gesture recognition, Localization memory,</td>
<td>content congruence, Cross modal</td>
<td></td>
<td>correlated with age. No sex differences</td>
<td>correlated with</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Face &amp; object recognition memory, Sequential face pairs memory and four</td>
<td>matching</td>
<td></td>
<td>were found. Correlations with age were</td>
<td>age. No information</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>auditory receptive subtests.</td>
<td></td>
<td></td>
<td>highest with Inverted faces, Face and</td>
<td>on SES or ethnicity (Shapiro et al., 1993).</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Object recognition memory &amp; Localization</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>memory tasks. ADHD group differed from control group in tasks of prosody/content congruence &amp; cross-modal matching.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DANVA 2</td>
<td>3-99</td>
<td>To examine perception of facial expression, paralanguage (emotional aspect of prosody)</td>
<td>5 subtests (2 faces, 2 paralanguage &amp; 1 posture)</td>
<td>Receptive:</td>
<td>Affective-</td>
<td>Reliability: Internal consistency</td>
<td></td>
<td>Time: ~60 min</td>
</tr>
<tr>
<td>Nowicki &amp;</td>
<td></td>
<td>and understanding of body postures in children and adults.</td>
<td>Adult facial (AF) expressions 2, Child facial (CF) expressions 2,</td>
<td>Adult paralanguage 2, Child</td>
<td>happy, sad,</td>
<td>DANVA 2 AF- 158 college students; DANVA 2</td>
<td></td>
<td>Format: electronic administration;</td>
</tr>
<tr>
<td>Duke, 1994</td>
<td></td>
<td></td>
<td>Adult paralanguage (AP) 2, Child paralanguage (CP) 2, Adult postures 2</td>
<td>paralanguage 2</td>
<td>angry, fearful</td>
<td>2 CF-across 10 studies with children aged 4-16 years; DANVA 2 AP-33.5 years, N=20; DANVA 2 CP-8 year old (N=32) &amp; 10 year old (N=31); DANVA 2 adult postures- college students (N=54)</td>
<td></td>
<td>manual scoring</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Ethnicity comparable to community rates. Groups matched for SES &amp; IQ.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Mean &amp; SDs of errors on adult &amp; child facial expressions &amp; paralanguage subtests are provided for ages 3-99. Data for child postures subtest-5-14 years; adult posture subtest-15-50 years.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test</td>
<td>Age (in years)</td>
<td>Purpose and population</td>
<td>Subtests</td>
<td>Prosody subtests</td>
<td>Aspects of prosody</td>
<td>Normative sample and Inclusion criteria</td>
<td>Reliability /Validity</td>
<td>Feasibility</td>
</tr>
<tr>
<td>--------------------------</td>
<td>----------------</td>
<td>-----------------------------------------------------------------------------------------</td>
<td>---------------------------------------------------------------------------</td>
<td>------------------</td>
<td>-------------------</td>
<td>----------------------------------------</td>
<td>------------------------</td>
<td>-------------</td>
</tr>
<tr>
<td>Aprosodia battery</td>
<td>&gt;17</td>
<td>To examine reception and expression of affective prosody in adults with acquired neurological disorders</td>
<td>Four subtests (2 receptive &amp; 2 expressive)</td>
<td>Receptive:</td>
<td>Affective-happy, sad, angry, surprised, neutral, disinterested</td>
<td>22 brain damaged subjects and 16 control subjects mean aged 49.4±/− 15.2, M/F=9/7. Six weeks post stroke patients with unilateral hemispheric infarctions were included. Controls had no history of previous neurological problems, strongly right-handed as determined by a score of +70 on the Edinburgh Inventory (Oldfield, 1971). Groups were matched for sex &amp; education (Ross et al. 1997). Results confirmed that the mean coefficient of variation calculated is an appropriate measure of affective performance &amp; that the performance of control &gt; LHD &gt;&gt; RHD</td>
<td>Time: ~60 min</td>
<td>Form: electronic administration; manual scoring. Made for: research use.</td>
</tr>
<tr>
<td>FAB</td>
<td>&gt;17</td>
<td>To assess perception of affective prosody in adults with neurologic or psychiatric disorders.</td>
<td>10 subtests (5 facial, 3 prosodic, &amp; 2 cross modal).</td>
<td>Receptive:</td>
<td>Affective-happy, sad, angry, fearful, neutral</td>
<td>164 participants aged 17-85 years, right handed, primarily Caucasian, living in the Southeastern US, no psychopathology at the time of testing. Norms are provided for young adults (N=53, 18-30 years), middle aged adults (N=42, 31-60 years), older adults (N=49, 61-70 years) &amp; elderly adults (N=20, 71-84 years) for each of the 10 subtests (Bowers et al., 1999) &amp; individuals with neurologic disorders (Blonder, Bowers, &amp; Heilman, 1991; Bowers, Blonder, Slomine, &amp; Heilman, 1996).</td>
<td>2 weeks test retest reliability in young adults (N=20, 18-30 years) &amp; middle age adults in their early 50’s (N=12) ranged between 0.89 to 0.97.</td>
<td>Time: ~60 min</td>
</tr>
<tr>
<td>ACS</td>
<td>16-70</td>
<td>To assess social functioning deficits in adults.</td>
<td>4 subtests</td>
<td>Receptive:</td>
<td>Affective-happy, sad, angry, surprised, fearful, disgusted, neutral, sarcastic</td>
<td>800 participants aged 16-70 years, no history of neurological, psychiatric, developmental, or medical condition affecting cognitive functioning. Demographics matched to US 2005 census data for ethnicity &amp; education level.</td>
<td>Internal consistency: social perception total (a=0.70-0.84), prosody (a=0.64-0.79), and pairs (a=0.78-0.85).</td>
<td>Time: ~30-45 min</td>
</tr>
</tbody>
</table>
**Intended purpose and target population**

Tables 4 and 5 indicate the different purposes for which the tools were developed. These include a focus on assessing: 1) expressive prosodic skills, 2) perception of affective prosody in adults with neurologic and psychiatric disorders, 3) social and cognitive functioning in adults, and 4) prosodic skills in children with ASD and hearing loss. There are differences in the conceptual frameworks on which these tools are based: psycholinguistic (PEPS-C, PPAT), neurolinguistic (Aprosodia battery), and social and cognitive functioning (DANVA 2, MNTAP, FAB, ACS), reflecting the many identified roles of prosody in social communication.

The PEPS-C has been used in a number of research studies investigating prosodic skills in typically developing children and children with communication disorders (Catterall, Howard, Stojanovik, Szczerbinski, & Wells, 2006; Foley et al., 2011; Martínez-Castilla & Peppé, 2008; Peppé, & McCann, 2003; Peppé et al., 2007; Stojanovik, 2010). The original PEPS (Profiling Elements of Prosodic Systems; Peppé, 1998) test described in Peppé’s dissertation was designed for clinical use with adults; however, the present version of PEPS-C is specifically designed for use with children. Data on typical adults using PEPS are reported in Peppé, Maxim, and Wells (2000) and some clinical results for adults are reported in Peppé, Bryan, Maxim, and Wells (1997). Martínez-Castilla, Sotillo, and Campos (2011) used the Spanish version of the PEPS-C (Peppé et al., 2010) to assess the prosodic abilities of Spanish-speaking adolescents and adults with Williams syndrome. The PPAT and MNTAP have been used to assess prosodic impairments in children with hearing loss and attention deficit hyperactivity disorder (Klieve & Jeanes, 2001; Shapiro, Hughes, August, & Bloomquist, 1993). The Aprosodia battery, FAB, and ACS have been mainly used to assess affective prosodic difficulties in conditions such as brain injury and neurologic and psychiatric disorders in adults. Nowicki (2006) reported that the DANVA 2, which was developed for use with children and adults, has been widely used for research purposes in
various clinical groups, including hearing loss, ASD, traumatic brain injury, and learning disability.

Crystal (2009) reported that there is inadequate diagnosis in terms of identifying prosodic difficulties in children (mastering prosodic contrasts) and adults (managing the organizational role of prosody in speech production) and in determining the prosodic difficulties caused by access to limited auditory information in adults and children (e.g. due to hearing loss). Our review indicates that tools such as the PEPS-C, DANVA 2, PROP, and PVSP are applicable for wide age ranges (see Tables 4 and 5) and diverse clinical populations. There is, in our opinion, sufficient information to indicate the intended purpose and target population for these tools but a dearth of published studies exploring their clinical utility. Perhaps the best way forward is to encourage researchers to use these available tools to generate evidence for clinical use.

Domains of prosody assessed
Turk (2009) reported that an ideal prosody assessment tool should involve assessment of function, phonological representation, surface-level implementation, and perception of prosody (i.e. a comprehensive assessment should be able to evaluate receptive and expressive skills across different aspects of prosody). An important variation among the assessment tools reviewed here is the attributes of prosody that they evaluate. This review reveals PEPS-C as a comprehensive tool that is useful to assess the perception and production of different aspects of prosody. The receptive component of the PEPS-C includes subtests to assess sentence type (question versus statement; Turn-End Reception), speaker’s attitude (liking or disliking of food items; Affect Reception), phrase boundaries (the distinction between simple and compound nouns and groupings of adjectives; Chunking Reception), placement of contrastive stress/accent (Contrastive Stress Reception) and auditory discrimination for long (Long Item Discrimination) and short (Short Item Discrimination) tones. The expressive component includes six subtests analogous to the receptive subtests. The PPAT assesses the
perception of phonetic features of prosody such as pitch, intensity, and duration (Apart from Linguistic Context) and grammatical (Grammatical Class and Compound and Abutting Words), emotional (Tone and Affect), and pragmatic (Stress) aspects of prosody. Thus, the PPAT is useful for the comprehensive assessment of receptive prosodic skills in children. However, the PPAT does not assess expressive prosody. The MNTAP, FAB, ACS, and DANVA 2 only assess the perception of vocal emotions conveyed using prosodic cues. The Aprosodia battery includes receptive and expressive prosody subtests but assesses only affective prosody. The PROP and PVSP include only expressive prosody subtests to evaluate features such as pitch, phrasing, stress, and loudness.

The Aprosodia battery and the FAB focus on evaluating affective prosodic deficits in neurologic or psychiatric patients, however, difficulties in affective and other aspects of prosody are observed in other conditions such as hearing loss (Hopyan-Misakyan, Gordon, Dennis, & Papsin, 2009; Most & Peled, 2007). Individuals with sensorineural hearing loss have difficulties perceiving subtle changes in pitch, loudness, and duration (Moore, 1987; Moore & Carlyon, 2005), which are the major acoustic cues for perception of prosody in English. Inaccurate perception of these acoustic cues by individuals with hearing loss are manifested as difficulties in perceiving different aspects of prosody such as differentiating question from statement, word stress, distinguishing word/phrase boundaries, and vocal emotion recognition (Hopyan-Misakyan et al., 2009; Meister et al., 2009; Most & Peled, 2007). The PPAT uses a phonetic and phonological perspective to describe receptive prosodic skills in children using CIs. This is appropriate for individuals with hearing loss as it considers both acoustic factors and linguistic functions of prosody. Unfortunately, prospective trials to evaluate the clinical utility of this tool are missing. Thus, there is a lack of comprehensive, valid, and reliable assessment tools to assess different domains of prosody in individuals with hearing loss.
Norms, validity and reliability data

Normative data: None of the tools described in this review is standardised. However, norms are provided by test developers for the PEPS-C, DANVA 2, ACS, FAB, and PVSP. Norms for the PPAT, Aprosodia battery, and MNTAP were obtained from studies that used these measures to compare the performance of disordered populations and control groups (see Tables 4 and 5). Normative data for these tests were obtained for sample sizes ranging from 16 to 800. Diehl and Paul (2009) reported that compared to the PROP, PVSP, and PEPS-C, language assessment measures such as CELF-IV and PPVT were normed on larger normative samples with stratified norms based on gender, race, geographic location, and other factors (Dunn & Dunn, 2007; Strauss, Sherman, & Spreen, 2006). Stratified normative data are not available for the prosody assessment tools identified by this review. Clinicians need to be cautious in implementing the normative scores provided by the test developers to their target clinical population. For example, the PEPS-C was normed on a sample of 120 British-English speaking children ages 5-14 years, and the DANVA 2 Child Paralanguage subtest was normed on North Americans ages 8-10 years. These norms may not be appropriate for use with children who belong to different ethnic groups as there are linguistic and cultural prosody differences even among speakers of English (Coggshall, 2008).

Validity: The PVSP test developers report good instrumental and perceptual criterion validity (Shriberg et al., 1992). Instrumental procedures were used to estimate the criterion validity of more than 300 audiotaped exemplars that were selected to teach the coding procedures. Where the criterion validity of these perceptual coding decisions could not be determined by instrumental means, comparisons with the perceptual decisions of a panel of expert listeners were used. Very few studies have used the MNTAP, FAB, DANVA 2, PEPS-C, ACS, the Aprosodia battery, and PPAT tools to discriminate between typical and atypical populations (discriminant validity; Bowers et al., 1999; Klieve & Jeanes, 2001; Ross et al., 1997; Shapiro et al., 1993) which suggest a need for further validation.
One aspect of validity, face validity, can be addressed by determining relevance of test items to real-life communication. For example, the Affect and Turn-End Reception subtests of the PEPS-C use single-word test items (names of food items) rather than a sentence context. A positive feature of the ACS, DANVA 2, and FAB is that they use sentence-level stimuli, which are more naturalistic than word-level stimuli. While assessing prosodic skills, it would be appropriate to balance the advantages of psychometric robustness (using normed tests) against the advantages of ecological validity (as in careful analysis and profiling of naturalistic conversational data).

Reliability: Data on the reliability of the nine assessment tools reviewed in this study are shown in Tables 4 and 5. Reliability has generally been assessed using internal consistency measures; a measure based on the correlations between different items on the same test. For the DANVA 2 and ACS, internal consistency was demonstrated by computing Cronbach’s alpha. Further empirical confirmation of the tool beyond its initial construction by the original developers was undertaken for DANVA 2 (Nowicki, 2006). Internal consistency was measured using the co-efficient of variation for the Aprosodia battery. Test-retest reliability data are provided for the FAB and PEPS-C. Good intra rater reliability (i.e. degree of agreement among raters) and inter rater reliability (i.e. consistency of a measure when administered by different examiners) were reported for the expressive subtests of the PEPS-C by Wells, Peppé, and Goulandris (2004). This tool has been used by an increasing number of researchers (Foley et al., 2011; Martínez-Castilla & Peppé, 2008; Stojanovik, 2010).

There is considerable difficulty in directly comparing the different tools described in this review. For example, the sensitivity and specificity of each test varies according to the population studied and the cut-off scores considered abnormal (see Tables 4 and 5). Overall, this review supports the recommendation by Diehl and Paul (2009) that there is a need for a prosody assessment tool that 1) is standardised relative to a large representative normative
Feasibility

As noted by Crystal (2009), prosody is often neglected in terms of assessment, and “it is difficult to think of another medical area where a set of potentially relevant symptoms would be treated with such unconcern” (p. 257). Green and Tobin (2009) reported that a phonetic (surface-level features) and phonological (functions) analysis of prosody is useful in both typical and atypical speech. Given the clinical relevance of prosody, clinicians should take the time to investigate prosodic skills routinely in their client groups despite time constraints. However, practical considerations are important in a clinical setting, affecting both clients and professionals involved in the assessment process, especially as time is necessarily limited. Clinicians working in the field might need to consider having access to equipment such as a laptop, loudspeaker, sound level meter, and digital voice recorder that would be required to administer some of the tools (e.g. PROP, PVSP, PEPS-C, DANVA 2). Automatic and computerized tests such as PEPS-C and DANVA 2 are simple, the user manuals provide clear instructions, and a minimum amount of training is required for clinicians. In contrast, the MNTAP, PROP, and PVSP require manual scoring and prosody transcription. Other factors relevant to the clinical utility of the assessment tools include the time taken to administer the test, ease of scoring (manual or computerized), appropriateness of the test stimuli for use with children (colour photographs or black and white photographs), test format (computerized or paper pencil test), and age range. A clinician should consider these factors before selecting an assessment tool.

Administration time: The time taken to administer each test varies depending on the number of subtests involved and the population assessed (typical or disordered population). The PEPS-C is a comprehensive test that takes approximately 45-60 minutes to administer both receptive and expressive subtests. Wells and Local (2009) described the PEPS-C as not time
consuming, whereas Diehl and Paul (2009) considered it to be very long for a clinical measure. Peppé (2009) reported that the PEPS-C is short compared with the process of conversation analysis, and long considering that it tests only prosody. The DANVA 2, MNTAP, FAB, and ACS take relatively less time (approximately 20-30 minutes) compared to the PEPS-C, but only assess perception of affective prosody. The PPAT and Aprosodia battery take approximately one hour to administer. Longer test duration would not be appropriate for young children and some clinical populations such as attention deficit hyperactivity disorder or ASD. Tests that are in paper-pencil format and require manual scoring (MNTAP, FAB, PPAT) would take longer for the clinician than automatic and computerized tests (PEPS-C receptive subtests, DANVA 2). The time taken to transcribe prosodic elements (PROP, PVSP, expressive subtests of Aprosodia battery, and PEPS-C) would depend on the expertise of the clinician.

Appropriateness: The appropriateness of test stimuli should be considered, particularly when assessing children and atypical populations. Colour photographs as in PEPS-C and DANVA 2 would appear more realistic and have higher ecological validity (Diehl & Paul, 2009) than black and white photographs (MNTAP). The colour of test stimuli can have a positive effect on performance levels in children (Jeanes et al., 1997). The number of response items can affect the chance performance level and the cognitive demands of the task. The PEPS-C receptive subtests use a simple two-alternative forced choice format (50% chance performance level). The ACS, DANVA 2, and FAB tools have a minimum of four response options (25% or lower chance performance level), and hence may not be feasible for very young children and some clinical populations.

The version of English used to record the test stimuli can differ from that of the target population; hence locally developed norms may be required for prosody assessment tools using audio-recorded material. The DANVA 2 test stimuli are recorded by native English speakers from the United States, so its suitability for assessing prosodic skills in speakers of
other versions of English such as Australian or New Zealand-English needs further evaluation. A positive feature of the PEPS-C tool is that recordings are available in four different versions of English, including British, Australian, North American, and Scottish-English. Clinicians need to consider the cross-dialect prosodic variations while assessing speakers of Afro-Caribbean, Singaporean, or Indian versions of English.

Sensitive to development: Diehl and Paul (2009) reported that there is a need for a prosody assessment tool that is developmentally sensitive and can be used with different age groups. Of the tools reviewed here, only the DANVA 2 has different forms for different ages. Tools without different age versions can have ceiling effects for older children and adults and floor effects for younger children (e.g. Wells & Peppé, 2003). Also, different subtests may not be equally difficult for the age group tested, making it difficult for the clinician to determine relative strengths and weaknesses across areas for the purpose of intervention unless good normative data are available (Diehl & Paul, 2009). For example, the PVSP can be used for a wide age range but does not indicate what percentage of correct prosody would be appropriate for various age groups. Few studies have reported age-related developmental changes on the prosodic skills in children using the PEPS-C (Foley et al., 2011; Gibbon & Smyth, 2013; Wells et al., 2004). A summary of the results and discussion section is provided in Table 6.
Table 6. Summary of the features of prosody assessment tools

<table>
<thead>
<tr>
<th>Tools</th>
<th>Target population</th>
<th>Subtests involved</th>
<th>Psychometric data</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Children</td>
<td>Adults</td>
<td>Receptive subtest</td>
</tr>
<tr>
<td>PROP</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>PVSP</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>PEPS-C</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>PPAT</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>MNTAP</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>DANVA 2</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Aprosodia battery</td>
<td>-</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>FAB</td>
<td>-</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>ACS</td>
<td>-</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

* Norms for control groups were derived from previous studies; # few studies have used these tools to discriminate between typical and atypical populations.

Conclusion

SLPs frequently encounter prosodic impairments in persons with various communication disorders; hence, knowledge regarding assessment for persons with prosodic impairments is important. The aims of this review were to identify the tools that are available to assess prosodic skills in children and adults and to evaluate the clinical utility of the tools. In recent years, methodological paradigms such as acoustic analysis of speech productions, direct measurement of articulatory movements, judgments and reaction times obtained during identification and discrimination tasks, measurements of brain activity, and patterns of attention in babies have been used in prosody research (Prieto, 2012). However, these techniques are time consuming and not feasible in a clinical environment. This paper described the nine behavioural assessment tools that are available to aid clinicians who wish to examine prosodic skills in typical and disordered populations. The relatively small number of tools available to evaluate prosody compared to other aspects of language indicates that, although prosody is a topic which is clinically relevant, it is often overlooked in terms of
formal assessment. Clearly, there is no widespread recognition of the need for assessment of prosody, as evidenced by the small number of assessment tools available.

Consistent with Diehl and Paul (2009), this literature review identified very few assessment tools available to evaluate prosody compared to the large number of standardised assessment tools for other aspects of language like syntax, vocabulary, and phonology. Many existing prosody tools are narrow in scope or have not been robustly validated. This is an important gap that warrants future research. Most tools had been carefully constructed but lack generalizability across prosodic disorders, and five out of the nine tools identified focus on one particular aspect of prosody. If a clinician is working with client groups such as neurologic impairments and psychiatric disorders in which affective prosody is the main focus, then Table 5 will be useful. When selecting a tool, clinicians should consider the target population, time required to administer the tool, ease of administration and format of the tool, and access to equipment, as well as normative and reliability data. The various assessment tools reported in this review have explained prosody using different perspectives; hence, clinicians should have a specific idea about which aspect of prosody they want to assess and on which dialects of English it can be used safely.

Assessment of prosody is currently constrained by the lack of normative data. Three of the identified tools are devoid of norms. Only four tools are available which focus on two or more aspects of prosody. Among these, PROP and PVSP require a high level of expertise to transcribe prosodic elements and are time consuming. The PPAT was originally developed for use with children using cochlear implants and therefore focuses only on assessing receptive prosody skills. The PPAT may work well with children with hearing loss, but limited empirical data are available. The PEPS-C may be a good choice in clinical practice as it covers a wide age range, is easy to administer and score, provides a good user guide and manual, and has been used with a number of different clinical populations. However, its application to adults needs further investigation. The lack of knowledge of developmental
norms for different domains of prosody makes it difficult to derive standardised scores for these tools in children.

Some reasons why prosody assessment is neglected by clinicians may be due to lack of training or awareness of existing assessments, time constraints, lack of normative data for comparison with atypical populations, lack of evidence-based studies regarding the intervention of prosodic difficulties, and lack of culturally appropriate and developmentally sensitive measures. The remarkable lack of published studies on intervention for prosodic deficits might be due to the lack of appropriate assessment tools. This review highlights the needs for prosody assessment tools which are sensitive to developmental changes in children and that comprehensively and reliably assess relevant aspects of prosody in children and adults.

It is worth mentioning here that previous research explaining the articulatory movements and acoustic features associated with various prosodic contrasts (e.g. Grigos & Patel, 2007, 2010; Patel & Grigos, 2006; Snow, 1994, 1998) are not tied to any of the assessment tools described in this review, and a good way forward is to use these theoretical data to develop future tools. Future research should gather empirical data on the acquisition of prosody using the available tools and further explore the establishment of standardised diagnostic measures to evaluate prosody. Validity and reliability have been addressed to varying degrees – thoroughly in some tools and not at all in others.

The effectiveness of these tools in highlighting specific aspects of prosody warranting clinical intervention, and then guiding the intervention approach, has received little attention. Therefore, when considering these tools for the clinical assessment of prosody, caution is required to ensure that the time investment is warranted in terms of improved clinical outcomes for clients with communication difficulties. Given the considerable evidence for the importance of prosody in everyday communication, this is an important area for future work.
Chapter 3
Prosody perception abilities in typically developing school-aged children

Introduction

Prosody serves to convey emotions and attitudes (affective prosody), indicate question-statement contrasts, distinguish word boundaries (grammatical prosody), and emphasize new and relevant information (pragmatic prosody) of speech (Christophe, Gout, Peperkamp, & Morgan, 2003; Johnson & Jusczyk, 2001; Port, 2003; Roach, 2000; Wells et al., 2004). The use of high pitch at the end of the utterance to signal turn-taking (Couper-Kuhlen & Selting, 1996; Schegloff, 1996) and pitch accents to convey new and already given information are examples of the pragmatic functions of prosody (De Ruiter, 2014; Wonnacott & Watson, 2008). Research on affective speech has reported that prosodic cues are used to express vocal emotions and attitudes (d’Alessandro, 2006; Johnstone & Scherer, 2000; Murray & Arnott, 1993). Accurate recognition of vocal emotions is important from a developmental perspective because auditory signals can capture attention from someone who is not visually attending to the speaker, as occurs between infants and toddlers and their caregivers (Burnham, 1993; Fernald, 1993). Significant correlations between theory of mind (Dunn, 2003), verbal abilities (Rosnay, Fink, Begeer, Slaughter, & Peterson, 2014), academic achievement (Nowicki & Duke, 1994) and emotion recognition have been reported in typically developing children.

It is important to know how children understand different prosodic functions during communication at different ages and the degree of variability that might be expected within an age group. Several studies have examined production of prosodic contrasts (Ballard, Djaja, Arciuli, James, & Doorn, 2012; Grigos & Patel, 2007; Patel & Brayton, 2009; Patel & Grigos, 2006; Snow, 1994), but less is known about prosody perception in children (Crystal, 2009). Although prosodic difficulties have been reported in various communication disorders
(McCann & Peppé, 2003; Nakata et al., 2012; Peng et al., 2008; Stojanovik et al., 2007) there is a lack of normative data on prosody perception in typically developing children (Diehl & Paul, 2009). Assessment of prosodic skills in clinical settings is currently constrained by this lack of normative data (Crystal, 2009). The current study examined prosody perception abilities in 7-12 year old typically developing New Zealand-English speaking children using the receptive prosody subtests of the PEPS-C (Peppé & McCann, 2003) and the Child Paralanguage subtest of the DANVA 2 (Nowicki & Duke, 1994). Wells et al. (2004) used the PEPS-C test to examine perception of prosodic contrasts in typically developing British-English speaking children ($N = 120$, aged between 5-13 years). They reported that the ability to discriminate question-statement and like-dislike contrasts were mostly acquired by 8 years, however in their sample the ability to understand contrastive stress patterns and chunking continued to develop between 10 and 13 years. Gibbon and Smyth (2013) used the Irish-English version of PEPS-C and reported that 4 year old typically developing children performed at chance to weak ability levels on all subtests compared to 5-6 year olds. The current study is the first of its kind to report prosody perception abilities in New Zealand-English speaking children.

Children’s ability to understand low emotional intensity items is important given that in everyday settings emotional expressions are often subtle (Russell & Barrett, 1999). Differences between typical and atypical populations in recognizing emotions may also be less evident when emotional expressions are depicted with greater intensity than when less intense emotion is expressed. The DANVA 2 Child Paralanguage subtest can be used to examine the effects of level of emotion intensity and emotion category on production and perception of emotion. Mazefsky and Oswald (2007) reported that children with high functioning autism were less accurate in understanding vocal emotions expressed with low emotion intensity compared to children with Asperger’s syndrome and typically developing peers. There were no significant differences between the groups on perception of high
emotion intensity items. In an earlier study Mazefsky (2002) found that lower accuracy for 
DANVA 2 low intensity tone of voice cues was related to greater social impairment and 
lower social competence measured using Child Behaviour Checklist (Achenbach, 1991; 
Achenbach & Rescorla, 2001) and Scales of Independent Behaviour-Revised (Bruininks, 
Woodcock, Weatherman, & Hill, 1996). High emotion intensity facial expressions and tone 
of voice cues were not related to any of these measures. These findings are consistent with 
Baum and Nowicki’s (1998) findings that greater accuracy on DANVA 2 low emotional 
intensity items, but not high emotion intensity items, was related to better social competence 
teacher ratings using Child Behaviour Checklist) in typically developing 2nd-6th grade 
children. These few studies investigating the ability to recognise subtle vocal emotion cues in 
children suggest that this type of assessment could be valuable for early detection of impaired 
emotion processing.

There are differences in acoustic cues used to produce different emotions. For 
example, high values of F0 are used for anger, fear, and happiness, whereas low values of F0 
reflect sadness and disgust (Juslin & Laukka, 2001). Largest F0 variations (standard 
deviations) are reported for happiness, followed by anger, then disgust, and the smallest for 
sadness and fear. Anger and happiness are produced with high voice intensity, followed by 
disgust, fear, and sadness (Banse & Scherer, 1996). Juslin and Laukka (2001) reported the 
effects of emotion intensity on acoustic cues; F0 variations are greater for strong compared to 
weak intensity items, with largest effects for anger and disgust. There are differences in voice 
intensity, speech rate, pause proportion, attack time, and voice quality depending on the level 
of emotion intensity and emotion category. Juslin and Laukka (2003) reported that acoustic 
cues are used probabilistically and continuously so that cues are not perfectly reliable but 
have to be combined to produce or perceive emotions. They also suggested that the cues are 
combined in an additive fashion, and there is a certain amount of “cue trading” in emotional 
expressions. For example, if speakers cannot vary pitch to express anger, they may
compensate by varying loudness more. This means that in reality different emotions may be conveyed acoustically in a variety of ways. Hence, assessment of emotion recognition abilities in atypical and typical populations should include a range of different emotions at different levels of emotion intensity.

It may also be useful clinically to have access to tests of prosody perception that use different speakers and different prosodic tasks. The purpose of this study was therefore to 1) investigate receptive prosody skills in typically developing 7-12 year old children, and 2) report normative performance in New Zealand-English speaking children using PEPS-C and DANVA 2 tests.

**Method**

*Participants*

Forty-five typically developing children (21 boys and 24 girls) participated. Participants were selected by age to form three groups: 7 to 8 year olds ($M_{age} = 7.84, SD = 0.35$, age range: 7.34-8.68 years, $n = 14$), 9 to 10 year olds ($M_{age} = 10.13, SD = 0.59$, age range: 9.13-10.92 years, $n = 16$), and 11 to 12 year olds ($M_{age} = 11.90, SD = 0.49$, age range: 11.22-12.93 years, $n = 15$) (Table 7). Informed written consent was obtained from caregivers/parents and participation was voluntary. All children met the inclusion criteria of normal hearing (passed pure tone audiometry and immittance audiometry screening), spoke New Zealand-English as their primary mode of communication, and had no history of speech, language, and/or hearing difficulties as reported by parents. Testing took place in a quiet room.
### Table 7. Participant characteristics

<table>
<thead>
<tr>
<th>Age group</th>
<th>N</th>
<th>Gender</th>
<th>Age (in decimal years)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Boys/Girls</td>
<td>M</td>
</tr>
<tr>
<td>7-8</td>
<td>14</td>
<td>6/8</td>
<td>7.83</td>
</tr>
<tr>
<td>9-10</td>
<td>16</td>
<td>8/8</td>
<td>10.13</td>
</tr>
<tr>
<td>11-12</td>
<td>15</td>
<td>7/8</td>
<td>11.90</td>
</tr>
</tbody>
</table>

**Materials**

*Profiling Elements of Prosody in Speech-Communication (PEPS-C)*

Four receptive prosody subtests of PEPS-C (Turn-end, Affect, Chunking, and Contrastive Stress Reception) were used. These receptive subtests involve simple binary choices, with low memory and processing demands (Wells & Peppé, 2003). A conservative pass criterion of 75% was used based on the recommendation by Wells and Peppé (2003) to exclude chance scores. The review presented in Chapter 2 revealed that PEPS-C is a comprehensive tool that is useful for assessing the perception and production of different aspects of prosody. The receptive subtests of PEPS-C include options to assess sentence type (question versus statement; Turn-End Reception), speaker’s attitude (liking or disliking of food items; Affect Reception), phrase boundaries (the distinction between simple and compound nouns and groupings of adjectives; Chunking Reception), placement of contrastive stress/accent (Contrastive Stress Reception) and auditory discrimination for long (Long Item Discrimination) and short (Short Item Discrimination) tones.

1. **Turn-end Reception**: This subtest assesses the ability to differentiate question-statement contrasts. The names of food-items are presented as stimuli and the difference in tones indicates whether the item is ‘read’ or ‘stated’ as opposed to ‘offered’ or voiced as a question/inquiry.

2. **Affect Reception**: In order to assess the use of prosody to convey affective meaning, PEPS-C uses the distinction between expressing strong liking (happy) as opposed to reservation/dislike (sad). The test items used are names of food-items.
3. **Chunking Reception**: Chunking refers to boundary-signalling or prosodic delineation of the utterance into units for grammatical, semantic, or pragmatic purposes. PEPS-C uses the minor phrase boundaries that can be used to distinguish between items in a list. For example, colour combinations (PINK and BLACK&GREEN socks vs. PINK&BLACK and GREEN socks) or single and compound food-items (FRUIT, SALAD, and MILK vs. FRUIT-SALAD and MILK).

4. **Contrastive Stress Reception**: Contrastive stress refers to the speaker’s use of phonetic prominence to indicate which word or syllable is most important in an utterance. For example, BLUE and green socks (emphasis on the first colour) vs. blue and GREEN socks (emphasis on the second colour).

The pre-recorded auditory stimuli were presented using a laptop computer through a GENELEC 6010A active portable loudspeaker (placed directly in front of the participant) at a comfortable level in the normal conversational range (65 - 75 dB SPL) measured using a sound level meter at the position of the participant’s seat. The computer response screen of the PEPS-C involves a split-screen display of cartoon-type pictures. Participants were instructed to either point to the correct item on the screen or to give a verbal response. Before each task, demonstration items and practice items were played to ensure participants’ understanding of the task. Details of PEPS-C subtests and instructions for administration and scoring are described in Peppé and McCann (2003) and on the PEPS-C website (http://www.peps-c.com). Reviews of the strengths and weaknesses of the PEPS-C test are provided in Gibbon and Smyth (2013), Peppé (2009), and Diehl and Paul (2009).

**Child Paralanguage subtest of Diagnostic Analysis of Non Verbal Accuracy 2 (DANVA 2)**

The DANVA 2 test (Baum & Nowicki, 1998) was developed to measure competence in affect recognition by reading facial expressions and voice tone (affective prosody). It includes five subtests: 1) Child Faces, 2) Adult Faces, 3) Child Paralanguage, 4) Adult
Paralanguage, and 5) Child and Adult Posture. The current study used the Child Paralanguage subtest of DANVA 2 to assess emotion recognition using voice only. This 24-item subtest (4 alternative forced choice response paradigm) involved a sentence “I am going out of the room now but I will be back later” presented in happy, sad, angry, and fearful tones at two levels (high and low) of emotion intensity (12 items per intensity level) by male and female speakers (in random sequence). Low emotion intensity items have reduced acoustic saliency or cue intensity of emotional information compared to high emotion intensity items. DANVA 2 provides a more comprehensive assessment of affective prosody than PEPS-C. The PEPS-C Affect Reception task uses single word test items (names of food) rather than a sentence context, which is less likely to happen in real life situations (i.e., less ecological validity; Diehl & Paul, 2009). DANVA 2 uses sentence level stimuli to assess perception of four different emotions (happy, sad, angry, and fearful) whereas the PEPS-C Affect Reception subtest includes only two emotions (like/dislike).

The auditory stimuli were presented through a loudspeaker (using a similar procedure to the PEPS-C) and participants either gave a verbal response by saying if the person sounded happy, sad, angry, or fearful or pointed to the correct emotional smiley faces showing these emotions (Figure 1). Tables showing the number of errors for each emotion, number of errors for high and low emotion intensity items, number of errors for emotion by intensity, and the responses that were chosen when there was an error were generated using the DANVA 2 automatic scoring. Error profiles were used to identify the pattern of difficulty. The computer scoring programme provides error profiles as they are helpful for quickly identifying the patterns of difficulties that participants have.
Figure 1. Response alternatives (happy, sad, angry, and fearful faces respectively) for DANVA 2 Child Paralanguage subtest. After each stimulus is presented, participants made decisions to choose their responses from one of these emotions.

Statistical analyses

Nonparametric tests were used as the data were not normally distributed. Kruskal-Wallis ANOVA tests (Howell, 2014) were used to examine between group differences on PEPS-C and DANVA 2 scores. Post-hoc Mann Whitney U tests were conducted to investigate significant main effects. Friedman ANOVA was used to determine within group differences in scores across PEPS-C tasks and DANVA 2 emotional categories. Post-hoc analyses using Wilcoxon Signed-Rank tests were conducted to examine significant main effects. A Bonferroni correction factor was applied when multiple post-hoc comparisons were performed.

Results

Age group differences on PEPS-C receptive prosody tasks

When performance for the three age groups was compared using a Kruskal Wallis ANOVA significant main effects of age on Chunking ($\chi^2_{(2, 45)} = 13.15, p = .001$), Contrastive Stress ($\chi^2_{(2, 45)} = 13.14, p = .001$), and PEPS-C total scores ($\chi^2_{(2, 45)} = 21.79, p = .001$) were found. There were no effects of age group on Turn-end and Affect Reception scores (all $p > .300$). Post-hoc Mann Whitney U tests (significance value set at $p < .005$ ($.05/9$)) showed that scores obtained by 7-8 year olds were significantly poorer than those obtained by 9-10 and 11-12 year olds for Chunking ($p \leq .003$), Contrastive Stress ($p \leq .003$), and PEPS-C total ($p = .001$). There were no significant differences in scores obtained by the two older age groups across
PEPS-C tasks ($p \geq .072$). The PEPS-C data for the two older groups were therefore combined for further descriptive and statistical analyses. Mann Whitney U tests were used to investigate differences in performance between the youngest (7-8 years) and the combined older (9-12 years) age group for the four PEPS-C subtests and PEPS-C total (significance value set at $p < .01 \ (\text{.05/5})$). Scores obtained by 7-8 year olds were significantly poorer, with large effect sizes, than those obtained by the combined 9-12 year olds for Chunking ($U = 80.00$, $p = .001$, $r = .54$), Contrastive Stress ($U = 75.50$, $p = .001$, $r = .53$), and PEPS-C total ($U = 27.50$, $p = .001$, $r = .89$; Table 8). There were no significant differences in scores obtained by the two groups for Turn-end ($U = 140.50$, $p = .043$, $r = .30$) and Affect Reception tasks ($U = 161.00$, $p = .156$, $r = .21$). These results match those obtained when the three age groups were compared.

Table 8 shows the mean percent correct scores, standard deviations, and ranges of scores on PEPS-C tasks for the three age groups. Mean percent correct scores obtained by 7-8 year olds on PEPS-C tasks were lower than the scores for the combined 9-12 year olds ($M_{\text{age}} = 10.99$, $SD = 1.05$, $n = 31$; Figure 2). High standard deviations and the wide ranges of scores obtained by the youngest group indicate greater intersubject variability in their performance (Figures 2 and 3). Compared to 7-8 year olds, smaller standard deviations and narrow ranges of scores were obtained by 9-12-year olds across the PEPS-C tasks. Most children (90%) in the 9-12 year old combined older age group performed above the chance level of 75%, with most achieving ceiling scores on the four PEPS-C subtests (Figure 3). Outliers were present for three out of the four tasks for the older group, however. Thus, even though the majority of the older children are successful at each PEPS-C task, five children (3 boys and 2 girls) performed very poorly compared to their peers. Ceiling effects were found for all tasks for some of the younger children. Among the 7 to 8 year olds, below chance level performance ($< 75\%$, Wells & Peppé, 2003) occurred for one participant for the Turn-end and Affect Reception tasks and four participants for the Contrastive Stress Reception task.
Figure 2. Means and 95% confidence intervals for PEPS-C subtests by age group.
Table 8. Mean percent correct scores, standard deviations, medians, and ranges of scores for PEPS-C subtests by age group

<table>
<thead>
<tr>
<th>Age group</th>
<th>Turn-end</th>
<th>Affect</th>
<th>Chunking</th>
<th>Contrastive Stress</th>
<th>PEPS-C total</th>
</tr>
</thead>
<tbody>
<tr>
<td>7-8 years (n = 14)</td>
<td>M (SD)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>89.85 (10.58)</td>
<td>85.92 (9.67)</td>
<td>88.92 (7.94)</td>
<td>81.42 (13.25)</td>
<td>86.53 (4.59)</td>
</tr>
<tr>
<td>Mdn</td>
<td>94</td>
<td>88</td>
<td>91</td>
<td>84</td>
<td>87.75</td>
</tr>
<tr>
<td>Range</td>
<td>69-100</td>
<td>69-100</td>
<td>75-100</td>
<td>56-100</td>
<td>78.25-92.25</td>
</tr>
<tr>
<td>9-10 years (n = 16)</td>
<td>M (SD)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>96.93 (5.10)</td>
<td>88.87 (8.35)</td>
<td>97.37 (3.77)</td>
<td>95.06 (6.48)</td>
<td>94.56 (4.00)</td>
</tr>
<tr>
<td>Mdn</td>
<td>100</td>
<td>91</td>
<td>100</td>
<td>94</td>
<td>95.50</td>
</tr>
<tr>
<td>Range</td>
<td>81-100</td>
<td>75-100</td>
<td>94-100</td>
<td>75-100</td>
<td>82.75-100</td>
</tr>
<tr>
<td>11-12 years (n = 15)</td>
<td>M (SD)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>95.13 (5.35)</td>
<td>91.00 (6.27)</td>
<td>97.20 (4.45)</td>
<td>94.26 (6.48)</td>
<td>94.40 (3.04)</td>
</tr>
<tr>
<td>Mdn</td>
<td>94</td>
<td>94</td>
<td>100</td>
<td>94</td>
<td>94.00</td>
</tr>
<tr>
<td>Range</td>
<td>88-100</td>
<td>81-100</td>
<td>88-100</td>
<td>81-100</td>
<td>89.00-98.50</td>
</tr>
<tr>
<td>Combined 9-12 years (n = 31)</td>
<td>M (SD)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>96.06 (5.22)</td>
<td>89.90 (7.38)</td>
<td>97.29 (4.05)</td>
<td>94.67 (6.38)</td>
<td>94.48 (3.51)</td>
</tr>
<tr>
<td>Mdn</td>
<td>100</td>
<td>94</td>
<td>100</td>
<td>94</td>
<td>95.50</td>
</tr>
<tr>
<td>Range</td>
<td>81-100</td>
<td>75-100</td>
<td>88-100</td>
<td>75-100</td>
<td>82.75-100</td>
</tr>
</tbody>
</table>

Differences in performance based on PEPS-C task

Among the 7 to 8 year olds, there were no significant differences in scores across PEPS-C tasks ($\chi^2_{(3, 14)} = 5.347, p = .148$). However, there were significant differences in scores among 9-12 year old children ($\chi^2_{(3, 31)} = 22.568, p = .001$) depending on the task. Post-hoc analyses with Wilcoxon Signed-Rank tests (significance level set at $p < .008 (.05/6)$) showed that scores obtained by 9-12 year olds on PEPS-C Affect were significantly poorer than those obtained on PEPS-C Turn-end ($Z = -3.106, p = .002$) and Chunking ($Z = -3.856, p = .001$). There were no significant differences in scores between any other pairs of tasks ($p \geq .012$).
Figure 3. Box plots representing percent correct scores obtained by two age groups on PEPS-C subtests. The median scores are indicated by the thick horizontal line. Boxes indicate the data falling between the 25th and 75th percentile and the whiskers indicate the 95% confidence intervals.

**Age group differences for the DANVA 2 Child Paralanguage subtest**

The normative data for the DANVA 2 Child Paralanguage subtest are presented as error scores (Nowicki & Duke, 1994). Table 9 shows the percentage of errors made by three age groups of children on two levels of emotion intensity and four emotional categories. Overall more errors were made by 7-8 year-olds, followed by 9-10 year old children, with fewest errors were made by 11-12 year olds. Kruskal Wallis ANOVAs were used to determine the effects of age on DANVA 2 total errors (four emotions combined) across the two levels of emotion intensity (Tables 9 and 10). There were significant main effects of age for high emotion intensity errors ($\chi^2 (2, 45) = 6.831, p = .033$), but not for low emotion intensity errors ($\chi^2 (2, 45) = 3.404, p > .05$). Mann Whitney U tests (significance value set at $p < .016 (.05/3)$) showed that, for high emotion intensity, 7-8 year olds made more errors than 9-10 year olds ($U = 51.00, p = .008$) but did not differ from 11-12 year olds ($U = 80.50, p > .05$). For low emotion intensity items, total scores did not differ significantly across age groups ($p \geq .188$), although Table 11 shows a trend for fewer errors for the 11-12 year olds.
Table 9. Total errors (percentage) for each age group across the four emotional categories and two emotion intensities (24 items in total, 12 per intensity, 6 per emotion) on DANVA 2 subtest

<table>
<thead>
<tr>
<th>Age group</th>
<th>n</th>
<th>Happy</th>
<th>Sad</th>
<th>Angry</th>
<th>Fearful</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Low High</td>
<td>Low High</td>
<td>Low High</td>
<td>Low High</td>
<td>Low High</td>
</tr>
<tr>
<td>7-8 years</td>
<td>14</td>
<td>29 14</td>
<td>40 21</td>
<td>19 5</td>
<td>33 19</td>
<td>30 16</td>
</tr>
<tr>
<td>9-10 years</td>
<td>16</td>
<td>27 4</td>
<td>27 6</td>
<td>25 2</td>
<td>31 13</td>
<td>28 7</td>
</tr>
<tr>
<td>11-12 years</td>
<td>15</td>
<td>22 6</td>
<td>22 16</td>
<td>16 4</td>
<td>29 18</td>
<td>22 12</td>
</tr>
<tr>
<td>Combined 9-12 years</td>
<td>31</td>
<td>25 5</td>
<td>25 11</td>
<td>20 3</td>
<td>30 15</td>
<td>25 9</td>
</tr>
</tbody>
</table>

Differences in DANVA 2 scores based on emotion intensity and emotional category

Wilcoxon Signed-Rank tests showed that the error scores for high emotion intensity items ($M = 1.26$, $SD = 1.23$) were significantly lower than the error scores for low emotion intensity items ($M = 3.20$, $SD = 1.60$, $Z = -4.984$, $p = .001$; Table 11). A Friedman ANOVA showed significant differences between emotional categories ($\chi^2_{(3, 45)} = 10.881$, $p = .012$). Irrespective of the levels of emotion intensity, participants made more errors on items expressing fear, followed by sadness, then happiness, and had relatively few errors for anger (Table 9). Post-hoc analyses using Wilcoxon Signed-Rank tests revealed that the error scores obtained for fear stimuli were significantly higher than the error scores obtained for angry stimuli ($Z = -2.969$, $p = .003$).

Wilcoxon Signed-Rank tests (significance value was set at $p < .012$ (.05/4)) were performed to determine the effects of emotion intensity on the errors obtained within the four emotion categories. There was no significant difference between high and low emotion intensity error scores for fear ($Z = -2.439$, $p = .015$; Table 10). Error scores for the other three emotion categories were lower for high emotion intensity (happiness: $Z = -3.774$, $p = .001$; sadness: $Z = -2.641$, $p = .008$; anger: $Z = -3.977$, $p = .001$; Table 11).
Table 10. Means and standard deviations (error scores) for DANVA 2 Child Paralanguage subtest by emotion intensity (low and high) and emotion categories

<table>
<thead>
<tr>
<th>Age group</th>
<th>Emotion intensity</th>
<th>Happy (/3)</th>
<th>Sad (/3)</th>
<th>Angry (/3)</th>
<th>Fearful (/3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>7-8 years</td>
<td>Low M (SD)</td>
<td>0.85 (0.86)</td>
<td>1.21 (1.05)</td>
<td>0.57 (0.75)</td>
<td>1.00 (1.03)</td>
</tr>
<tr>
<td></td>
<td>Range</td>
<td>0-2</td>
<td>0-3</td>
<td>0-2</td>
<td>0-3</td>
</tr>
<tr>
<td></td>
<td>High M (SD)</td>
<td>0.42 (0.64)</td>
<td>0.64 (0.84)</td>
<td>0.14 (0.36)</td>
<td>0.57 (0.75)</td>
</tr>
<tr>
<td></td>
<td>Range</td>
<td>0-2</td>
<td>0-2</td>
<td>0-1</td>
<td>0-2</td>
</tr>
<tr>
<td>9-12 years</td>
<td>Low M (SD)</td>
<td>0.74 (0.68)</td>
<td>0.74 (0.81)</td>
<td>0.61 (0.66)</td>
<td>0.90 (0.83)</td>
</tr>
<tr>
<td></td>
<td>Range</td>
<td>0-2</td>
<td>0-2</td>
<td>0-2</td>
<td>0-3</td>
</tr>
<tr>
<td></td>
<td>High M (SD)</td>
<td>0.16 (0.37)</td>
<td>0.32 (0.59)</td>
<td>0.09 (0.30)</td>
<td>0.45 (0.56)</td>
</tr>
<tr>
<td></td>
<td>Range</td>
<td>0-1</td>
<td>0-2</td>
<td>0-1</td>
<td>0-2</td>
</tr>
</tbody>
</table>

Table 11. Mean error scores and standard deviations on four emotional categories at two levels of emotion intensity for DANVA 2

<table>
<thead>
<tr>
<th>Emotion intensity</th>
<th>Happy</th>
<th>Sad</th>
<th>Angry</th>
<th>Fearful</th>
<th>Total (/12)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low M (SD)</td>
<td>0.77  (0.73)</td>
<td>0.88  (0.91)</td>
<td>0.60  (0.68)</td>
<td>0.93  (0.88)</td>
<td>3.20 (1.60)</td>
</tr>
<tr>
<td>Range</td>
<td>0-2</td>
<td>0-3</td>
<td>0-2</td>
<td>0-3</td>
<td>1-7</td>
</tr>
<tr>
<td>High M (SD)</td>
<td>0.24  (0.48)</td>
<td>0.42  (0.69)</td>
<td>0.11  (0.31)</td>
<td>0.48  (0.62)</td>
<td>1.26 (1.23)</td>
</tr>
<tr>
<td>Range</td>
<td>0-2</td>
<td>0-2</td>
<td>0-1</td>
<td>0-2</td>
<td>0-5</td>
</tr>
</tbody>
</table>

Emotion confusion matrix

Table 12 shows the emotion confusion matrix for the entire group of participants (N = 45).

The emotion that was most correctly identified was anger (88%), followed by happiness (83%), then sadness (78%), and finally fear (76%). Fear and sadness were the emotions that participants had the most difficulty identifying. Fear was most often confused with sadness (15% of the error responses for fearful emotions were sad) and vice versa (12% of the error responses for sad emotions were fearful). The confusion matrix shows that the errors were not randomly distributed, some pairs of emotions were confused with one another more often than others.
Table 12. Emotion confusion matrix for the entire group of participants \((N = 45)\) on DANVA 2 subtest (% confusions)

<table>
<thead>
<tr>
<th>Stimulus</th>
<th>Happy</th>
<th>Sad</th>
<th>Angry</th>
<th>Fearful</th>
</tr>
</thead>
<tbody>
<tr>
<td>Happy</td>
<td>82.96</td>
<td>8.88</td>
<td>3.33</td>
<td>4.81</td>
</tr>
<tr>
<td>Sad</td>
<td>6.29</td>
<td>78.14</td>
<td>3.33</td>
<td>12.22</td>
</tr>
<tr>
<td>Angry</td>
<td>1.48</td>
<td>8.88</td>
<td>88.14</td>
<td>1.48</td>
</tr>
<tr>
<td>Fearful</td>
<td>5.55</td>
<td>15.92</td>
<td>2.22</td>
<td>76.29</td>
</tr>
</tbody>
</table>

Note. The percentage of correctly identified emotions is given on the main diagonal in boldface type.

Gender differences

Mann Whitney U tests showed no significant effects of gender for any PEPS-C task (all \(p > .868\); Table 13) or DANVA 2 subtest (all \(p > .161\)).

Table 13. Gender wise comparisons using PEPS-C subtests and DANVA total scores

<table>
<thead>
<tr>
<th>Gender</th>
<th>Turn-end</th>
<th>Affect</th>
<th>Chunking</th>
<th>Contrastive Stress</th>
<th>PEPS-C Total (% correct)</th>
<th>DANVA 2 Total (# errors)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>Mdn</td>
<td>IQR</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Girls</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(n = 24)</td>
<td>94.37</td>
<td>6.65</td>
<td>94.00</td>
<td>6.00</td>
<td>92.04</td>
<td>4.12</td>
</tr>
<tr>
<td></td>
<td>90.00</td>
<td>8.83</td>
<td>93.04</td>
<td>11.25</td>
<td>90.75</td>
<td>2.27</td>
</tr>
<tr>
<td></td>
<td>90.00</td>
<td>8.12</td>
<td>94.00</td>
<td>12.00</td>
<td>94.00</td>
<td>4.00</td>
</tr>
<tr>
<td></td>
<td>94.00</td>
<td>94.00</td>
<td>94.00</td>
<td>17.25</td>
<td>92.37</td>
<td>4.00</td>
</tr>
<tr>
<td></td>
<td>4.00</td>
<td>2.27</td>
<td>4.00</td>
<td>7.63</td>
<td>94.00</td>
<td>10.00</td>
</tr>
<tr>
<td>Boys</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(n = 21)</td>
<td>93.85</td>
<td>9.00</td>
<td>100.00</td>
<td>9.00</td>
<td>91.97</td>
<td>4.85</td>
</tr>
<tr>
<td></td>
<td>87.14</td>
<td>7.47</td>
<td>96.57</td>
<td>13.00</td>
<td>90.33</td>
<td>2.12</td>
</tr>
<tr>
<td></td>
<td>88.00</td>
<td>4.05</td>
<td>100.00</td>
<td>12.00</td>
<td>94.00</td>
<td>4.00</td>
</tr>
<tr>
<td></td>
<td>100.00</td>
<td>4.00</td>
<td>94.00</td>
<td>5.50</td>
<td>94.00</td>
<td>2.50</td>
</tr>
<tr>
<td></td>
<td>9.00</td>
<td>13.00</td>
<td>6.00</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. IQR = Interquartile range

Discussion

Overall the results indicate that prosodic competence develops significantly during the early school years. The younger children (7-8 year olds) performed significantly poorer than 9-12 year olds on PEPS-C Chunking and Contrastive Stress Reception tasks, indicating that some prosodic abilities assessed by the PEPS-C are still developing in young school-aged children. The reduced standard deviation scores and narrow ranges of scores obtained by 9-12 year olds compared to the youngest group indicate more reliable receptive prosody skills in older
children. Most children in the oldest group achieved ceiling scores on the four PEPS-C subtests. The lack of gender effects on PEPS-C or DANVA 2 performance is consistent with results reported by Wells et al. (2004) and Peppé et al. (2000).

Even though a general age-related improvement in perception scores was observed across PEPS-C tasks, there were variations in the acquisition pattern for different aspects of prosody. The older group performed significantly better than the 7-8 year olds on Chunking and Contrastive Stress Reception tasks. However, there were no significant differences between the older and younger age groups on Turn-end and Affect Reception tasks. Turn-end and Affect Reception tasks involve discrimination of simple pitch movements. The PEPS-C Chunking subtests require judging speakers’ use of timings cues and the PEPS-C Contrastive Stress subtest requires children to understand the use of accent/focus. These skills must be acquired later and/or more gradually in young school-aged children. This is consistent with Wells et al. (2004) who found that comprehension of chunking and contrastive stress continued to develop between 10 and 13 years in British-English speaking children.

Differential patterns for the acquisition of different prosodic skills are supported by prosody production studies. For example, Grigos and Patel (2010) investigated articulatory movements associated with the production of words with and without focus in 4, 7, and 11 year olds, and adults. Significant differences in duration, displacement, and velocity between focused and unfocused productions were seen between 7 and 11 year olds and adults, and there were differences between 11 year olds and adults. Grigos and Patel (2010) concluded that the ability to produce sentential stress starts to develop between 7 and 11 years and continues throughout adolescence. The developmental time course of different aspects of prosody perception has been investigated in a number of studies. Doherty, Fitzsimons, Asenbauer, and Staunton (1999) examined prosody perception in typically developing children ($N = 40$, aged between 5-9 years) using linguistic (discrimination of compound noun vs. noun phrase pairs and differentiation of questions/statements/commands) and affective
prosody tasks and found that perceptual abilities develop significantly up to 8.5 years. Vocal emotion recognition in children developed later than the corresponding linguistic ability. In the current study performance on the PEPS-C Affect Reception task was significantly poorer than performance on Turn-end and Chunking Reception tasks for the 9-12 year olds, indicating that the Affect Reception task was the most difficult task for the older children. For the 7-8 year olds all PEPS-C tasks were equally difficult.

Balogh, Swinney, and Tigue (1998) reported that the ability to respond to contrastive stress is related to a general sensitivity to prosodic cues and is distinct from syntactic and pragmatic knowledge. Contrastive stress scores improved significantly in the older children in the current study, but there were several outliers in the older age group indicating that not all children had mastered this skill. Consistent with this, Ito, Bibyk, Wagner, and Speer (2014) reported age-related improvements in interpreting contrastive stress in English-speaking children aged between 6 and 11 years, however even the 11 year olds showed delayed responses compared to adults. This suggests that it may take many years for children to fully acquire the prosodic meaning of pitch accent. Earlier mastery of the question-statement distinction compared to contrastive stress patterns could be related to greater exposure and familiarity effects. The infant directed speech literature suggests that motherese includes large amount of emotional information and utterances in the question-statement form (e.g. Durkin, Rutter, & Tucker, 1982; Soderstrom, Blossom, Foygel, & Morgan, 2008; Trainor, Austin, & Desjardins, 2000). In conversational English, stress usually occurs in the final word position of a sentence while the PEPS-C Contrastive Stress task uses stress in different word positions (e.g. I wanted BLUE and green socks (emphasis on the first colour) vs. I wanted blue and GREEN socks (emphasis on the second colour). This may not be the familiar pattern for children, which may contribute to the difficulty of the PEPS-C contrastive stress task.
DANVA 2 subtest results showed that 7-8 year olds made most errors in emotion recognition, followed by 9-10 year olds, and then 11-12 year olds. Although these findings suggest improvements in affect perception with increasing age in children, the effect of age on DANVA 2 scores did not reach statistical significance. Nowicki and Duke (1994) reported significant age-related changes in 6-10 year old American-English speaking children (N = 1001) on the DANVA 2 Child Paralanguage subtest, perhaps because of the inclusion of younger children and larger sample size. They also reported a strong correlation between vocal emotion recognition and academic achievement in children while DANVA 2 Child facial expression and posture recognition subtests did not show any correlation. Emotion processing in children has been mainly assessed through the visual modality by using facial expression tasks, and there has been less focus has been given to emotion recognition using voice. Difficulties in vocal emotion processing should be identified and treated early, however, in order to support children’s social and emotional development. The auditory system matures earlier than the visual system (Gottlieb, 1971; Jusczyk, 1998) and understanding of vocal emotion expressions plays a major role in early emotional development (Denham, 1998; Dunn, 2003). Halberstadt and Eaton (2002) reported that reduced family expressiveness of emotions through facial expressions and voice were associated with poor emotion understanding and expression in children.

DANVA 2 provides a more comprehensive assessment of affective prosody than PEPS-C. The PEPS-C Affect Reception task uses single word test items (names of food) rather than a sentence context, which is less likely to happen in real life situations (less ecological validity; Diehl & Paul, 2009). DANVA 2 uses sentence level stimuli to assess perception of four different emotions (happy, sad, angry, and fearful) whereas the PEPS-C Affect Reception subtest includes only two emotions (like/dislike). DANVA 2 errors varied considerably depending on the level of emotional intensity. Emotions presented at high intensities were recognised significantly better than those presented at low intensities for all
emotions, except for fear. These findings are consistent with Juslin and Laukka’s (2001) findings that adults (19 - 44 years) were able to detect happiness, sadness, anger, fear, and disgust vocal emotions presented at strong emotion intensity better than weak emotion intensity. Bänziger and Scherer (2005) found an increase in F0 mean and F0 range with increasing intensity, allowing easier detection of high emotion intensity stimuli. They found that F0 parameters such as mean, range, and minimum and maximum F0 peak for low emotion intensities (e.g. ‘sadness’, ‘calm joy’, and ‘anxious fear’) were generally lower than the F0 values for emotions with high intensities such as ‘despaired sadness’, ‘elated joy’, ‘panic fear’, and ‘hot anger’. It is important to know how well children understand low emotion intensity cues, as in real life situations emotional expressions are often subtle (Russell & Barrett, 1999). Emotion intensity has not been systematically varied in studies comparing atypical and typical populations. This is an important issue because emotion processing difficulties in atypical populations may be underestimated if only high intensity stimuli are used. Baum and Nowicki (1998) reported that accurate perception of low emotion intensity cues, but not high intensity cues, was related to social competence in typically developing children. Considering the level of emotion intensity as a factor may be useful in identifying typical error patterns associated with different disorders (e.g. Castelli, 2005; Greimel et al., 2010; Grossman & Tager-Flusberg, 2012; Mazefsky & Oswald, 2007).

The lowest accuracy on the DANVA 2 was observed for fearful emotions followed by sadness. Highest accuracy was noted for angry followed by happiness, consistent with the results from previous studies (Johnstone & Scherer, 2000; Juslin & Laukka, 2001). Bänziger and Scherer (2005) reported specific differences in F0 contours for different emotion categories that make certain emotions easier to identity than others. For emotions such as ‘hot anger’, ‘cold anger’, and ‘elation joy’ the F0 excursions in the second part of the utterance tend to be larger than for sadness or happiness. The shape of the F0 contour also changes depending on the emotion category; steeper final falls were observed for anger compared to a
progressive decrease (sadness) and increase (happiness) in F0 until the final fall. The additional F0 information associated with anger and happiness could be the reasons why these emotions are perceived more accurately than others by the children. Most of the confusions between emotions reported in the current study can be described as symmetrical (a term borrowed from Juslin & Laukka, 2001). For example, sadness was often confused with fear, and fear was confused with sadness. The same is true for sad and happy emotions and fear and happy emotions. These confusions mostly occurred for low emotion intensity items; suggesting that subtle acoustic cues are insufficient to accurately discriminate different emotions (Banse & Scherer, 1996; Bänziger & Scherer, 2005; Juslin & Laukka, 2001). Asymmetrical confusions were also present, such as anger was mostly confused with sadness, but sadness was rarely confused with anger. However, sadness was the most frequently chosen incorrect alternative. There is minimal research to suggest that there are differences in acquisition of understanding vocal emotions depending on the emotion category. Further research into the mechanisms by which children develop the ability to recognise different emotions may be useful for identifying therapy approaches to enhance affective prosody perception in young children and in children with prosody disorders.

**Conclusion**

The current study revealed a number of significant findings regarding prosody perception abilities in typically developing 7-12 year old children. Four receptive prosody subtests of PEPS-C and the Child Paralanguage subtest of DANVA 2 were used. This research is the first of its kind to report normative performance for prosody perception tasks in New Zealand English-speaking children using PEPS-C and DANVA 2 tests. Performance was affected by age but not gender. Overall findings are consistent with other studies indicating that prosodic competence develops significantly before 11 years. A differential pattern of acquisition for different aspects of prosody was found; chunking and contrastive stress reception skills
develop at a later age compared to turn-end and affect recognition. A trend for improvement in performance on DANVA 2 with age was observed, however this was not statistically significant. DANVA 2 scores varied depending on the level of emotion intensity, with high emotion intensity stimuli perceived more accurately than low emotion intensity items for all emotions.

Many researchers have reported that, of all the aspects of speech and language, prosody is less frequently assessed in clinical settings despite its key role in communication. Clinical assessment of prosody perception is currently limited by lack of normative comparison sample. The current study fills a gap in the literature by providing normative data for perception of different aspects of prosody in typically developing New Zealand-English speaking children. We hope that the results of this study will encourage clinicians to more routinely assess prosody perception abilities in clinical groups at risk for prosodic difficulties, such as children with hearing loss or autism spectrum disorder.
Chapter 4
Comparison of prosody perception and production in children with
hearing loss and age-and gender-matched controls

Introduction
Although the role of prosody in successful communication is well established (Peppé & McCann, 2003; Roach, 2000; Wells et al., 2004), prosody still remains a neglected area when assessing individuals using hearing aids and/or CIs. Auditory development in children with hearing loss, including the perception of prosody, depends on having adequate input from CIs and/or hearing aids (Gordon, Papsin, & Harrison, 2003). Lack of adequate auditory stimulation can lead to delayed speech and language development (Blamey et al., 2001; Rotteveel, Snik, Vermeulen, Cremers, & Mylanus, 2008; Tye-Murray, Spencer, & Woodworth, 1995). Nevertheless, prosody perception and production in people with hearing loss have received less attention than other aspects of language (Crystal, 2009; Diehl & Paul, 2009). The perception of auditory information conveyed through prosody using variations in the pitch, amplitude, and duration of speech (Coutinho & Dibben, 2012) is not usually evaluated clinically. Spoken language is introduced to the child as a vocal performance, and children attend to its musical features first. Brandt, Gebrian, and Slevc (2012) described language as a specific type of music in which referential discourse is bootstrapped onto a musical framework.

Developments in hearing technology have provided enhanced understanding of speech in quiet and in noise (Hällgren, Larsby, Lyxell, & Arlinger, 2005; Müller, Schön, & Helms, 2002; Ricketts, Lindey, & Henry, 2001; Valente, Schuchman, Potts, & Beck, 2000) and sound localisation (Dunn, Perreau, Gantz, & Tyler, 2010; Grieco-Calub & Litovsky, 2010; Nawaz, McNeill, & Greenberg, 2014) for people with hearing loss. However, there is limited evidence for improved access to prosodic cues in speech. Children using CIs can
differentiate utterances on the basis of number of syllables (Carter, Dillon, & Pisoni, 2002; Most & Peled, 2007), but they have considerably more difficulty differentiating utterances with contrastive intonation contours or affective prosody (Most & Peled, 2007; Hopyan-Misakyan et al., 2009; Peng et al., 2008). Although children with hearing loss aged between 7 to 15 years could distinguish statements from yes/no questions at modest but above chance levels (Most & Peled, 2007; Peng et al., 2008), they were unable to distinguish among happy, sad, angry, and fearful versions of the same sentences (Hopyan-Misakyan et al., 2009; Nakata et al., 2012). To date, no study has comprehensively investigated different aspects of prosody perception in children with hearing loss.

Prosody serves various communicative functions such as question-statement distinctions, vocal emotion recognition, differentiating word boundaries, and understanding the use of contrastive stress (Roach, 2000; Wells et al., 2004). Accuracy in perceiving the emotional content of speech at an early age has been linked to emotional and social development in children (Mellon, 2009; Schorr et al., 2009). Mellon (2009) reported that lack of understanding of emotional cues during conversation by children with hearing loss was misunderstood as lack of empathy and social problems. Accurate emotion recognition is associated with better quality of life in children using CIs (Schorr et al., 2009). Previous research suggests that prosodic sensitivity facilitates children’s reading development (Miller & Schwanenflugel, 2008; Whalley & Hansen, 2006) and language acquisition (Morgan & Demuth, 1996; Soderstrom, Seidl, Nelson, & Jusczyk, 2003; Jusczyk, Houston, & Newsome, 1999; Thiessen, Hill, & Saffran, 2005). Therefore, assessment and intervention for prosodic difficulties should be considered in children with hearing loss, who are at risk for reading and language delay (Allen, 1986; Geers, Tobey, Moog, & Brenner, 2008; Moog & Geers, 1985).

Typically developing children produce prosodic features of speech in their vocalisations from a very young age (Jusczyk, 1993; Jusczyk et al., 1993; Snow, 1994, 1998). Prosody perception and production in typically developing children improve with age (Patel
Perception and production of prosodic features may be challenging for children with hearing loss, due to the limitations of CIs in presenting pitch information and the general limitations of the human auditory system with electric simulation (Kalathottukaren, Purdy, & Ballard, 2015; Wilson et al., 1991; Vandali, Whitford, Plant, & Clark, 2000). This is discussed further in Chapter 5. Peng et al. (2008) reported that children and young adults with CIs \((n = 26, \text{aged } 7.44 - 20.74 \text{ years})\) performed significantly poorer than controls \((n = 17)\) on perception and production of question versus statement contrasts. Moderate correlations between production and perception of prosodic contrasts were found in this study. Chin, Bergeson, and Phan (2012) used a sentence imitation task to examine how adult listeners judged the intelligibility of prosody productions (declarative, interrogative, happy, and sad sentences) in children with CIs compared to controls. Children with CIs were rated significantly lower than controls in the production of interrogative sentences; no significant differences were found for other ratings. These findings have implications for aural habilitation and speech-language therapy intervention for children with prelingual deafness using CIs. Few studies in this area have included children using hearing aids. Few intervention studies have targeted prosodic difficulties in children with hearing loss compared to other aspects of language. This may be due to the lack of information regarding the nature of prosodic difficulties in children with hearing loss. Peppé (2009) reported that this relative neglect of prosody was a result of the somewhat invisible nature of prosodic errors compared to errors in pronunciation and word choice. Samuelsson (2010) conducted a six-week evidence-based intervention for prosodic production difficulties in a child with language impairment and reported improved prosodic skills at word and phrasal level. Klieve and Jeanes (2001) reported that children using CIs improved in perceiving prosodic cues in linguistic context after 10 weeks of intervention, suggesting that prosodic difficulties can be treated and hence assessment of prosodic difficulties is worthwhile.
The long-held view that music and language share acoustic and evolutionary commonalities (Coutinho & Dibben, 2012; Fitch, 2006; Juslin & Laukka, 2003; Masataka, 2009) has led to an ongoing enquiry about possible links between music and prosody. Studies have investigated the influence of musical training on sensitivity to emotional prosody. Findings in this area had been contradictory. For example, Thompson, Schellenberg, and Husain (2004) reported that musically trained adults performed significantly better than their untrained counterparts on emotion decoding tasks conveyed by tone sequences and spoken utterances. In a similar study, Trimmer and Cuddy (2008) investigated whether musical training was associated with greater sensitivity to emotional prosody in speech. University undergraduates ($N = 100$) were asked to identify emotion (anger, fear, joy, or sadness) conveyed in both semantically neutral utterances and melodic analogues. Participants also completed a questionnaire about music education and a battery of tests to assess emotional intelligence, music perception, and memory. They found that emotional intelligence (perceiving emotions, facilitating thought about emotions, understanding emotions, and managing one’s own emotions) was associated with the identification of emotion in speech and melodic analogues, but not music training or music perception abilities.

The current study investigated whether there was a link between musicality and children’s prosody perception. Two musical pitch discrimination subtests of the Montreal Battery of Evaluation of Amusia (MBEA) were used to assess pitch contour and interval perception. Information regarding the musical experience of children was collected from parents/caregivers. Aims of the study were to: 1) compare prosody perception and production abilities of children with hearing loss and children with normal hearing and 2) investigate associations between age, hearing level, and musicality and children’s prosody perception.
Materials and methods

Participants

Participants were a convenience sample of children (N = 32) recruited through advertisement in the community and audiology services, 16 with hearing loss (10 boys, 6 girls, M_{age} = 8.71, SD = 1.35, range = 7;2 - 11;11 years) and 16 age- and gender-matched typically developing children with normal hearing (M_{age} = 8.87, SD = 1.47, range = 7;2 - 12;2 years). Children with normal hearing had bilateral pure tone air conduction thresholds ≤ 15 dB at octave frequencies between 250 and 8000 Hz and normal tympanometric peak pressure (measured using a GSI-39 portable pure tone and tympanometer screener). Children with normal hearing were from English-speaking homes and did not have any known cognitive or other developmental delay as reported by parents. Children with normal hearing were age-matched within ± 6 months of their chronological age of a child with hearing loss.

The hearing loss group included: four children with unilateral hearing loss and 12 children with bilateral hearing loss. These children differed in their use of amplification devices and degree of hearing loss (Table 14). Age at diagnosis of hearing loss ranged from at birth to 48 months (M = 24.9, SD = 14.9). All of the children with hearing loss attended mainstream school, used oral communication, spoke English as their first language, had received speech-language therapy, and had no additional disability (motor and/or visual) as reported by parents. Unaided air conduction thresholds for children with hearing loss were obtained between six months to one year period prior to this study.
<table>
<thead>
<tr>
<th>Age (years)</th>
<th>M/F</th>
<th>Hearing status</th>
<th>Device</th>
<th>Age at diagnosis (months)</th>
<th>Better ear PTA (dB HL)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Right ear</td>
<td>Left ear</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>8;2</td>
<td>M</td>
<td>Normal hearing</td>
<td>Mild SNHL</td>
<td>HA</td>
</tr>
<tr>
<td>2</td>
<td>7;6</td>
<td>M</td>
<td>Normal hearing</td>
<td>Moderate CHL</td>
<td>HA</td>
</tr>
<tr>
<td>3</td>
<td>7;10</td>
<td>M</td>
<td>Normal hearing</td>
<td>Profound HL</td>
<td>HA</td>
</tr>
<tr>
<td>4</td>
<td>8;3</td>
<td>M</td>
<td>Profound HL</td>
<td>Normal hearing</td>
<td>CI</td>
</tr>
<tr>
<td>5</td>
<td>7;2</td>
<td>M</td>
<td>Mild HL</td>
<td>Mild HL</td>
<td>HA</td>
</tr>
<tr>
<td>6</td>
<td>9;4</td>
<td>M</td>
<td>Moderate SNHL</td>
<td>Moderate SNHL</td>
<td>HA</td>
</tr>
<tr>
<td>7</td>
<td>7;6</td>
<td>M</td>
<td>Moderate SNHL</td>
<td>Moderate SNHL</td>
<td>HA</td>
</tr>
<tr>
<td>8</td>
<td>9;4</td>
<td>M</td>
<td>Moderate SNHL</td>
<td>Mild SNHL</td>
<td>HA</td>
</tr>
<tr>
<td>9</td>
<td>10;4</td>
<td>M</td>
<td>Moderately severe HL</td>
<td>Severe SNHL</td>
<td>HA</td>
</tr>
<tr>
<td>10</td>
<td>11;11</td>
<td>F</td>
<td>Severe SNHL</td>
<td>Moderately severe SNHL</td>
<td>HA</td>
</tr>
<tr>
<td>11</td>
<td>7;4</td>
<td>F</td>
<td>Profound HL</td>
<td>Profound HL</td>
<td>CI</td>
</tr>
<tr>
<td>12</td>
<td>11;3</td>
<td>F</td>
<td>Profound HL</td>
<td>Profound HL</td>
<td>CI</td>
</tr>
<tr>
<td>13</td>
<td>9;3</td>
<td>F</td>
<td>Profound HL</td>
<td>Profound HL</td>
<td>CI</td>
</tr>
<tr>
<td>14</td>
<td>9;1</td>
<td>F</td>
<td>Profound HL</td>
<td>Profound HL</td>
<td>CI</td>
</tr>
<tr>
<td>15</td>
<td>8;3</td>
<td>F</td>
<td>Profound HL</td>
<td>Profound HL</td>
<td>-</td>
</tr>
<tr>
<td>16</td>
<td>7;7</td>
<td>M</td>
<td>Mild HL</td>
<td>Mild HL</td>
<td>-</td>
</tr>
</tbody>
</table>
Test materials and procedure
Six receptive subtests of the PEPS-C developed by Peppé and colleagues (Wells et al., 2004) and the Child Paralanguage subtest of the DANVA 2 (Nowicki & Duke, 1994) were used in the current study to evaluate receptive prosody in children. Test-retest reliability and internal consistency information are available for these assessments (Chapter 2).

Participants completed the assessment in one session lasting 1 - 2 hours. Informed consent was obtained for all participants and the procedure was approved by the University of Auckland Human Participants Ethics Committee. Six receptive subtests of PEPS-C, Child Paralanguage subtest of the DANVA 2, and Contour and Interval subtests of the MBEA were used (Table 15). Test items were presented using a laptop computer through a GENELEC 6010A active portable loudspeaker at a comfortable level in the normal conversational range (65 - 75 dB SPL) measured using a sound level meter at the position of the participant’s seat. For children with hearing loss using hearing instruments (n = 15), all tests were completed with hearing devices set at their everyday listening setting. Testing was performed in a quiet room. Order of presentation of tasks was randomised across participants.
<table>
<thead>
<tr>
<th>Test materials</th>
<th>Subtest</th>
<th>Task</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>PEPS-C</td>
<td>Short Item Discrimination</td>
<td>Indicating whether two short tones sounded “same” or “different”</td>
<td>Each subtest consists of 16 test items and two practice items. Uses a two alternative forced choice paradigm and hence the pass criterion was set at 75% inorder to avoid chance performance. No feedback was given during the test.</td>
</tr>
<tr>
<td></td>
<td>Long Item Discrimination</td>
<td>Indicating whether two long tones sounded “same” or “different”</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Turn-end</td>
<td>Indicating whether an utterance is a question or statement</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Affect</td>
<td>Indicating liking or disliking with respect to food items</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Chunking</td>
<td>Identifying prosodic phrase boundaries, as in difference between “fruit-salad and milk” versus “fruit, salad, and milk”</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Contrastive Stress</td>
<td>Indicating emphasis on a particular word in a sentence, as in difference between “GREEN and blue socks” versus “green and BLUE socks”</td>
<td></td>
</tr>
<tr>
<td>DANVA 2</td>
<td>Child Paralanguage</td>
<td>The sentence “I am going out of the room now but I will be back later” and participants decide to choose one of the four emotions that best suits the voice.</td>
<td>24 recorded repetitions of the same neutral sentence “I am going out of the room now, but I will be back later” is presented at four emotional states (happy, sad, angry, and fearful) at either high or low emotion intensity. A score of six or below out of 24 items was considered chance performance.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Print out of the emotional smiley faces expressing the four emotions on a piece of paper were placed in front of the participants inorder to assist children remember the response options.</td>
<td></td>
</tr>
<tr>
<td>MBEA</td>
<td>Contour&amp; Interval</td>
<td>Each subtest consists of 31 trials (including one catch trial).</td>
<td>Contour subtest- consists of contour-violated melodies created by changing one note of the melody such that its pitch height relative to its neighbouring notes is reversed, but the key is not altered. Interval subtest- consists of interval violated melodies that are created by modifying the pitch distance between two notes, while maintaining the original contour and key.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>The participant hears the two melodies and indicates on an answer sheet if they are same or different. Correct score of 15 or less was considered as chance performance.</td>
<td></td>
</tr>
</tbody>
</table>
**MBEA and musical experience**

The MBEA was developed by Peretz, Champod, and Hyde (2003) for diagnosing amusia in adults. Peretz et al. (2013) reported that the MBEA was not useful to assess children younger than 10 years of age as the test length is excessive for children. The current study was conducted before the MBEMA (Montreal Battery of Evaluation of Musical Abilities - child version of MBEA, Peretz et al., 2013) was developed with reduced length of the melodies (from 10 notes to 7) and fewer test items (from 30 to 20). The current study used the Contour and Interval subtests of the MBEA to assess musical pitch discrimination abilities in children. The children took 20 minutes to complete the tasks.

A history regarding any formal musical training was obtained from parents/caregivers. Musical experience was coded as a dichotomous variable; “yes” refers to more than one year of formal music training (piano, guitar, or singing) and “no” refers to no formal music training. For the purpose of statistical analyses, music scores and experience were combined and referred to as musicality. Categories of musicality included: 0 = scores below or at chance on both MBEA tasks and no musical experience, 1 = scores above chance on either of the MBEA task and no musical experience, 2 = scores above chance on both MBEA tasks and has musical experience.

**Perceptual rating of prosody production**

Children read a passage from the Wheldall Assessment of Reading Passages (Madelaine & Wheldall, 2002; developed for use with 2nd - 6th grade children). The passage about a horse and a girl has some humour and surprise elements and hence would typically be read with varied intonation and pitch. The reading samples of children with hearing loss \( (n = 15) \) and controls \( (n = 15) \) were rated by nine experienced listeners (speech-language therapists and audiologists). Raters were blinded to the group assignment. Ratings were made using three perceptual prosody rating scales (adapted from Nadig & Shaw, 2012; Table 16). Children’s age and gender information were provided to the raters who were instructed to use their first
impression to rate each sample. Pitch and pitch variations were rated on a 7-point scale (1 = low, 4 = normal, 7 = high) and the overall impression of prosody production was rated using a 4-point scale (1 = atypical, 4 = normal). Raters listened to the 30 recordings in a fixed random order via headphones connected to a laptop at their comfortable loudness level. A practice trial was presented to familiarise them with the procedure. Each recording (one minute in duration) was played only once. PEPS-C expressive subtests were not used in this study as Wells et al. (2004) reported poor inter-rater reliability for PEPS-C expressive tasks ranging from 1.4% to 14.9%. They suggested that this poor reliability was caused by the complex scoring procedure for these tasks.

**Table 16.** Perceptual prosody rating scale

<table>
<thead>
<tr>
<th>1. Pitch</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td></td>
<td></td>
<td></td>
<td>Normal</td>
<td></td>
<td></td>
<td>High</td>
</tr>
<tr>
<td>2. Pitch variations</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>Low</td>
<td></td>
<td></td>
<td></td>
<td>Normal</td>
<td></td>
<td></td>
<td>High</td>
</tr>
<tr>
<td>3. Overall impression</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Atypical</td>
<td></td>
<td></td>
<td></td>
<td>Normal</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Results**

**Prosody measures**

The automatic scoring provided mean percent correct scores for PEPS-C subtests and raw correct scores for DANVA 2 subtest. PEPS-C total scores were calculated as the average of the scores from the four prosody subtests. DANVA 2 total refers to the total raw correct scores obtained across four emotions at high and low emotion intensities.

**Music scores and experience**

All children with musical experience (n = 4 children with hearing loss, n = 7 children with normal hearing) scored above chance (> 18 out of 30 correct) on both MBEA tasks. Four children with hearing loss (n = 2 bilateral hearing loss, n = 2 unilateral hearing loss) who scored above chance on both MBEA tasks used hearing aids and had mild to moderate hearing loss. Six children with no musical experience (n = 2 children with hearing loss, n = 4 children
with normal hearing) scored above chance on one of the MBEA tasks. Fifteen children with no musical experience \((n = 10 \text{ children with hearing loss}, n = 5 \text{ children with normal hearing})\) performed at or below chance on both MBEA tasks. None of the children without musical experience scored above chance on both MBEA tasks.

Between-group differences on PEPS-C and DANVA 2 Child Paralanguage tests

Table 17 shows means and standard deviations for the two groups on PEPS-C and DANVA 2 subtests. The two groups were considered as matched pairs and hence Wilcoxon signed-rank tests were performed to check between-group differences on PEPS-C and DANVA 2 tests. Wilcoxon signed-rank tests showed differences in performance between children with hearing loss and controls for some prosody perception measures (Table 17). As Wilcoxon tests are low powered tests (LaVange & Koch, 2006), Bonferroni correction was not used for multiple comparisons, instead a conservative significance value of \(p < .01\) was used. Scores on all PEPS-C subtests and the total score were significantly lower for children with hearing loss compared to controls \((p < .01)\), except on Long Item Discrimination, Chunking, and Contrastive Stress Reception (Figure 4). DANVA 2 group differences depended on emotional categories (Figure 5) and level of emotion intensity (Figure 6). Scores obtained by children with hearing loss on happy, sad, and fearful emotions were significantly poorer than for controls (Table 17). The hearing loss group performed more poorly overall than the control group for DANVA 2 low \((V = 14.00, p = .005)\) but not for high \((V = 17.50, p = .015)\) emotion intensity items (Figure 6).
Table 17. Mean (SD) and Wilcoxon signed-rank test statistic for PEPS-C and DANVA 2 tests in control and hearing loss groups

<table>
<thead>
<tr>
<th>Tests</th>
<th>Control group</th>
<th>Hearing loss group</th>
<th>V statistic</th>
<th>p</th>
<th>Cohen’s d</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>PEPS-C</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Short Item Discrimination</td>
<td>97.68 (6.38)</td>
<td>87.56 (14.01)</td>
<td>0.00</td>
<td>.005*</td>
<td>1.58</td>
</tr>
<tr>
<td>Long Item Discrimination</td>
<td>89.37 (7.02)</td>
<td>80.68 (15.19)</td>
<td>16.5</td>
<td>.044</td>
<td>1.23</td>
</tr>
<tr>
<td>Turn-end</td>
<td>95.06 (6.48)</td>
<td>86.87 (14.75)</td>
<td>18.0</td>
<td>.010*</td>
<td>1.26</td>
</tr>
<tr>
<td>Affect</td>
<td>89.31 (6.69)</td>
<td>74.06 (13.83)</td>
<td>6.50</td>
<td>.006*</td>
<td>2.27</td>
</tr>
<tr>
<td>Chunking</td>
<td>93.87 (7.21)</td>
<td>86.12 (10.34)</td>
<td>10.00</td>
<td>.024</td>
<td>1.07</td>
</tr>
<tr>
<td>Contrastive Stress</td>
<td>90.81 (8.51)</td>
<td>82.68 (12.89)</td>
<td>5.00</td>
<td>.043</td>
<td>0.95</td>
</tr>
<tr>
<td>Total</td>
<td>92.26 (4.15)</td>
<td>82.43 (8.49)</td>
<td>4.00</td>
<td>.001*</td>
<td>2.36</td>
</tr>
<tr>
<td><strong>DANVA 2</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Happy</td>
<td>4.87 (0.80)</td>
<td>4.00 (1.41)</td>
<td>14.50</td>
<td>.010*</td>
<td>1.08</td>
</tr>
<tr>
<td>Sad</td>
<td>4.93 (0.85)</td>
<td>4.00 (1.89)</td>
<td>8.00</td>
<td>.050</td>
<td>1.09</td>
</tr>
<tr>
<td>Angry</td>
<td>5.31 (1.01)</td>
<td>5.06 (1.06)</td>
<td>18.0</td>
<td>.624</td>
<td>0.24</td>
</tr>
<tr>
<td>Fearful</td>
<td>4.81 (1.10)</td>
<td>3.06 (1.87)</td>
<td>7.50</td>
<td>.008*</td>
<td>1.59</td>
</tr>
<tr>
<td>High Intensity</td>
<td>10.75 (0.93)</td>
<td>9.12 (2.06)</td>
<td>17.50</td>
<td>.015</td>
<td>1.75</td>
</tr>
<tr>
<td>Low Intensity</td>
<td>9.18 (1.37)</td>
<td>7.00 (2.19)</td>
<td>14.00</td>
<td>.005*</td>
<td>1.59</td>
</tr>
<tr>
<td>Total</td>
<td>19.93 (1.73)</td>
<td>16.12 (3.89)</td>
<td>7.00</td>
<td>.004*</td>
<td>2.20</td>
</tr>
</tbody>
</table>

Figure 4. Mean percent correct scores and standard deviations on PEPS-C subtests for control vs. hearing loss groups; SID=Short Item Discrimination, LID=Long Item Discrimination.
Figure 5. Mean scores on four emotion categories using DANVA 2 Child Paralanguage subtest for control vs. hearing loss groups.

Figure 6. Mean scores on low and high emotion intensity stimuli of the DANVA 2 Child Paralanguage subtest for control vs. hearing loss groups.

Association between prosody perception, age, hearing level, and musicality

Table 18 shows Spearman correlations between age, hearing level (better ear PTA), musicality and prosody perception measures. Nonparametric tests were used for correlations involving ordinal data (musicality and prosody production). PEPS-C and DANVA 2 total scores were significantly correlated with age, hearing level, and musicality (Table 18). There was a significant correlation between DANVA 2 low emotion intensity scores and musicality. This
significant correlation between DANVA 2 low intensity scores and musicality ($r_s = .507, p = .004$) suggests that musical skills may help children to recognise vocal emotions presented with subtle emotional cues. No significant correlation was found between DANVA 2 high intensity scores and musicality ($r_s = .114, p = .540$).

**Table 18.** Correlations between age, hearing, musicality, and prosody perception ($N = 32$)

<table>
<thead>
<tr>
<th>Test</th>
<th>Age</th>
<th>Hearing level</th>
<th>Musicality</th>
</tr>
</thead>
<tbody>
<tr>
<td>Turn-end</td>
<td>.276</td>
<td>-.191</td>
<td>.298</td>
</tr>
<tr>
<td>Affect</td>
<td>.269</td>
<td>-.527**</td>
<td>.231</td>
</tr>
<tr>
<td>Chunking</td>
<td>.361*</td>
<td>-.411*</td>
<td>.379*</td>
</tr>
<tr>
<td>Contrastive Stress</td>
<td>.313</td>
<td>-.215</td>
<td>.386*</td>
</tr>
<tr>
<td>PEPS-C total</td>
<td>.406*</td>
<td>-.546**</td>
<td>.473**</td>
</tr>
<tr>
<td>Happy</td>
<td>.278</td>
<td>-.424*</td>
<td>.359*</td>
</tr>
<tr>
<td>Sad</td>
<td>.524**</td>
<td>-.030</td>
<td>-.150</td>
</tr>
<tr>
<td>Angry</td>
<td>-.088</td>
<td>-.170</td>
<td>.269</td>
</tr>
<tr>
<td>Fearful</td>
<td>.152</td>
<td>-.599**</td>
<td>.441*</td>
</tr>
<tr>
<td>Low intensity</td>
<td>.305</td>
<td>-.482**</td>
<td>.507**</td>
</tr>
<tr>
<td>High intensity</td>
<td>.407*</td>
<td>-.399*</td>
<td>.110</td>
</tr>
<tr>
<td>DANVA 2 total</td>
<td>.375*</td>
<td>-.504**</td>
<td>.423*</td>
</tr>
</tbody>
</table>

*Note.** $**p < .01; *p < .05; only PEPS-C receptive prosody subtests are included in the table.*

Stepwise regression analyses were performed using IBM SPSS Statistics 21 to examine associations between age, hearing level, musicality and prosody measures (PEPS-C and DANVA 2 total). Hearing level alone explained 29.5% and 37.2% of the variance in PEPS-C and DANVA 2 total scores, respectively. Age and hearing level together accounted for 55.4% and 56.7% of the variance in PEPS-C and DANVA 2 total scores, respectively (Table 19).
Table 19. Stepwise regression analyses of prosody perception measures as a function of age (in years) and hearing level (in dB HL)

<table>
<thead>
<tr>
<th></th>
<th>PEPS-C Total</th>
<th>DANVA 2 Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>b</td>
<td>SE</td>
</tr>
<tr>
<td>Constant</td>
<td>64.346</td>
<td>6.735</td>
</tr>
<tr>
<td>Age</td>
<td>3.149</td>
<td>.767</td>
</tr>
<tr>
<td>Hearing level</td>
<td>-.157</td>
<td>.031</td>
</tr>
</tbody>
</table>

Note. The dependent variables were PEPS-C total ($R^2 = .554$) and DANVA 2 total scores ($R^2 = .567$); $p < .01$.

Although the Spearman analysis showed a significant correlation between prosody scores and musicality, this was excluded in the regression model ($p > .05$). This lack of statistical significance suggests colinearity between the predictor variables (e.g. hearing loss and musicality). There was no significant association between age and musicality ($r_s = .043, p = .814$). Spearman tests showed a negative correlation between hearing level and musicality ($r_s = -.354, p = .047$). That is, poorer hearing was associated with less musical experience and lower MBEA scores. Four children with hearing loss who had musical experience and good MBEA scores had mild to moderate hearing loss. None of the children with severe to profound hearing loss had taken music lessons and all of them performed at chance on both MBEA tasks.

Unilateral hearing loss

As the regression analyses used the hearing level in the better ear, children with unilateral hearing loss were treated in the model in the same way as children with normal hearing. Table 20 shows the PEPS-C and DANVA 2 total scores obtained by the four children with unilateral hearing loss. Although these children have normal hearing thresholds in their better ear, all of them performed more poorly (range 72.25 - 91.00) than the control group average based on their PEPS-C total scores (control group $M = 92.26$, $SD = 4.15$) (Table 17). Two children with moderate and profound unilateral hearing loss had poorer scores than the mean DANVA 2 total score obtained by the control group ($M = 19.97$, $SD = 1.73$). Thus, for the small number
of participants investigated here, unilateral hearing loss appeared to be associated with poorer
prosody perception.

Table 20. PEPS-C total and DANVA 2 total scores of children with unilateral hearing loss

<table>
<thead>
<tr>
<th>Age</th>
<th>PTA (poorer ear) in dB HL</th>
<th>PEPS-C total</th>
<th>DANVA 2 total</th>
<th>Aetiology</th>
<th>Device used</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>7;6</td>
<td>50</td>
<td>91</td>
<td>18</td>
<td>Otitis Media</td>
</tr>
<tr>
<td>2</td>
<td>7;10</td>
<td>93.33</td>
<td>89.25</td>
<td>18</td>
<td>Meningitis</td>
</tr>
<tr>
<td>3</td>
<td>8;2</td>
<td>35</td>
<td>72.25</td>
<td>20</td>
<td>Unknown</td>
</tr>
<tr>
<td>4</td>
<td>8;3</td>
<td>78.33</td>
<td>86</td>
<td>20</td>
<td>Unknown</td>
</tr>
</tbody>
</table>

Prosody production

Inter-rater agreement computed using intra-class correlation coefficients showed significant
moderate level agreement between the nine raters for the pitch, pitch variation, and overall
impression rating scales ($p = .001$; Table 21). The hearing loss group had greater variability
for all ratings compared to the control group (Figure 7). On average, children with hearing loss
($M = 2.94$, $SD = .702$) were rated as having more atypical prosody overall than the control
group ($M = 3.14$, $SD = .495$), but this difference was not significant ($p > .05$). Repeated
measures ANOVA with pitch and pitch variations as within-subject variables and group as a
between-subject variable showed a significant difference between ratings of pitch and pitch
variation ($F = 56.761$, $p = .001$; with Greenhouse-Geisser correction for degrees of freedom)
and a trend for higher pitch and greater pitch variation in the hearing loss group ($F = 3.389$, $p$
$= .076$). Spearman correlations between ratings of prosody production and prosody perception
(PEPS-C or DANVA 2 total scores) were not significant ($p > .05$).

Table 21. Inter-rater agreement between raters on perceptual prosody rating data

<table>
<thead>
<tr>
<th>Prosody feature</th>
<th>% agreement</th>
<th>$p$</th>
<th>Lower 95% CI</th>
<th>Upper 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pitch</td>
<td>86.9</td>
<td>.001</td>
<td>.786</td>
<td>.929</td>
</tr>
<tr>
<td>Pitch changes</td>
<td>75.1</td>
<td>.001</td>
<td>.597</td>
<td>.864</td>
</tr>
<tr>
<td>Overall</td>
<td>83.9</td>
<td>.001</td>
<td>.728</td>
<td>.914</td>
</tr>
</tbody>
</table>
Prosody perception and production in children with hearing loss

Children with hearing loss performed significantly more poorly than their age- and gender-matched control peers on PEPS-C and DANVA 2 tests, despite having received early diagnosis and speech-language therapy services. This suggests that children with a range of hearing losses using different types of hearing technology are not able to accurately detect subtle variations in acoustic cues that are necessary for adequate perception of prosodic features in speech. These results are consistent with previous findings that children with hearing loss are poorer than controls in perceiving affective prosody (Hopyan-Misakyan et al., 2009; Nakata et al., 2012), question-statement contrasts (Peng et al., 2008), and lexical stress (Most & Peled, 2007). As with any aspect of speech and language, early intervention could be beneficial to children with prosodic difficulties. The PEPS-C test is useful for evaluating specific strengths and weaknesses on different aspects of prosody (linguistic and emotional) in
children and could be useful for developing individualised intervention plans. Different methods for treating prosodic difficulties have been described in the literature including Dynamic temporal and tactile cueing for speech motor learning (DTTC, Strand, Stoeckel, & Bass, 2006), Melodic intonation therapy (MIT, Sparks & Holland, 1976), and Lee Silverman voice treatment (LSVT®, Ramig, Countryman, Thompson, & Horii, 1995). The Prosody Treatment Program (Rothstein, 2013) is a newly developed tool that provides exercises to train clients in receptive and expressive prosodic skills across various disorders (Odell & Shriberg, 2001; Peppé, 2011; Hargrove, Anderson, & Jones, 2009). This software package incorporates the principles of empirically validated prosody treatment strategies (such as DTTC, MIT and LVST) and is also based on expert professional practice. There is shortage of empirical research into the efficacy of such treatments targeting prosody in people with hearing loss (Hargrove et al., 2009; Peppé, 2009).

The perceptual rating experiment revealed a trend for children with hearing loss to be rated as having more atypical prosody than controls. This is consistent with Peng et al.’s (2008) finding that children with CIs scored significantly lower than controls when their prosody production (question vs. statement contrasts) was rated for utterance type accuracy and contour appropriateness. Previous studies using narratives, reading, picture description, and imitation tasks (Diehl et al., 2009; Grossman, Edelson, & Tager-Flusberg, 2013; Sharda et al., 2010) have shown greater pitch variations in children with HFA compared to typically developing peers, but results vary for narrative retelling and spontaneous speech tasks. Diehl et al. (2009) found no difference in mean pitch production between children with HFA and typically developing peers using a narrative retelling task, whereas Sharda et al. (2010) found significantly higher mean pitch in their HFA group compared to controls using a spontaneous speech sample. Prosody samples in the current study were elicited using a passage reading task of one minute in duration. Further studies of prosody production using a spontaneous speech sample (as used by Sharda et al., 2010) should examine whether differences in prosody
production are evident between children with hearing loss and controls. Another suggestion is to conduct more objective and precise acoustic analyses.

Effects of age and hearing

Regression analyses showed that increasing age and better hearing were associated with better prosody perception abilities. This result is consistent with studies showing that prosody perception improves in the early school years and reaches a plateau at about 9 years (Gibbon & Smyth, 2013; Wells et al., 2004). Better PEPS-C and DANVA 2 total scores were associated with better hearing levels. This is consistent with Marx et al.’s (2015) finding that CI users with thresholds better than 60 dB HL in the low frequencies had near-normal question/statement discrimination abilities. Poor perception of affective prosody and music in people with hearing loss is often attributed to inadequate access to pitch cues (Kong, Cruz, Jones, & Zeng, 2004; Looi, McDermott, McKay, & Hickson, 2008; McDermott 2004). Cullington and Zeng (2010) reported that bimodal stimulation allows for better acoustic transmission of pitch cues compared to bilateral CIs. This suggests that combined use of hearing aids and CIs may improve prosody perception. Straatman, Rietveld, Beijen, Mylanus, and Mens (2010) compared question/statement discrimination performance in 17 children with CIs when the children were wearing or were not wearing their contralateral hearing aid. They reported that using residual hearing, correct recognition improved by 6% to 11%. The current study included three children with bilateral hearing loss using unilateral CIs. Bimodal stimulation could be beneficial for music and prosody perception in these children.

Musical training and prosodic sensitivity

Long-term musical training modulates representation in the human auditory cortex as revealed in auditory evoked potentials (AEP) and magnetoencephalography (MEG) (Pantev et al., 1998; Pantev, Roberts, Schulz, Engelien & Ross, 2001). Recent studies have demonstrated functional brain reorganisation even after short periods of musical training (Gaab et al., 2006; Stewart et al., 2003), although the extent to which these changes endure after training is unknown. Several studies have also reported structural (cortical and subcortical) brain
plasticity induced by musical practice in people with normal hearing (Draganski & May, 2008; Hyde et al., 2009; Bermudez & Zatorre, 2005). Fauvel, Groussard, Eustache, Desgranges, and Platel (2013) reported that as musical learning puts high demand on pre-existing skills and cognitive functions resulted in transfer of improvements from one activity (music making) to other (e.g. language skills, executive functions). Musical training has been associated with cognitive enhancement (Schellenberg, 2004; 2005), improved working memory (Forgeard, Winner, Norton & Schlaug, 2008), attention (Strait, Kraus, Parberry-Clark & Ashley, 2010), and executive function skills (Bialystok & DePape, 2009). However, a limited number of studies have investigated the influence of musical training on prosody perception. In particular, this has not considered children with hearing loss.

Thompson et al. (2004) assessed musically trained and untrained adults on their ability to detect emotions (happy, sad, angry, or fearful) conveyed by tone sequences that mimicked pitch and temporal structure of spoken phrases. Musically trained adults performed significantly better than the untrained counterparts. Thompson et al. (2004) also investigated judgement of emotional meaning of tone sequences and spoken utterances in musically trained and untrained adults. They reported enhanced decoding accuracy among trained adults compared to the untrained group for sad and fearful emotions, but not for happy and angry emotions. Interestingly, the pattern of decoding accuracy reported in this study was similar to the normative data for adults reported by Juslin and Laukka (2003). Generally, happy and angry emotions are easier to decode compared to sad and fearful as they are more acoustically salient. In the current study children with hearing loss were poorer at recognising happy, sad, and fearful emotions on the DANVA 2 compared to controls, but only for the more difficult task with a low level of emotional intensity. Limo and Castro (2011) found that musical expertise enhanced the recognition of emotion in speech. Short sentences expressing anger, disgust, fear, happiness, sadness, and surprise emotions were presented in a forced-choice identification task. Forty highly trained musicians (at least eight years of musical training and regular practice) had significantly faster reaction times and enhanced recognition of emotional
prosody compared to controls. Children in the current study had an average of 2.5 years of formal musical training. Although there was a moderate correlation between musicality and children’s prosody perception, this needs to be investigated further with a larger sample, and using a prospective study design.

Conclusions

Children with hearing loss have marked difficulty in perceiving different aspects of prosody compared to their typically developing peers. There was a trend for the hearing loss group to be rated as having more atypical prosody compared to controls. These findings suggest that prosodic skills should be assessed routinely in children with hearing loss. Prosody treatments such as the Prosody Treatment Program (Rothstein, 2013) require further research. A few programmes are available for the treatment of impaired prosody production; however there is a lack of empirical research into the efficacy of treatments targeting prosody perception. Future research is needed to establish the contribution of musical training to prosody perception in children with hearing loss. Aspects of prosody other than emotion such as understanding sentence type (question/statement/command), idioms, sarcasm, word boundaries, and emphasis/prominence in a sentence require investigation. Auditory habilitation programmes for children with hearing loss should include development of age-appropriate prosodic skills sufficient for them to function without accommodations at school and in the wider community (Mecklenburg, et al., 1990). Given the importance of prosody in successful communication, prosodic deficits in children with hearing loss should be investigated and treated in clinical practice.
Chapter 5
Prosody perception and musical pitch discrimination in adults using cochlear implants

Introduction

Electrical stimulation of the auditory nerve using CIs can restore hearing in severe to profound sensorineural hearing loss (SNHL). Previous studies have examined different aspects of speech perception in CI users, in particular consonant and vowel recognition, speech in noise perception, and localization and lateralization abilities (Davidson, Geers, & Brenner, 2010; van Hoesel & Tyler, 2003; Wolfe et al., 2009), but there are only few studies in the area of prosody (Hopyan-Misakyan et al., 2009; Kong, Stickney, & Zeng, 2005; Meister et al., 2009; Most, Harel, Shpak, & Luntz, 2011; Most & Peled, 2007; Nakata et al., 2012; Peng et al., 2008; Saindon, 2010; Straatman et al., 2010; Torppa et al., 2010, 2014). Advances in CI technology and speech processing strategies have significantly improved CI users’ ability to recognise speech sounds (Wilson, 2000; Vandali et al., 2000), however no studies have documented advances in prosody perception with changes in CI technology (Nakata et al., 2012).

Prosody has a role in indexical, grammatical, emotional, and pragmatic levels of communication, conveying a speaker’s emotional state (e.g. happy vs. sad), gender and identity, information on sentence type (question vs. statement), and word boundaries within phrases (Wells & Peppe, 2003). Auditory cues responsible for prosody perception include fundamental frequency (F0) changes during an utterance (average, range, shape, and direction of pitch contour), the syllable duration or the rate of speech production, and the average intensity and intensity changes along the utterance (Chun, 2002). The extent to which different acoustic cues dominate depends on the kind of prosodic contrast. Coutinho and Dibben (2012) reported that loudness, tempo and speech rate, melodic and intonation contour, spectral centroid, and sharpness are the psychoacoustic cues important for making judgements of
emotion in speech. Timing and pitch cues are used to convey turn-taking during conversation (Hirschberg, 2002).

Accurate recognition of prosodic patterns depends on the ability to resolve subtle changes in frequency and amplitude over time. A normal functioning cochlea is capable of accurately conveying these auditory cues to listeners, but this is not the case with individuals with SNHL who have impaired frequency and temporal resolution (Moore & Carlyon, 2005). Moore and Carlyon (2005) reported that as a consequence of poor frequency resolution and reduced ability to use temporal fine structure information, listeners with hearing loss (HL) are more reliant on temporal envelope cues for pitch information than listeners with normal hearing (NH).

F0 is a major factor in prosody perception, and it is assumed that the difficulties CI users have perceiving pitch will be associated with difficulties in the perception of prosody. Temporal pitch and place pitch are two basic mechanisms for pitch perception in CIs. Temporal pitch is related to the rate of stimulation of electrodes and place pitch is related to the location of electrode stimulation along the cochlea (Kong & Carlyon, 2010). The number of electrodes provided in CI systems is small compared to estimates of the number of channels (tonotopic organization) in the normal auditory system (Moore, 2003). The finite number of electrodes inserted in the cochlea (12 to 22 electrodes) covering a frequency range up to 5000-10,000 Hz and the limited temporal information available from band-specific envelope cues restricts pitch perception in CI users compared to NH listeners (Wilson et al., 1991; Vandali et al., 2000).

The cochlear mechanics and hair cell transduction involved in acoustic stimulation allow neural responses to show phase locking for frequencies up to 2 kHz (Johnson, 1980). Auditory nerve responses to electrical stimulation with CIs lack the stochastic activity or travelling-wave delay across nerve fibers that occur with NH (Carlyon, Deeks, & McKay, 2010). Moreover, modern CIs primarily use a monopolar electrode configuration which produces a relatively broad electrical field that reduces place specificity (Kral, Hartmann,
Mortazavi, & Klinke, 1998). Not surprisingly, previous studies suggest that place and temporal pitch coding mechanisms in CIs do not yield the same pitch perception experienced by NH listeners (Carlyon, van Wieringen, Long, Deeks, & Wouters, 2002; Chen & Zeng, 2004; Zeng, 2002).

Current speech processing strategies such as Advanced Combination Encoder (ACE), Continuous Interleaved Sampling (CIS), and Spectral Peak (SPEAK) use a bank of bandpass filters to filter the signal waveforms and extract envelope information from these filtered waveforms using full-wave rectification and low-pass filtering (typically with 200-400 Hz cutoff frequencies). The compressed envelope information is then used to modulate biphasic pulses. Biphasic pulses with amplitudes proportional to the envelope are delivered to the CI electrodes at a constant rate in a non-overlapping fashion. Pulse rate, pulse duration, stimulation order, and compression can be varied in CIS processors to optimize speech recognition. Perception of F0 in CIS-like strategies is related to envelope temporal fluctuations in each band (Geurts & Wouters, 2004). However, pitch perception using this temporal coding saturates at approximately 300 Hz (Green, Faulkner, Rosen, & Macherey, 2005; Oxenham, 2008) and this imposes a limitation on prosody perception. In contrast to CIS strategies, which only extract envelope information, more recent fine structure processing (FSP) strategies such as FS4 and FS4p used in MED-EL CIs are designed to better represent signal fine structure (Arnoldner et al., 2007). This is intended to provide better frequency coding in the low to mid frequencies by allowing auditory nerve fibres to better phase-lock to the sound signal, as is the case in NH in this frequency region. In addition, the lower border frequency is decreased to 70 Hz compared to 250 Hz for the CIS strategy. This improved access to the fine structure of low frequency sounds has benefits for speech perception in noise and music appreciation.

Pitch, melody, rhythm, and timbre are four basic elements of music (Krumhansl & Iverson, 1992) and information related to F0 (of individual notes), duration, and the temporal fine structure are cues for music perception (Cooper, Tobey, & Loizou, 2008). CI users are reported to have poor perception of music due to inadequate access to these cues (Heng,
Cantarero, Elhilali, & Limb, 2011; Looi et al., 2008). Strategies like F0/F1/F2 and MPEAK (not in current use) convert an estimate of F0 directly into the stimulation rate, so prominent cues like F0 and first two formants are coded; CIS-like strategies only retain the temporal envelope information while the temporal fine structure information is eliminated.

There is evidence that temporal fine structure information is more important for musical pitch (Smith, Delgutte, & Oxenham, 2002) and timbre perception (Heng et al., 2011) than envelope cues. The feature-extraction strategies used in early CIs were not able to process multiple pitches from simultaneously presented sound sources, as would occur in music. No significant association between music perception scores and speech processing strategy has been reported (ACE, CIS, SPEAK, MPEAK) (McKay & McDermott, 1993; Gfeller et al., 2002a, Leal et al., 2003). CI users perform significantly worse than NH listeners on music perception tasks such as melodic contour and musical pitch discrimination (Gfeller, Woodworth, Robin, Witt, & Knutson, 1997; Hopyan, Peretz, Chan, Papsin, & Gordon, 2012). Adult CI users are significantly less accurate than NH adults on perception tasks associated with processing of spectral information, including timbre recognition (musical instrument recognition), timbre appraisal (sound quality rating) (Gfeller et al., 2002b; Looi, 2008; McDermott, 2004), frequency discrimination, pitch ranking, and discrimination of brief pitch patterns (Kong et al., 2004, 2005; Laneau, Wouters, & Moonen, 2004; Looi, 2008).

Most studies of pitch perception in CI listeners use single sources of sound. Little research is available to indicate how CI users process pitch given the presence of multiple pitch-carrying sounds (e.g. music). Melodies and harmonies are comprised of sequential and concurrently presented pitches respectively; poor transmission of pitch results in difficulties perceiving melodies with or without harmony (Galvin, Fu, & Nogaki, 2007; Kong et al., 2004; Leal et al., 2003; Looi, 2008; McDermott, 2004).

In this paper, we report the prosody perception abilities of adults with CIs evaluated using the PEPS-C test (Peppé, 1998) and the DANVA 2 Adult Paralanguage subtest (Baum & Nowicki, 1998). In addition, pitch contour and interval discrimination for melodic tones was
evaluated using the MBEA test (Peretz et al., 2003). This study also examined the relationship between prosody perception and non-linguistic auditory measures, demographic variables (age, duration of implant use, and duration of musical experience), and speech perception (HINT and CNC tests). Previous results for NH listeners are available for each of the tasks and these were used to compare the CI users’ performance to norms.

Methodology

Participants

Twelve adult CI users (4 males, 8 females) aged 25;5 to 78;0 years (M_age = 58.4, SD = 16.1) with bilateral severe to profound SNHL participated. CI participants’ PEPS-C scores were compared to normative data collected from 15 New Zealand English speaking adults (7 males, 8 females) aged 23;0 to 57;0 years (M_age = 32.0, SD = 10.0) with NH (Gharahdaghi, 2013). Nine out of 12 CI participants had acquired SNHL while three participants had congenital HL (CI-4, CI-7, CI-11). The aetiology of participants’ HL varied, with the most frequent cause cited as “unknown”. CI participants used different Nucleus multichannel CI models (Table 22) and had fully inserted electrodes (information from CI surgical records). Ten participants used the ACE (Vandali et al., 2000) speech processing strategy with a constant stimulation rate of 900 Hz and two participants (CI-8, CI-9) used SPEAK (Skinner et al., 1994). Three participants (CI-10, CI-11, CI-12) used bimodal stimulation (a hearing aid in the contralateral ear) and were tested with both their CI and their hearing aid. The average aided pure tone threshold in the contralateral ear of bimodal users (n = 3) was 90 dB HL. The average unaided pure tone threshold in the contralateral ear of unilateral CI users (n = 9) was 105 dB HL. All CI participants had good aided pure tone thresholds in their implanted ear (≤ 25 dB HL). Most participants had at least 6 months experience with their CI, except CI-12 (4 months). The duration of CI use ranged from 0;4 to 17;4 years (M = 5.4, SD = 5.8). Seven out of 12 CI participants (including two bimodal users; CI-10 and CI-12) in the present study had music lessons or participated in musical activities (mean duration 5.2 years) during their high school
years. With their permission, details regarding duration of CI use, speech perception scores, and CI model were obtained from participants’ files maintained at the University of Auckland Audiology Clinic (Table 22).
Table 22. Participant characteristics

<table>
<thead>
<tr>
<th>Participants</th>
<th>Age (years)</th>
<th>M/F</th>
<th>Device(s)</th>
<th>Implant type</th>
<th>Processing strategy</th>
<th>Implant use (years)</th>
<th>HL onset (years)</th>
<th>Hearing aid use (years)</th>
<th>Aetiology of HL</th>
<th>Musical training (years)</th>
<th>CNC%</th>
<th>HINT%</th>
</tr>
</thead>
<tbody>
<tr>
<td>CI-1</td>
<td>73;4</td>
<td>F</td>
<td>CI</td>
<td>Nucleus 24</td>
<td>ACE</td>
<td>3;4</td>
<td>44</td>
<td>26</td>
<td>Unknown</td>
<td>0</td>
<td>24</td>
<td>82</td>
</tr>
<tr>
<td>CI-2</td>
<td>69;9</td>
<td>M</td>
<td>CI</td>
<td>Nucleus24</td>
<td>ACE</td>
<td>2;2</td>
<td>37</td>
<td>42</td>
<td>Unknown</td>
<td>5</td>
<td>76</td>
<td>100</td>
</tr>
<tr>
<td>CI-3</td>
<td>52;9</td>
<td>F</td>
<td>CI</td>
<td>Nucleus24</td>
<td>ACE</td>
<td>4;7</td>
<td>1</td>
<td>46</td>
<td>Unknown</td>
<td>0</td>
<td>68</td>
<td>100</td>
</tr>
<tr>
<td>CI-4</td>
<td>48;1</td>
<td>F</td>
<td>CI</td>
<td>Nucleus422</td>
<td>ACE</td>
<td>1;1</td>
<td>36</td>
<td>10</td>
<td>Unknown</td>
<td>0</td>
<td>92</td>
<td>100</td>
</tr>
<tr>
<td>CI-5</td>
<td>78;0</td>
<td>F</td>
<td>CI</td>
<td>Nucleus512</td>
<td>ACE</td>
<td>3;2</td>
<td>39</td>
<td>32</td>
<td>Unknown</td>
<td>6</td>
<td>60</td>
<td>100</td>
</tr>
<tr>
<td>CI-6</td>
<td>72;2</td>
<td>M</td>
<td>CI</td>
<td>Nucleus512</td>
<td>ACE</td>
<td>2;9</td>
<td>18</td>
<td>32</td>
<td>Unknown</td>
<td>2</td>
<td>48</td>
<td>95</td>
</tr>
<tr>
<td>CI-7</td>
<td>25;5</td>
<td>M</td>
<td>CI</td>
<td>Nucleus24</td>
<td>ACE</td>
<td>4;6</td>
<td>2</td>
<td>18</td>
<td>Unknown</td>
<td>0</td>
<td>60</td>
<td>96</td>
</tr>
<tr>
<td>CI-8</td>
<td>56;3</td>
<td>F</td>
<td>CI</td>
<td>Nucleus22</td>
<td>SPEAK</td>
<td>17;4</td>
<td>15</td>
<td>0</td>
<td>Unknown</td>
<td>3</td>
<td>62</td>
<td>-</td>
</tr>
<tr>
<td>CI-9</td>
<td>66;1</td>
<td>M</td>
<td>CI</td>
<td>Nucleus22</td>
<td>SPEAK</td>
<td>17;4</td>
<td>17</td>
<td>48</td>
<td>Otosclerosis</td>
<td>6</td>
<td>72</td>
<td>92</td>
</tr>
<tr>
<td>CI-10</td>
<td>68;2</td>
<td>F</td>
<td>CI+HA</td>
<td>Nucleus422</td>
<td>ACE</td>
<td>1;5</td>
<td>20</td>
<td>38</td>
<td>Unknown</td>
<td>10</td>
<td>79</td>
<td>96</td>
</tr>
<tr>
<td>CI-11</td>
<td>38;6</td>
<td>F</td>
<td>CI+HA</td>
<td>Nucleus24</td>
<td>ACE</td>
<td>6;2</td>
<td>2;6</td>
<td>36</td>
<td>Hereditary</td>
<td>0</td>
<td>56</td>
<td>89</td>
</tr>
<tr>
<td>CI-12</td>
<td>61;9</td>
<td>F</td>
<td>CI+HA</td>
<td>Nucleus422</td>
<td>ACE</td>
<td>0;4</td>
<td>35</td>
<td>7</td>
<td>Otosclerosis</td>
<td>5</td>
<td>72</td>
<td>90</td>
</tr>
</tbody>
</table>

Note: Performance of CI-1 on CNC test ranged between 44% to 24%, only the recent result is reported here; CNC=Consonant Nucleus Consonant test; HINT=Hearing in Noise Test; CI=Cochlear Implant; CI+HA=Cochlear Implant and Hearing Aid.
Test materials

Profiling Elements of Prosody in Speech-Communication (PEPS-C)

PEPS-C was developed based on a psycholinguistic framework to assess receptive and expressive prosodic abilities using analogous tasks (Péppe, 1998). Four receptive tasks such as Turn-End (interaction), Affect (emotion), Chunking (grammar), and Contrastive Stress (pragmatics) Reception subtests were used. PEPS-C uses the prosodic opposition between high-rise (question) versus low-fall (statement) in the Turn-End Reception subtest. The Affect Reception tasks make use of rise-fall pitch movement to convey strong liking and a fall-rise pitch indicates dislike. Chunking refers to prosodic delimitation of an utterance into units (or intonational phrases), for example /FRUIT-SALAD/and MILK/ versus /FRUIT/SALAD/and MILK/. In the second utterance, there are three intonational phrases, each with its own accent, which makes it an utterance with a list of three food items. In the first utterance, the absence of a separate intonation phrase for SALAD makes FRUIT-SALAD a compound noun. The pragmatic use of prosody to indicate phonetic prominence is assessed in the Contrastive Stress Reception subtest which uses prosodic prominence by making use of falling pitch together with initial high pitch on the main syllable of the item (e.g. /CHOCOLATE AND HONEY/ or /CHOCOLATE AND HONEY/).

The stimulus set includes prerecorded auditory stimuli (e.g., names of food items), and their corresponding pictures. The computer response screen involves a split-screen display of cartoon-type pictures. The software uses a two alternative forced choice method (2AFC). The automatic scoring provides the raw score, percentage scores, standard deviation from the normative mean, a pass/fail indicator, and percentile ranks. The pass criterion was set at 75% to avoid the possibility of chance scoring. Each subtest in PEPS-C includes two example items, two practice items, and 16 test items. The normative data in PEPS-C version 7 is for 5-14 year olds from London (U.K.). The language version used for the current study is British English.
Diagnostic Analysis of Non Verbal Accuracy 2 (DANVA 2)

The DANVA 2 was developed by Baum and Nowicki (1998) to measure competence in affect recognition by reading facial expressions and voice tone (emotion/affect). It includes five subtests: 1) Child Faces, 2) Adult Faces, 3) Child Paralanguage, and 4) Adult Paralanguage, and 5) Child and Adult Posture. The current study used the Adult Paralanguage subtest to assess affect recognition using voice only. This 24-item subtest involves a semantically neutral sentence “I am going out of the room now but I’ll be back later” presented in happy, sad, angry, and fearful tones at two levels of emotional intensity (12 items per intensity level) by male and female speakers (in random sequence). Participants respond to say if the person sounds happy, sad, angry, or fearful. The automatic computer scoring generates tables that show the number of errors for each emotion, number of errors for high and low intensity stimuli, number of errors for emotion by intensity, and the responses that were chosen when there was an error. Error profiles can be used to identify the pattern of difficulty. Additional information about the DANVA 2 test can be found on http://psychology.emory.edu/clinical/interpersonal/

Montreal Battery of Evaluation of Amusia (MBEA)

MBEA was developed by Peretz et al. (2003) as a test for diagnosing amusia. The current study used two of the six MBEA subtests (Contour and Interval) which rely on perception of pitch, because pitch is regarded as one of the primary cues for perception of prosody in spoken English (Chun, 2002). The contour subtest consists of contour-violated melodies created by changing one note of the melody such that its pitch height relative to its neighbouring notes is reversed, but the key is not altered. The interval subtest consists of contour-preserved, interval violated melodies that are created by modifying the musical interval between two notes, while maintaining the original contour and key (Peretz et al., 2003). The average pitch interval change from the original pitch is 4.2 semitones. Each subtest consists of 31 trials (including one catch trial), where a trial comprises two short
musical tones that are either identical or different at a single point (the nature of the difference differs according to the subtest). The participant hears the two melodies and indicates on an answer sheet if they are same or different. Each test is 10 minutes long, preceded by two examples with feedback for each subtest. No feedback is given during the test. Participants are advised that some differences may be difficult to hear. Scoring is done manually and the participants should correctly respond to a catch trial (which has a very obvious difference between stimuli) in order for their responses to be considered valid. All CI participants responded correctly to the catch trial.

Procedure

Testing was carried out in a sound treated audiometric booth, with the participant seated approximately 1 m from the loudspeaker placed at 0° azimuth. Test items were presented through a GENELEC 6010A active portable loudspeaker at a comfortable level in the normal conversational range (65-75 dB SPL) measured using a sound level meter at the position of the participant’s seat. Participants were tested in their normal listening conditions (using CI and hearing aid in case of bimodal use) with regular settings. Before each test, the examiner explained each task and introduced a few practice items to ensure participant understanding of the tasks. Upon the participant’s request, repetition of test items is allowed for the DANVA 2 Adult Paralanguage subtest (as per the test manual instructions). The order of presentation of different tests was randomized across participants.
Table 23. Individual test score profiles for CI participants

<table>
<thead>
<tr>
<th>Participants</th>
<th>M/F</th>
<th>Age (years)</th>
<th>PEPS-C TER</th>
<th>PEPS-C AR</th>
<th>PEPS-C CR</th>
<th>PEPS-C CSR</th>
<th>DANVA 2 (correct/24)</th>
<th>MBEA Contour (correct/30)</th>
<th>MBEA Interval (correct/30)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CI-1</td>
<td>F</td>
<td>73;4</td>
<td>100</td>
<td>63</td>
<td>100</td>
<td>75</td>
<td>12</td>
<td>21</td>
<td>16</td>
</tr>
<tr>
<td>CI-2</td>
<td>M</td>
<td>69;9</td>
<td>88</td>
<td>38</td>
<td>100</td>
<td>94</td>
<td>14</td>
<td>18</td>
<td>11</td>
</tr>
<tr>
<td>CI-3</td>
<td>F</td>
<td>52;9</td>
<td>88</td>
<td>56</td>
<td>88</td>
<td>88</td>
<td>9</td>
<td>19</td>
<td>12</td>
</tr>
<tr>
<td>CI-4</td>
<td>F</td>
<td>48;1</td>
<td>94</td>
<td>100</td>
<td>100</td>
<td>94</td>
<td>15</td>
<td>21</td>
<td>20</td>
</tr>
<tr>
<td>CI-5</td>
<td>F</td>
<td>78;0</td>
<td>88</td>
<td>50</td>
<td>100</td>
<td>100</td>
<td>12</td>
<td>18</td>
<td>24</td>
</tr>
<tr>
<td>CI-6</td>
<td>M</td>
<td>72;2</td>
<td>50</td>
<td>38</td>
<td>94</td>
<td>69</td>
<td>13</td>
<td>22</td>
<td>14</td>
</tr>
<tr>
<td>CI-7</td>
<td>M</td>
<td>25;5</td>
<td>81</td>
<td>63</td>
<td>100</td>
<td>81</td>
<td>12</td>
<td>11</td>
<td>14</td>
</tr>
<tr>
<td>CI-8</td>
<td>F</td>
<td>56;3</td>
<td>56</td>
<td>44</td>
<td>100</td>
<td>100</td>
<td>16</td>
<td>19</td>
<td>19</td>
</tr>
<tr>
<td>CI-9</td>
<td>M</td>
<td>66;1</td>
<td>69</td>
<td>31</td>
<td>94</td>
<td>94</td>
<td>12</td>
<td>16</td>
<td>18</td>
</tr>
<tr>
<td>CI-10</td>
<td>F</td>
<td>68;2</td>
<td>88</td>
<td>81</td>
<td>94</td>
<td>94</td>
<td>18</td>
<td>17</td>
<td>16</td>
</tr>
<tr>
<td>CI-11</td>
<td>F</td>
<td>38;6</td>
<td>81</td>
<td>56</td>
<td>88</td>
<td>100</td>
<td>11</td>
<td>17</td>
<td>16</td>
</tr>
<tr>
<td>CI-12</td>
<td>F</td>
<td>61;9</td>
<td>50</td>
<td>31</td>
<td>94</td>
<td>88</td>
<td>13</td>
<td>13</td>
<td>14</td>
</tr>
<tr>
<td>Mean (SD)</td>
<td></td>
<td>58.4 (16.18)</td>
<td>77.75 (17.29)</td>
<td>54.25 (20.67)</td>
<td>96.0 (4.67)</td>
<td>89.75 (10.08)</td>
<td>13.08 (2.39)</td>
<td>17.66 (3.22)</td>
<td>16.16 (3.63)</td>
</tr>
</tbody>
</table>

Note. TER-Turn-End Reception; AR-Affect Reception; CR-Chunking Reception; CSR-Contrastive Stress Reception.
Results

Comparison of PEPS-C scores in CI users and adults with NH

PEPS-C normative data collected from 15 New Zealand English speaking adults (M<sub>age</sub> = 32.0, SD = 10.0) with NH (Gharahdaghi, 2013) were compared to the PEPS-C scores obtained from the 12 adults using CIs (Table 23). An independent sample t test showed poorer performance in the CI group for three of the four PEPS-C subtests. Figure 8 shows that the Turn-End scores (M = 77.75, SD = 17.2) for CI users were significantly poorer (<i>t</i>(11.5) = 4.01, <i>p</i> = .002) than for the NH group (M = 98.00, SD = 2.9). Similarly, Affect (M = 54.25, SD = 20.67) and Contrastive Stress scores (M = 89.75, SD = 10.08) in CI users were also significantly poorer than Affect (M = 85.93, SD = 10.19; <i>t</i>(25) = 5.21, <i>p</i> = .001) and Contrastive Stress (M = 97.20, SD = 4.45; <i>t</i>(14.4) = 2.38, <i>p</i> = .032) scores of listeners with NH.

Both groups did well on the Chunking Reception subtest, with high scores for CI (M = 96.00, SD = 4.67) and NH groups (M = 98.80, SD = 2.48; <i>t</i>(15.9) = 1.87, <i>p</i> > .05). Table 24 presents means, standard deviations, and ranges of PEPS-C scores for the two groups. The ranges for scores on Turn-End, Affect, and Contrastive Stress Reception subtests for CI and NH groups indicate that there is a wider variation in the CI group’s performance than the NH group.

Both groups scored the lowest for the Affect Reception subtest. The mean scores of NH listeners showed an overall ceiling effect for the Chunking Reception subtest. Three participants in the NH group performed differently from the rest of the group (outliers in Figure 8), however. These three participants were aged 25, 32, and 57 years. Two of these participants were bilingual, but other NH adults performing at ceiling for this task were also bilingual, so it is not clear why these three NH adults performed more poorly.
Figure 8. Percent correct scores obtained by CI users and adults with NH on four subtests of PEPS-C. Boxes indicate the data falling between the 25th and 75th percentile and the whiskers indicate the 95% confidence intervals. The median scores are indicated by the thick horizontal line. Scores on all tests were significantly better in the NH group compared to the CI group, except on Chunking Reception subtest (outlier in the figure indicates scores for three participants). A percent correct score of 75 or above indicates responses above chance.
Table 24. Means, standard deviations, and ranges for scores on PEPS-C subtests and comparison of mean scores between CI users and adults with NH

<table>
<thead>
<tr>
<th>Group</th>
<th>Turn-End Reception</th>
<th>Affect Reception</th>
<th>Chunking Reception</th>
<th>Contrastive Stress Reception</th>
</tr>
</thead>
<tbody>
<tr>
<td>CI (N=12)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>77.75</td>
<td>54.25</td>
<td>96.0</td>
<td>89.75</td>
</tr>
<tr>
<td>SD</td>
<td>17.29</td>
<td>20.67</td>
<td>4.67</td>
<td>10.08</td>
</tr>
<tr>
<td>Minimum</td>
<td>50</td>
<td>31</td>
<td>88</td>
<td>69</td>
</tr>
<tr>
<td>Maximum</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>NH (N=15)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>98.0</td>
<td>85.93</td>
<td>98.8</td>
<td>97.2</td>
</tr>
<tr>
<td>SD</td>
<td>2.92</td>
<td>10.19</td>
<td>2.48</td>
<td>4.45</td>
</tr>
<tr>
<td>Minimum</td>
<td>94</td>
<td>69</td>
<td>94</td>
<td>88</td>
</tr>
<tr>
<td>Maximum</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>p values</td>
<td>.002</td>
<td>.001</td>
<td>.079#</td>
<td>.032</td>
</tr>
</tbody>
</table>

Note. p < .05 is threshold of significance; #No significant difference between the means, p > .05.

Performance on the DANVA 2 Adult Paralanguage subtest

The DANVA 2 Adult Paralanguage subtest has four response options for each trial (happy, sad, angry, and fearful) so a score of six correct out of 24 items represents chance performance. Saindon (2010) reported that the average score for NH group aged 19-58 years ($M_{age} = 29.0$, $SD = 13.8$; $N = 12$) on the DANVA 2 Adult Paralanguage subtest was 19.3 ($SD = 2.3$). The present study utilized the mean scores of the NH group from this previous study to interpret the performance of CI users. The average score obtained for the CI users was 13.08 ($SD = 2.39$), significantly below Saindon’s (2010) NH group score (Z test, $p < .001$). The mean score of 13.08 obtained by CI users in the present study is slightly better than the average score of 10.8 ($SD = 3.30$) reported by Saindon (2010) for six CI users (speech processing strategy not mentioned) with mean age 62.2 ($SD = 13.0$). Figure 9 shows the mean score and standard deviation for the DANVA 2 Adult Paralanguage subtest for Saindon’s (2010) NH group and individual scores for CI participants from the present study. Only three of the CI users (CI-4, CI-8, CI-10) in the current study performed within two SDs of the NH mean reported by Saindon (2010), with the remaining participants scoring more
than two SDs below the norm and one CI user performing at close to chance levels (CI-3). No significant correlation (Spearman’s) was observed between DANVA 2 Adult Paralanguage subtest and PEPS-C or MBEA scores.

**Figure 9.** Mean and standard deviation of the DANVA 2 Adult Paralanguage subtest scores for listeners with NH (N = 12) (Saindon, 2010) (grey bar) and individual scores for CI users from present study (N = 12) (black bars). A score of six correct responses indicate chance performance; three participants (CI-4, CI-8, & CI-10) performed within +/- two SDs of NH mean reported by Saindon (2010); CI-8 & CI-9 used SPEAK strategy.

Table 25 shows the individual error profiles of CI participants on the DANVA 2 Adult Paralanguage subtest. Overall, participants made more errors for low intensity items (M = 6.08, SD = 2.27) than high intensity items (M = 4.83, SD = 1.8) but this difference was not statistically significant (Wilcoxon Signed Ranks test, Z = 1.25, p > .05).

**Contour and Interval subtests of the MBEA**

MBEA Contour and Interval subtests have two response options per trial so a score of 15 correct out of 30 items represents chance performance. Mean scores for CI users in were 17.6 (SD = 3.2) and 16.1 (SD = 3.6) for Contour and Interval subtests. These results are similar to the mean scores of approximately 18 correct responses for Contour and Interval subtests.
obtained for 12 CI users ($M_{age} = 50$) reported by Cooper et al. (2008). Only two CI participants (CI-4, CI-8) had MBEA scores more than 1 SD above a chance level of 15 for both tasks. These two participants performed relatively better than others on PEPS-C Chunking and Contrastive Stress Reception and DANVA 2 subtests.

Table 25. Individual error profiles of CI participants on the DANVA 2 Adult Paralanguage subtest

<table>
<thead>
<tr>
<th>Participant</th>
<th>Age</th>
<th>Errors on High Intensity</th>
<th>Errors on Low Intensity</th>
<th>Composite error score</th>
</tr>
</thead>
<tbody>
<tr>
<td>CI-1</td>
<td>73;4</td>
<td>7</td>
<td>5</td>
<td>12</td>
</tr>
<tr>
<td>CI-2</td>
<td>69;9</td>
<td>4</td>
<td>6</td>
<td>10</td>
</tr>
<tr>
<td>CI-3</td>
<td>52;9</td>
<td>6</td>
<td>9</td>
<td>15</td>
</tr>
<tr>
<td>CI-4</td>
<td>48;1</td>
<td>5</td>
<td>4</td>
<td>9</td>
</tr>
<tr>
<td>CI-5</td>
<td>78;0</td>
<td>4</td>
<td>8</td>
<td>12</td>
</tr>
<tr>
<td>CI-6</td>
<td>72;2</td>
<td>6</td>
<td>5</td>
<td>11</td>
</tr>
<tr>
<td>CI-7</td>
<td>25;5</td>
<td>8</td>
<td>4</td>
<td>12</td>
</tr>
<tr>
<td>CI-8</td>
<td>56;3</td>
<td>4</td>
<td>4</td>
<td>8</td>
</tr>
<tr>
<td>CI-9</td>
<td>66;1</td>
<td>5</td>
<td>7</td>
<td>12</td>
</tr>
<tr>
<td>CI-10</td>
<td>68;2</td>
<td>2</td>
<td>4</td>
<td>6</td>
</tr>
<tr>
<td>CI-11</td>
<td>38;6</td>
<td>2</td>
<td>11</td>
<td>13</td>
</tr>
<tr>
<td>CI-12</td>
<td>61;9</td>
<td>5</td>
<td>6</td>
<td>11</td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>58.4 (16.18)</td>
<td>4.83 (1.8)</td>
<td>6.08 (2.27)</td>
<td>10.91 (2.39)</td>
</tr>
</tbody>
</table>

Peretz et al. (2003) reported mean correct responses and standard deviations for the 30 experimental trials for Contour and Interval subtests obtained from 160 adults with NH aged 14-79 years. Mean correct responses for Contour and Interval subtests were 27 (SD = 2.2) and 26 (SD = 2.4), respectively. Z tests showed significantly poorer performance for the CI group in the current study compared to Peretz et al’s normative data for both Contour and Interval subtests ($p < .001$). Figure 10 shows scores for MBEA Contour and Interval subtests obtained by Peretz et al’s (2003) NH group and the CI group from the present study. Only one CI user (CI-5) performed within two SDs of the NH mean (Peretz et al., 2003) for the
Interval subtest while the remaining participants scored more than two SDs below the NH mean for both subtests (Table 23).

Figure 10. Mean scores and standard deviations for MBEA Contour and Interval subtests obtained by NH group (Peretz et al., 2003) and CI group from the current study. MBEA-Montreal Battery of Evaluation of Amusia; NH group mean scores were 27 (SD = 2.2) and 26 (SD = 2.4) for MBEA Contour and Interval subtests respectively; CI group mean scores were 17.6 (SD = 3.2) and 16.1 (SD = 3.6) for MBEA Contour and Interval subtests respectively.

Demographic variables and speech perception measures

PEPS-C total scores were compared with age at time of testing, duration of implant use, duration of musical experience, and performance on the two speech perception measures (Consonant Nucleus Consonant (CNC) test and Hearing in Noise Test (HINT)). Percent correct scores obtained on the CNC test and HINT scores at +10 dB SNR are provided in Table 22. One participant (CI-8) was excluded from one of the correlation analyses as HINT scores were not available. Spearman rank correlations showed no significant association between PEPS-C total scores and demographic variables or speech recognition measures.
Correlations with hearing thresholds

Aided thresholds in the sound field for bimodal participants \(n = 3\) and unilateral CI users \(n = 9\) were obtained from the participants’ records. The unaided thresholds in the contralateral ear of the unilateral CI users were also obtained. If no response was recorded at the maximum output of the audiometer, the threshold was assumed to be 5 dB greater than the maximum level tested. Spearman correlation coefficients were calculated for PEPS-C total scores, DANVA 2, and MBEA Contour and Interval subtest scores with aided and unaided thresholds at 250, 500, 1000, 2000, and 4000 Hz. Due to multiple comparisons, a Bonferroni correction was applied. There were no significant findings \((p \geq .134)\).

Discussion

Performance of CI users on PEPS-C Turn-End, Affect, and Contrastive Stress Reception subtests was significantly poorer than that of a New Zealand sample of adults with NH (Gharahdaghi, 2013). Performance on the Chunking Reception subtest did not differ from the norm. Overall, the CI group had lower scores and higher individual variability across the PEPS-C subtests than the NH group, and both groups performed most poorly on the Affect Reception subtest. The PEPS-C Chunking subtest requires judging speakers’ use of timing cues whereas Turn-End, Affect, and Contrastive Stress subtest mainly require perception of subtle pitch cues. These results are consistent with previous research findings indicating that CI subjects perform significantly worse than NH listeners on pitch perception measures (Gfeller et al., 2002a; Looi, McDermott, McKay, & Hickson, 2004). F0 is not explicitly coded in ACE or SPEAK strategies, however F0 information can be extracted based on perception of amplitude modulation across implant channels (Vandali et al., 2000; Kiefer, Hohl, Stürzebecher, Pfennigdorff, & Gstöettner, 2001). Meister et al. (2009) investigated the effects of F0 cues on the discrimination of question/statement and sentence stress in German speaking CI users. They reported that most errors in question/statement discrimination task
occurred when there was limited variation in F0 and that CI recipients could not accurately interpret stress patterns even at very high alterations of F0.

Performance on the PEPS-C Chunking Reception subtest did not differ between CI and NH groups. This subtest involves distinguishing between two or three intonational units (e.g., /FRUIT-SALAD/and MILK/ versus /FRUIT/SALAD/and MILK/). Studies have provided evidence linking chunking (syntactic parsing) to differences in patterns of duration (pause duration, preboundary lengthening) and pitch changes (Cooper & Paccia-Cooper, 1980). Five sets of test stimuli used in PEPS-C Chunking task were randomly selected and analysed using PRAAT software (Boersma & Weenink, 2001). This showed that the items are distinguished mainly based on pause duration differences rather than pitch cues. The mean pause durations between the items in the two intonational phrase task, the first and second item within the three intonational phrase task, and that between the second and the third items within the three intonational phrase task were 258, 438, and 354 ms respectively. Good performance on this task indicates that the CI participants could use the prominent durational cues to perceive this prosodic contrast accurately. The pause durations between the items in the Chunking Reception subtest are much larger than gap DLs reported for CI users. Moore and Glasberg (1988) reported that CI users can identify temporal gaps of 4.3 ms and 4.7 ms at 1 and 2 kHz respectively for sinusoids, however, CI users do show reduced temporal resolution abilities compared to NH listeners for temporal modulation detection and time compressed speech recognition tasks (Fu, Galvin, & Wang, 2001; Won, Drennan, Nie, Jameyson, & Rubinstein, 2011).

Adult CI users are able to discriminate simple rhythmic patterns when presented at a moderate rate and they show superior performance on rhythm tasks relative to pitch perception tasks (Gfeller et al., 2002a; Cooper et al., 2008). Rhythm discrimination is generally good in CI users but not quite as good as listeners with NH (Leal et al., 2003; Kong et al., 2004; McDermott, 2004). Rhythmic differences are encoded in CIs as temporal gaps,
amplitude modulations, or both (Drennan & Rubinstein, 2008). Similar performance on the Chunking Reception subtest by CI and NH groups could be due to better temporal discrimination than pitch perception in the CI group. The use of high stimulation rates enables neural firing patterns to be similar to acoustic hearing resulting in relatively good temporal resolution (Wilson, Finley, Lawson, & Zerbi, 1997; Rubinstein, Wilson, Finley, & Abbas, 1999; Shannon, 1993). Most participants used ACE which has a much higher stimulation rate (900 Hz) than SPEAK (250 Hz). The moderate-rate 900 Hz ACE strategy is currently the default choice for the Nucleus 24 devices because of its reported higher speech perception scores compared to SPEAK (Holden, Skinner, Holden, & Demorest, 2002; Skinner et al., 2002). The two SPEAK users performed similarly to the ACE users on the Chunking Reception subtest, however, which could reflect the relatively obvious temporal cues in the chunking task.

Consistent with the PEPS-C Affect Reception subtest data, the CI group performed significantly poorer than Saindon’s (2010) NH group on the Adult Paralanguage subtest of DANVA 2. Performance on the Adult Paralanguage subtest by CI participants in the present study is similar to adult CI users reported by Saindon (2010). In both studies CI users were unsuccessful in differentiating vocal emotions. Murray and Arnott (1993) reported that pitch envelope including the level, range, shape and timing of pitch contours are the important parameters that allow listeners to differentiate between emotions. They also reported the pitch changes in different emotions relative to neutral speech as: 1) abrupt on stressed syllables (anger), 2) smooth upward inflections (happiness), 3) downward inflections (sadness), and 4) wide pitch range (fear). Identifying the direction of pitch change over the duration (pitch contour) of a syllable or sentence, or from one musical note to the next is relevant to affect recognition and music perception. Poor performance on the Adult Paralanguage subtest suggests that CI users are unable to perceive abrupt and subtle changes
in pitch contour accurately, consistent with their poor performance on the MBEA musical subtests.

Overall, the CI participants made more errors with low intensity test stimuli than high intensity stimuli for the DANVA 2 Adult Paralanguage subtest, but this varied across participants and there was no statistical difference in performance as a function of emotional intensity. No specific pattern of difficulty was observed for different emotions. Previous studies have reported large variation in accuracy of auditory perception across emotions in individuals with HL (Hopyan-Misakyan et al., 2009; Most & Aviner, 2009). Children with HL have been reported to identify happy and sad utterances better than other emotions (Saarni, 1999) which may be because these are the earliest emotion categories acquired by children (Markham & Adams, 1992) or the perceptual cues for these emotions may be more salient. Postlingual adult CI users such as the participants in the current study should have well established mental representations of various emotion categories (Nakata et al., 2012) and perception of different emotion categories should therefore depend on the ability to detect subtle pitch and durational changes (Murray & Arnott, 1993; Luo, Fu, & Galvin, 2007). Nine of the participants in the current study were postlingually deafened, suggesting that the poor performance overall of CI users reflects limited coding by the CI of cues important for affective prosody perception. PEPS-C Affect scores only reflect perception of like/dislike. Future research examining the identification of a greater range of emotional expressions in CI users and analyzing the acoustic characteristics of different emotions is recommended. The lack of correlation between PEPS-C scores and other measures used in this study could reflect a lack of statistical power. Further investigations with a larger sample size would clarify the nature of the relationship between PEPS-C and DANVA 2 scores.

As expected (Peretz et al., 2003; Cooper et al., 2008) CI users performed poorer than NH listeners on MBEA Contour and Interval subtests. Most of the CI participants in the present study performed at chance level on both MBEA tasks. This is in line with Wright and
Uchanski’s (2012) findings that CI users’ performance was not significantly different from chance level performance on MBEA Scale, Contour, and Interval subtests. The MBEA results in the current study indicate that pitch contours and pitch interval cues that are important for music perception is not accurately conveyed by CI strategies. CI users can discriminate pitch and hear musical intervals solely based on stimulation rate, but this temporal-pitch mechanism is functional only at low rates up to 300 pps/channel (Zeng, 2004). The high stimulation rate (900 Hz) used by most of the CI participants in this study would not support this mechanism for musical pitch perception. Although pitch can be perceived via temporal envelope cues (temporal gaps, amplitude modulations) this is reduced for CI users relative to NH listeners (Fu et al., 2001; Green, Faulkner, & Rosen, 2002).

Participants in this study were a convenience sample from the local implant program; only three adults with bimodal stimulation participated (Table 22). These participants differed in age, duration of deafness, and etiology of HL and hence it is not possible to draw conclusions regarding the effect of bimodal stimulation on prosody perception. Studies have reported that bilateral CIs and bimodal fitting in children and adults provide benefit over unilateral devices for speech perception in noise and sound localization (Litovsky, Johnstone, & Godar, 2006; Mok, Grayden, Dowell, & Lawrence, 2006; Sammeth, Bundy, & Miller, 2011) but to our knowledge few studies have examined prosody perception in bimodal CI users (Landwehr et al., 2007; Straatman et al., 2010; Most et al., 2011). Future research should be conducted on a larger sample with controlled variables to evaluate the effects of bimodal stimulation on perception of prosody and music.

**Summary and conclusions**

The current study provides evidence for poor performance on tasks involving different aspects of prosody in adult CI users. Although there are indications from previous studies, such as Meister et al. (2009), that CI users have difficulty perceiving questions versus
statements, so far no studies have investigated different aspects of prosody in CI users. All participants could be classified as ‘good’ CI users based on their mean percent scores on CNC ($N = 12, 64.0\%$) and HINT scores at $+10$ dB SNR ($N = 11, 94.5\%$). For example, these scores are better than those reported by Miller, Watson, Kistler, Wightman, and Preminger (2008) for adult CI users with mean age of 57.8 years (CNC, $N = 8, 47\%$; HINT scores at $+10$ dB SNR, $N = 8, 41\%$). Despite this, only three participants had emotional prosody performance within normal limits on the Adult Paralanguage subtest of DANVA 2 and all participants performed significantly poorer (except for Chunking Reception subtest) than the NH group on PEPS-C. Difficulty in accurately perceiving different aspects of prosody and music could be attributed to reduced pitch discrimination ability. The results are supported by comments from the participants. For example, when asked about musical experience after cochlear implantation, one participant in the current study reported that “listening to music is not what I would do for relaxing, because it requires a lot of effort and at the end there is no satisfaction”. It could be assumed that while many CI users enjoy success regarding speech recognition, most of them are still frustrated by their inability to accurately perceive music (Koelsch, Wittford, Wolf, Müller, & Hahne, 2004). Overall, the findings of the present study indicate the need for further development of CI processing strategies to provide accurate representation of pitch which is important for perceiving prosody and enjoyment of music.
Chapter 6: Discussion and conclusion

This concluding chapter summarises the major findings of the four studies presented in this doctoral thesis, describes its implications for clinical practice, and suggests directions for future research.

Prosody assessment in clinical setting

The review of literature outlined in Chapter 2 reported the nine published tools that can be used to assess receptive and expressive prosodic skills in children and adults with a range of communication disorders including ASD, hearing loss, and other neurologic and psychiatric disorders. The different conceptual frameworks on which the tools have been based: psycholinguistic (PEPS-C and PPAT), neurolinguistic (Aprosodia Battery), and social and cognitive functioning (DANVA 2, MNTAP, FAB, and ACS) reflect the many identified roles of prosody. The findings of this review suggested that PEPS-C, DANVA 2, PROP, and PVSP are applicable for a wide age range (Chapter 2) and diverse clinical populations. Tables 4 and 5 provide details of the subtests involved, domains of prosody assessed, normative data, and reliability and validity information for each test. This paper serves as a guide for clinicians who wish to assess prosodic skills using different paradigms in people with a range of communication disorders.

None of the tools described in this review was standardised. Norms were provided by test developers for the PEPS-C, DANVA 2, ACS, FAB, and PVSP tests. Norms for the PPAT, Aprosodia battery, and MNTAP reported in this review were obtained from studies that used these tools to compare the performance of disordered populations and control groups (Chapter 2). Stratified norms based on age, gender, and geographic locations are not available for any of these tools. Norms provided by test developers (e.g. PEPS-C was normed on 120 British-English speaking children and DANVA 2 on North Americans aged 8-10
years) may not be appropriate for use with children who belong to different ethnic groups as there are linguistic and cultural prosody differences even among speakers of English (Clopper & Smiljanic, 2011; Coggshall, 2008). Hence, clinicians need to be cautious when applying these norms to their target clinical population and should consider cross-dialect prosodic variations while assessing speakers of Afro-Caribbean, Singaporean, or Indian versions of English (Grabe & Low, 2002; Grabe, 2006; Low, Grabe, & Nolan, 2001).

This review revealed that the purpose of development and the content of the tools were well documented, but data on how feasible they were to use in practice was scarce. Each tool met some but not all the widely accepted criteria for validity and reliability. Most have not been sufficiently well tested for use in routine clinical practice. We noted that there are few psychometrically sound standardised tools in the area of typical prosody development, as suggested in Diehl and Paul (2009). The need to continue to develop and test tools for effective and comprehensive assessment of prosodic skills was highlighted in this review. Future research should gather empirical data using the available tools and further explore the establishment of standardised diagnostic measures to evaluate prosody. Future studies should investigate the effectiveness of these tools in highlighting specific aspects of prosody warranting clinical intervention.

**Differential pattern of prosody acquisition**

Chapter 3 reported significant age-related improvements in prosody perception in typically developing New Zealand-English speaking 7-12 year olds using the PEPS-C test. The results showed that 7-8 year olds performed significantly poorer than 9-12 year olds on PEPS-C Chunking and Contrastive Stress Reception subtests. However, there were no significant differences between the older and younger age groups on Turn-End and Affect Reception tasks. This study showed differential patterns for the acquisition of different prosodic skills, consistent with Wells et al.’s (2004) study of British-English speaking children, previous
studies examining children’s prosody production (Doherty et al., 1999; Grigos & Patel, 2010), and the infant directed speech literature (Durkin et al., 1982; Soderstrom et al., 2008; Trainor et al., 2000). This study described in Chapter 3 reported PEPS-C and DANVA 2 scores obtained from 45 children, divided into three age groups with mean ages 7.84, 10.13, and 11.90 years. This provides stratified norms based on age for typically developing New Zealand-English speaking children which can be used as a normative comparison sample to determine whether prosody perception is atypical in children with suspected prosodic difficulties. However, these findings need to be replicated with a larger sample and in younger age groups (5-6 year olds). Gibbon and Smyth (2013) assessed prosodic abilities of 30 typically developing 4 year olds using the Irish version of the PEPS-C test and reported that the majority (83%) of the children were able to complete the test. They reported that 4 year olds had lower scores than 5-6 year olds in all subtests. These results indicate that PEPS-C could be a valuable tool for assessing prosody in young and preschool children, however age appropriate norms need to be established.

**Emotion recognition (affective prosody)**

DANVA 2 scores obtained from typically developing children were analysed based on the level of emotion intensity and emotional category. Results varied across age groups for high emotion intensity items; 7-8 year olds made more errors than 9-10 year olds but did not differ from 11-12 year olds. For low emotion intensity items, total scores did not differ significantly across age groups although there was a trend for fewer errors amongst the 11-12 year olds (Table 9, Chapter 3). This finding is important given that in everyday settings emotional expressions are often subtle (Russell & Barrett, 1999). As suggested by previous studies, it was concluded that considering the level of emotion intensity may be useful in identifying error patterns associated with various communication disorders (Castelli, 2005; Greimel et al., 2010; Grossman & Tager-Flusberg, 2012; Mazefsky & Oswald, 2007). The confusion
matrix for DANVA 2 errors in typically developing children showed that some pairs of emotions were confused with one another more often than others (Chapter 3). Although the effect of various acoustic cues involved in emotion recognition was not probed in this study, the confusion matrix results are consistent with previous literature reporting variations in acoustic cues used to perceive different emotions and cue trading (Banse & Scherer, 1996; Juslin & Laukka, 2003). Further research into the mechanisms by which children develop the ability to recognise different emotions may be useful for identifying therapy approaches to enhance emotional prosody perception in children with prosody disorders. It would be worthwhile investigating the effect of level of emotion intensity in both facial and auditory affect perception in ASD, ADHD, and other neurologic and psychiatric disorders.

**Assessment and intervention of prosodic deficits in children with hearing loss**

The study outlined in Chapter 4 found that children with hearing loss performed significantly poorer than controls on PEPS-C and DANVA 2 tests, which suggest that clinical assessment and therapy services for children with hearing loss should be expanded to target prosodic difficulties. For the four children with unilateral hearing loss who participated in this study, unilateral hearing loss appeared to be associated with poorer prosody perception. This finding has implications for early identification and intervention of unilateral hearing loss in children. The finding that better hearing thresholds were associated with better prosody perception (Chapter 4) suggests that greater degrees of hearing loss can be detrimental for the development of age-appropriate prosodic skills. The improvements in prosody perception with age in children with hearing loss are consistent with the age effects reported in typically developing children (Chapter 3). The findings also suggest that PEPS-C and DANVA 2 tests can be used to differentiate prosodic abilities in children with hearing loss and controls (discriminant validity). Nevertheless, future studies should collect normative data for larger sample using PEPS-C and DANVA 2 for comparison with clinical populations. Small sample
size and heterogeneity of the sample in terms of degree and type of hearing loss and age of amplification precluded the comparison between amplification devices (hearing aids vs. CIs). This is an important gap that warrants future research. Future studies should examine the effects of bimodal stimulation on prosody perception in children. Studies have shown that using a combination of electric and acoustic stimulation has the potential to improve perceptual outcomes for CI users (Gfeller et al., 2007; Kong et al., 2005).

Researchers have reported that there is no solid evidence-base concerning treatment of impaired prosody (Arciuli, 2014; Diehl & Paul, 2009). Dynamic Temporal and Tactile Cueing (DTTC, Strand et al., 2006), Melodic Intonation Therapy (MIT, Sparks & Holland, 1976), and the Lee Silverman Voice Treatment (LSVT®, Ramig et al., 1995) have been used with adults with dysarthria and aphasia. A few programmes that have been developed for use with adults with acquired neurological disorders (imitative approach, cognitive-linguistic approach, knowledge of results, knowledge of performance, Visi-Pitch) are now being tried with children with Childhood Apraxia of Speech (McCabe, Macdonald-D’Silva, Van Rees, Ballard, & Arciuli, 2014; Van Rees, Ballard, McCabe, Macdonald-D’Silva, & Arciuli, 2012). There is a lack of empirical research into the efficacy of treatments targeting prosody in people with hearing loss. The recently developed Prosody Treatment Program (Rothstein, 2013) is a computer software program that can potentially be used to train receptive and expressive prosodic skills in children with hearing loss. However, the efficacy of this treatment program is not yet reported.

Another significant finding of the study of children with hearing loss (Chapter 4) was the correlation between musicality (composite MBEA scores and musical experience) and prosody perception scores. The significant correlation between DANVA 2 low emotion intensity scores and musicality suggests that musical experience improves children’s sensitivity to subtle acoustic/pitch cues, however this was a cross-sectional study hence causality cannot be determined. These findings suggest that musical training may be
beneficial for improving receptive prosody skills. Children who continue to have musical training may have better overall auditory perceptual skills from the outset and hence prospective longitudinal studies are needed to clarify the contribution of musical training to improved prosody perception. The correlation between musicality and prosody perception described in Chapter 4 is consistent with Torppa et al.’s (2014) report that prior musical experience was linked to perception of stress in children with hearing loss. There was no significant difference between the Finnish speaking children with CIs and children with normal hearing for the perception of word and sentence stress. The participants were implanted before 3 years of age and all had musical experience which could have contributed to their performance being similar to their hearing peers. The age of implantation of the participants in our study was more variable (Chapter 4). The correlation between musicality and children’s prosody perception needs further investigation with a larger sample and a prospective study design. Future research should also explore links between prosody and related areas such as perception of nonliteral language and reading in children (Goswami et al., 2010; Holliman et al., 2014).

**Perceptual rating of prosody**

The perceptual ratings of children’s prosody productions by nine experienced listeners revealed that there was greater variability for the ratings of pitch, pitch variations and overall prosody in the hearing loss group compared to the control group (Chapter 4). Inter-rater agreement showed significant agreement between the nine raters which suggests that prosodic features could be reliably assessed clinically without special equipment or analysis. This study provides preliminary evidence for the potential application of perceptual ratings for the clinical assessment of prosody in people with hearing loss. However, future research should examine how perceptual ratings relate to the acoustic measurements of prosodic features in children with hearing loss and also investigate if there are any consistent markers
of atypical prosody in these children. Nadig and Shaw (2013) compared atypical prosody in speakers with high-functioning autism (HFA) compared to controls using acoustic measurements of a conversational speech sample and a structured speech task and perceptual ratings of conversation. Increased pitch range was found in the HFA group during both conversation and structured communication. A similar study design could be used in future studies in children with hearing loss.

**Access to pitch cues through cochlear implants**

The study described in Chapter 5 reported that the performance of adult CI users on PEPS-C Turn-End, Affect and Contrastive Stress Reception subtests was significantly poorer than controls whereas performance on Chunking Reception subtest did not differ from the New Zealand adult norm. This finding indicates that CI users were able to discriminate duration cues better than pitch cues, as suggested by previous studies (Rubinstein et al., 1999; Shannon, 1993; Wilson et al., 1997). Chance level performance on both MBEA musical pitch discrimination tasks indicated that pitch contour and pitch interval cues, which are important for music perception, were not accurately perceived by CI users. Researchers have recognised that future speech processing strategies must improve the transmission of pitch cues important for perceiving prosody and music (Luo et al., 2007; Oxenham, 2008). Fine structure processing (FSP) strategies such as FS4 and FS4p used in MED-EL systems are intended to provide better signal fine structure compared to CIS and SPEAK strategies (Lorens, Zgoda, & Henryk, 2010; Müller et al., 2012). Research has shown that FSP strategies have benefits for speech perception in noise and music appreciation (Arnoldner et al., 2007; Looi, Winter, Anderson, & Sucher, 2011). Benefits of FSP strategies for prosody perception have not yet been investigated. Bimodal stimulation has been shown to improve speech understanding in quiet and in noise beyond that achieved by a hearing aid or cochlear implant alone (Gstoettner et al., 2004; Kiefer et al., 2005; Luo, Fu, & Galvin, 2006).
However, few studies have investigated prosody perception in bimodal users (Landwehr et al., 2006; Straatman et al., 2010) and their findings are limited by low sample size and lack of statistical power. There were only three bimodal adult CI users in the current study; more systematic research on a larger sample with controlled variables is warranted to evaluate effects of bimodal stimulation on perception of prosody and music.

**Closing remarks**

It is widely recognised that accurate perception and production of different aspects of prosody are critical for successful communication (Aziz-Zadeh et al., 2010; Crystal, 2009). Assessment of prosody is not included in the speech perception batteries commonly used to assess people fitted with hearing aids and/or CIs. Prosody is usually ignored in clinical settings which limits the validity of the assessment for determining a person’s everyday communication performance. Our studies recommend the use of several available assessment tools to generate evidence of typical and atypical prosody performance so that these tools can be routinely used in clinics. Together, the studies of children and adults with hearing loss suggest that prosodic deficits in people with hearing loss should be more routinely investigated and treated in clinical practice.
APPENDIX 1: Ethics approval

UNIVERSITY OF AUCKLAND HUMAN PARTICIPANTS ETHICS COMMITTEE

17-Dec-2012

MEMORANDUM TO:

Prof Suzanne Furdy
Psychology

Re: Application for Ethics Approval (Our Ref. 8870)

The Committee considered your application for ethics approval for your project entitled Suprasegmental perception in children using cochlear implants and hearing aids- A comparative study.

Ethics approval was given for a period of three years.

The expiry date for this approval is 12-Dec-2015.

If the project changes significantly, you are required to submit a new application to UAHPEC for further consideration.

In order that an up-to-date record can be maintained, you are requested to notify UAHPEC once your project is completed.

The Chair and the members of UAHPEC would be happy to discuss general matters relating to ethics approvals if you wish to do so. Contact should be made through the UAHPEC Ethics Administrators at humanethics@auckland.ac.nz in the first instance.

All communication with the UAHPEC regarding this application should include this reference number: 8870.

(This is a computer generated letter. No signature required.)

UAHPEC Administrators
University of Auckland Human Participants Ethics Committee

c.c. Head of Department / School, Psychology

Dr Elaine Ballard
Ms Rose Kalathottukaren
Ms Rose Kalathottukaren
Ms RoseKalathottukaren
Additional information:

1. Do not forget to fill in the 'approval wording' on the Participant Information Sheets and Consent Forms, giving the dates of approval and the reference number, before you send them out to your participants.

2. Should you need to make any changes to the project, write to the UAHPEC Administrators by email (humanethics@auckland.ac.nz) giving full details of the proposed changes including revised documentation.

3. At the end of three years, or if the project is completed before the expiry, please advise UAHPEC of its completion.

4. Should you require an extension, write to UAHPEC by email before the expiry date, giving full details along with revised documentation. An extension can be granted for up to three years, after which a new application must be submitted.

5. If you have obtained funding other than from UniServices, send a copy of this approval letter to the Manager - Funding Processes, UoA Research Office. For UniServices contracts, send a copy of the approval letter to the Contract Manager, UniServices.

6. Please note that UAHPEC may from time to time conduct audits of approved projects to ensure that the research has been carried out according to the approval that was given.
APPENDIX 2: Ethics approval amendment

UNIVERSITY OF AUCKLAND HUMAN PARTICIPANTS ETHICS COMMITTEE

13-Mar-2013

MEMORANDUM TO:
Prof Suzanne Purdy
Psychology

Re: Application for Ethics Approval (Our Ref. 8870)

The Committee considered your request for change for your project titled Suprasegmental perception in children using cochlear implants and hearing aids— A comparative study on 13-Mar-2013.

The Committee approved the following amendments:

1. To add a pilot study with adults to the project (all relevant documents included).
2. To distribute a brochure to recruit child participants among schools and among parents.

The expiry date for this approval is 12-Dec-2015.

If the project changes significantly you are required to resubmit a new application to the Committee for further consideration.

In order that an up-to-date record can be maintained, it would be appreciated if you could notify the Committee once your project is completed.

The Chair and the members of the Committee would be happy to discuss general matters relating to ethics approvals if you wish to do so. Contact should be made through the UAHPEC secretary at humanethics@auckland.ac.nz in the first instance.

All communication with the UAHPEC regarding this application should include this reference number: 8870.

(This is a computer generated letter. No signature required.)

Secretary
University of Auckland Human Participants Ethics Committee
APPENDIX 3: Recruitment brochure

Children wanted for research at the University of Auckland

Speech Science in the School of Psychology at the University of Auckland invites you to participate in this PhD research study.

We are investigating understanding of prosody in children aged between 7 to 12 years who have normal hearing or who have hearing loss. Prosody is the way we talk when we express different emotions or change the meaning of sentences by saying them with a different tone.

As the picture shows, children in the study listen to sounds from earphones and point to pictures on a computer screen.

Children would spend two hours with the student researcher (Rose) who will perform hearing, reading and music tests.

Children will be tested at the University of Auckland speech language therapy clinic or at the child’s school or home, whichever is convenient for the family.

A $10 gift voucher will be given to the child for their participation.

Families travelling for appointments will be offered a $20 petrol voucher to contribute to travel costs.

This study is approved by the University of Auckland Human Participants Ethics Committee (UAHPEC) 12 Dec 2012- 12 Dec 2015
If you would like to know more about this research, please contact us

**Rose Thomas Kalathottukaren**  
Audiologist & Speech Language Pathologist/PhD Student  
School of Psychology  
Tamaki Innovation Campus,  
The University of Auckland  
Private Bag 92019, Auckland  
Phone (09) 373 7599 ext. 84930  
rkal180@aucklanduni.ac.nz

**Professor Suzanne C. Purdy**  
Head of Discipline of Speech Science  
School of Psychology  
Tamaki Innovation Campus,  
The University of Auckland  
Private Bag 92019, Auckland  
Phone: (09) 373 7599 ext. 82073 or 86886

**Dr. Elaine Ballard**  
Senior Lecturer in Speech Science  
School of Psychology  
Tamaki Innovation Campus,  
The University of Auckland  
Phone: 373 7599 ext. 87502  
Email: e.ballard@auckland.ac.nz
APPENDIX 4A: Information sheet - Children with hearing loss

General information

- Name of the child (optional): Age: ……….years ……….months
- Date of Birth: Sex: Male/Female
- Languages spoken:
- Details of musical training:
- Address:

Phone: Email:

Details about hearing loss:

- Age at which hearing loss was identified: ……………years ……… months
- Type of hearing loss:
  - Right ear: Conductive/Sensorineural/Mixed hearing loss
  - Left ear: Conductive/Sensorineural/Mixed hearing loss
- Degree of hearing loss:
  - Right Ear: Minimal/Mild/Moderate/Moderately severe/Severe/Profound hearing loss
  - Left Ear: Minimal/Mild/Moderate/Moderately severe/Severe/Profound hearing loss
Details about hearing aids:

- Age at which your child started wearing hearing aids: .......... years .......... months
- Ear: Right ear Left ear Both ears
- Style: BTE /ITE/ITC/CIC
- Manufacturer: Phonak/Widex/Oticon Others:

Details about cochlear implants:

- Age at which your child started using cochlear implants:
- Ear: Right ear Left ear Both ears
- Style: Body worn Processor/ Ear level processor
- Manufacturer: Advanced Bionics/ Cochlear/ Med-El Others:

Details about speech language therapy / aural rehabilitation:

- Did your child receive speech language therapy services? Yes/ No
  If Yes, for how long: ..................Years .................. months
- Did your child receive aural rehabilitation services? Yes/ No
  If Yes, for how long: ..................Years .................. months

Parent/Guardian’s Name: .................. Date / /
APPENDIX 4B: Information sheet - Children with normal hearing

General Information:

- Name of the child (optional):
- Date of Birth:
- Age: ............ years ..............months
- Sex: Male/Female
- Languages spoken:
- Details of musical training:
- Any parental concerns about hearing or speech and language development?
- Address:
- Phone: 
- Email:
- Parent/Guardian’s Name: ................. Date / /
Study title: Perception and production of speech prosody in children with hearing loss

This study is being undertaken as a part of a PhD programme in Speech Science in the School of Psychology, The University of Auckland by Rose Thomas Kalathottukaren. The research project will be guided by Professor Suzanne C. Purdy and Dr. Elaine Ballard. The final report of this study will be reviewed by the University examiners to mark the completion of the project. The proposed research will be completed by May 1st 2015.

The current study requires participants aged between 7 – 12 years with hearing loss and with normal hearing to be a part of this study. We are interested in “prosody/rhythmic” aspects of speech because these are important to understand the emotions and meaning of the speaker. Speech prosody depends on “suprasegmental” cues such as pitch and loudness and timing. Children with hearing loss may have trouble hearing these cues that can convey different meanings and emotions, even when the words in the sentence are the same. Speech prosody plays a major role in day to day communication.

You will also be asked to complete an information sheet regarding your child’s hearing and if your child has a hearing loss we would like to know about their use of hearing instruments.

If you and your child agree to participate, we would like your child to do the following tasks:

1. The PEPS-C (Profiling Elements of Prosodic Systems-Children) will be used to assess your child’s understanding of speech prosody. Your child will listen to words and sentences from the computer that they will hear through a loudspeaker and they respond by pointing on to the computer screen (30 mins)
2. The DANVA 2 (Diagnostic Analysis of Non Verbal Accuracy 2) test will be used to assess child’s understanding of emotions in speech. The “Child Paralanguage” task involves listening to sentences through a loudspeaker and your child will respond by
pointing on to pictures/verbally to say if the person sounds happy, angry, or a different emotion (10 mins)

3. The Wheldall Assessment of Reading Passages (WARP) will be used to assess reading fluency and accuracy. This task involves reading three short passages aloud, for duration of 1 minute for each passage. The child’s responses will be audio recorded (5 mins)

4. Pitch discrimination tasks using music tones: Participants would listen to series of tones/music and would respond verbally to indicate whether they could hear a difference in pitch or not (10 mins)

**Note:** The time taken to complete all tasks depends on the individual and may be slightly longer or shorter than the estimations provided above. The test procedure will be explained to the children prior to the start of the test and if at any stage during testing if the child becomes tired or uncooperative the testing would be stopped immediately and resumed later or in another session if the child is happy to do this.

The testing will take place at the University of Auckland Speech Language Therapy Clinic or at child’s home with the consent of the parent, whichever is convenient for the families. Children will receive a $10 gift voucher for their participation. Families travelling for appointments will be offered $20 petrol vouchers to contribute to travel costs.

**Confidentiality:** Confidentiality will be preserved throughout the research process. If the information collected from children are reported or published, it will be done in a way that does not reveal the identity of the participant as the source of the information will be kept confidential.

Data collected from participants will be stored in a locked cabinet within the University premises for a period of 6 years. The data will be destroyed (paper shredded, disks erased) after 6 years following the completion of the research study. The child’s answers will be audio recorded but no photographs/video recording will be taken during the testing.

**Participation in this study is voluntary.** Participants will have the right to withdraw from the project at any time, and can also withdraw the data within 3 months after the completion of testing. Withdrawing from the project will in no way affect your further health care with your clinic or school or your relationship with the University of Auckland.
The child’s Principal / Advisor on Deaf Children / Speech-language Therapist / Teacher / Audiologist would provide assurance that the participation of your child in this study will not affect their learning support at schools, or any other support services.

Yours sincerely,

Rose Thomas Kalathottukaren  
Audiologist & Speech Language Pathologist/PhD Student  
School of Psychology  
Tāmaki Innovation Campus  
The University of Auckland  
Mobile: 021 08224421  
Email: rkal180@aucklanduni.ac.nz

Professor Suzanne C. Purdy  
Head of Discipline of Speech Science  
School of Psychology  
Tāmaki Innovation Campus  
The University of Auckland  
Phone: 373 7599 ext. 82073 or 86886  
Email: sc.purdy@auckland.ac.nz

Dr. Elaine Ballard  
Senior Lecturer in Speech Science  
School of Psychology  
Tāmaki Innovation Campus  
The University of Auckland  
Phone: 373 7599 ext. 87502  
Email: e.ballard@auckland.ac.nz

For any queries regarding ethical concerns please contact:
The Chair,  
The University of Auckland Human Participants Ethics Committee  
The University of Auckland  
Research Office-Office of the Vice Chancellor  
Private Bag 92019, Auckland.  
Tel. (09) 373 7599 ext. 87830

This study is approved by the University of Auckland Human Participants Ethics Committee (UAHPEC) 12 Dec 2012 to 12 Dec 2015, Reference number: 8870.
APPENDIX 6: Consent form - Parents/caregivers

Study title: Perception and production of speech prosody in children with hearing loss

Names of Researchers: Professor Suzanne C. Purdy (Head of Speech Science, The University of Auckland); Dr Elaine Ballard (Senior Lecturer, Speech Science, The University of Auckland); Ms Rose Thomas Kalathottukaren (Audiologist & Speech Language Pathologist/PhD Student, The University of Auckland)

- I have read the Participant Information Sheet and have understood the nature of the research designed to investigate speech perception in children with hearing loss and normal hearing children.

- I have had the opportunity to discuss the study, ask questions and have had them answered to my satisfaction.

- I agree that the researchers may access my child’s clinical records for the purpose of this study, if I need additional information about their hearing.

- I understand that my child’s answers will be audio recorded, and that these will be kept confidential. No photographs/video recording will be done during testing.

- I understand that the participation/non participation of my child in this study will not affect his/her future learning support in their school or audiology/speech language therapy clinic.

- I understand that my child may need to spend 1 session lasting for about 2 hours with the researcher (or 2 shorter sessions if this is preferred).

- I understand that the time required to complete testing may vary based on child’s cooperation and fatigue.

- I understand that there is a chance that your child’s hearing has changed from the previous tests and that we will provide copies of the hearing test so that this can be checked by you and your child’s clinicians.

- I understand that I am free to withdraw participation at any time

- I understand that I will be free to withdraw my data up to 3 months after the date on which my child was tested.
• I understand that if the findings are reported or published my identity and my child’s identity will be kept confidential

• I understand that data will be kept anonymous and stored for a period of 6 years, after which they will be destroyed.

• I know whom to contact if I have queries regarding the study.

• Do you wish to receive written reports of this investigation: Yes/No

Name _____________________________

Signature ___________________________ Date ________________

This study is approved by the University of Auckland Human Participants Ethics Committee (UAHPEC) 12 Dec 2012 to 12 Dec 2015, Reference number: 8870.
APPENDIX 7: Information sheet - Adults with hearing loss

General information

- Name (optional):
- Age: ..........years ..........months
- Date of Birth:
- Sex: Male/Female
- Languages spoken:
- Details of musical training
- Address:
- Phone: 
- Email:

Details about hearing loss:

- Age at which hearing loss was identified: .............years .......... months
- Type of hearing loss:
  Right ear: Conductive/Sensorineural/Mixed hearing loss
  Left ear: Conductive/Sensorineural/Mixed hearing loss
- Degree of hearing loss:
  Right Ear: Minimal/Mild/Moderate/Moderately severe/Severe/Profound hearing loss
  Left Ear: Minimal/Mild/Moderate/Moderately severe/Severe/Profound hearing loss
Details about hearing aids:

- Age at which you started wearing hearing aids: .......... years ........months
- Ear: Right ear                  Left ear                  Both ears
- Style: BTE /ITE/ITC/CIC
- Manufacturer: Phonak/Widex/Oticon       Others:

Details about cochlear implants:

- Age at which you started using Cochlear Implants:
- Ear: Right ear                  Left ear                  Both ears
- Style: Body worn Processor/ Ear level processor
- Manufacturer: Advanced Bionics/ Cochlear/ Med-El       Others:

Details about speech language therapy /aural rehabilitation:

- Did you receive speech language therapy services?  Yes/ No
  If Yes, for how long: ..................Years ................. months
- Did you receive aural rehabilitation services?    Yes/ No
  If Yes, for how long: ..................Years ................. months

Name: ..................                  Date / /

Signature:
APPENDIX 8: Participant information sheet - Adults

Study title: Perception of speech prosody in adults using cochlear implants

This study is being undertaken as a part of a PhD programme in Speech Science in the School of Psychology, The University of Auckland by Rose Thomas Kalathottukaren. The research project will be guided by Professor Suzanne C. Purdy & Dr. Elaine Ballard and the final report will be reviewed by the University examiners to mark the completion of the project. The proposed research will be completed by May 1st 2015.

The current study requires adult participants aged 18 years or older with hearing loss to be a part of this study. We are interested in “prosody” aspects of speech because these are important to understand the emotions and meaning of the speaker. Speech prosody depends on “suprasegmental” cues such as pitch and loudness and timing. Individuals with hearing loss may have trouble hearing these cues that can convey different meanings and emotions, even when the words in the sentence are the same. Speech prosody plays a major role in day to day communication.

The participants would be asked to complete an information sheet regarding their hearing status and use of hearing instruments.

If you wish to participate, we would like you to do the following tasks:

1. The PEPS-C will be used to assess your understanding of prosody. You will listen to words and sentences from the computer (through a loudspeaker) and you would need to respond by pointing on to the computer screen (30 mins)

2. The DANVA 2 (Diagnostic Analysis of Non-Verbal Accuracy 2) test will be used to assess your understanding of emotions in speech. The “Adult Paralanguage subtest” involves listening to sentences through loudspeaker and you will respond verbally/pointing to pictures to say if the person sounds happy, angry, or a different emotion (10 mins)

3. Pitch discrimination tasks music tones: Participants would listen to series of tones/music and would respond verbally to indicate whether they could hear a difference in pitch or not (10 mins)
The testing will take place at the University of Auckland Speech Language Therapy Clinic or at participant’s home, whichever is convenient. Participants will receive $10 gift voucher for their participation. A $20 petrol vouchers will be provided to contribute to travel cost if necessary.

Confidentiality:

Confidentiality will be preserved throughout the research process. If the information collected are reported or published, it will be done in a way that does not reveal the identity of the participant/clinic as the source of information will be kept confidential. Please be assured that your participation is completely voluntary, and that your decision either way will have no impact on the University of Auckland’s relationship with your clinic.

I have attached a copy of the participant information sheet and consent form for your information. If you wish to participate please return the consent forms to the researchers by an agreed date. I will then contact you to arrange a suitable time and location for testing. I will contact you by telephone with in the next two weeks to discuss this further. If you have any questions in meantime, please feel free to contact any of the researchers using the email addresses or telephone numbers given at the end.

Yours sincerely,

Rose Thomas Kalathottukaren  
Audiologist & Speech Language Pathologist/ PhD Student  
School of Psychology  
Tāmaki Innovation Campus  
The University of Auckland  
Phone +64 9 373 7599 ext. 84930  
rkal180@aucklanduni.ac.nz

Professor Suzanne C. Purdy  
Head of Discipline of Speech Science  
School of Psychology  
Tāmaki Innovation Campus,  
The University of Auckland  
Phone: 373 7599 ext. 82073 or 86886  
Email: sc.purdy@auckland.ac.nz

Dr. Elaine Ballard  
Senior Lecturer in Speech Science  
School of Psychology  
Tāmaki Innovation Campus  
The University of Auckland  
Phone: 373 7599 ext. 87502  
Email: e.ballard@auckland.ac.nz
For any queries regarding ethical concerns please contact:
The Chair,
The University of Auckland Human Participants Ethics Committee
The University of Auckland
Research Office-Office of the Vice Chancellor
Private Bag 92019, Auckland
Tel. (09) 373 7599 ext. 87830

This study is approved by the University of Auckland Human Participants Ethics Committee (UAHPEC) 13- Mar-2013 to 12-Dec- 2015, Reference number: 8870
APPENDIX 9: Consent form - Adults

Study title: Perception of speech prosody in adults using cochlear implants

Names of Researchers: Professor Suzanne C. Purdy (Head of Speech Science, The University of Auckland); Dr Elaine Ballard (Senior Lecturer, Speech Science, The University of Auckland); Ms Rose Thomas Kalathottukaren (Audiologist & Speech Language Pathologist/PhD Student, The University of Auckland)

• I have read the Participant Information Sheet and have understood the nature of the research designed to investigate speech perception in adults with hearing loss.

• I have had the opportunity to discuss the study, ask questions and have had them answered to my satisfaction.

• I agree that the researchers may access my clinical records for the purpose of this study, to get information about my hearing.

• I understand that my participation/nonparticipation in this study will not affect my future support from audiology/speech language therapy clinic.

• I understand that I might need to spend 1 session lasting for about 1 hour with the researcher.

• I understand that I am free to withdraw participation at any time and that I will be free to withdraw my data up to 3 months after the date on which I was tested.

• I understand that if the findings are reported or published my identity will be kept confidential.

• I understand that data will be kept anonymous and stored for a period of 6 years, after which they will be destroyed.

• I know whom to contact if I have queries regarding the study.

Name ___________________________

Signature ___________________________ Date ________________

This study is approved by the University of Auckland Human Participants Ethics Committee (UAHPEC), 13- Mar-2013 to 12-Dec- 2015, Reference number: 8870.
References


Cullington, H. E., & Zeng, F. G. (2010). Comparison of bimodal and bilateral cochlear implant users on speech recognition with competing talker, music perception,


(SLI) and/or dyslexia. *International Journal of Language and Communication Disorders, 44*, 466-488.


Shriberg, L. D., Paul, R., McSweeney, J. L., Klin, A. M., Cohen, D. J., & Volkmar, F. R. (2001). Speech and prosody characteristics of adolescents and adults with high-


