

School of Psychology and Speech Pathology

**Phonological and Speech Motor Abilities in Children with Childhood Apraxia
of Speech and Phonological Disorder**

Anne Marie Lewis

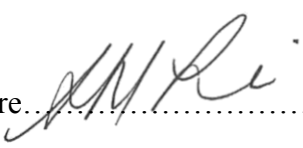
**This thesis is presented for the Degree of
Doctor of Philosophy
of
Curtin University**

May 2018

DECLARATION

To the best of my knowledge and belief this thesis contains no material previously published by any other person except where due acknowledgement has been made.

This thesis contains no material that has been accepted for the award of any other degree or diploma in any university.

Signature.....

Date.....20/01/2019

ACKNOWLEDGEMENTS

I would like to acknowledge numerous people for their support throughout this process. First and foremost, I would like to thank the children and families that participated in my studies, without whom there would be no research paper to write.

I would also like to acknowledge my supervisor, Dr. Neville Hennessey for his continual support and guidance in writing this dissertation. It has been a lengthy process and his unwavering support has been invaluable. His attention to detail and thoroughness has been resolute throughout this process. I cannot stress how grateful I am to him for supporting me through the numerous interruptions that I encountered at various times throughout this period. Additional thanks to Dr. Robert Kane for providing guidance and support for the very complex statistical analysis undertaken as part of this research.

I am very grateful to the speech pathologists that assisted in the recruitment of the children with CAS and PD, and for the teachers that helped with the recruitment of children with typical development. I would also like to thank Kym Mathews for her assistance establishing inter-rater reliability.

To my friends thank you. Despite this PhD taking significantly longer than had initially planned my friends have continued to provide support and words of encouragement when needed most. At no time did they ever give up on me and for that I am very grateful.

Last but not least I would like to acknowledge my family. They have been so patient and tolerant of all the time this project has demanded over a significant number of years. Thank you for your support throughout this time and for providing words of encouragement when needed. My husband has provided practical and emotional support throughout this time and for that I am very grateful. He has never

wavered in his belief that I will finish and has helped me through some very trying times when I felt like giving up. My children have allowed me to pursue my goals without judgement or consequence. I have endeavored to fulfill my role as a mother whilst on this journey but no doubt have failed at times, so thank you to my children for understanding how important this work is to me and for being so tolerant during this time. My parents instilled in me the absolute belief that I could do and be whatever I wanted and that there was no limit. My determination, resilience and persistence are a tribute to them. Growing up in a loving family and supported by three older sisters has given me the self-assurance I needed to complete this project and for that I am eternally grateful.

ABSTRACT

Childhood apraxia of speech (CAS) is a speech sound disorder (SSD) of unknown etiology. It is predominantly perceived to be a disorder in speech motor planning/programming disorder and for that reason distinct from other common forms of SSD, such as phonological disorder (PD) (ASHA, 2007). Children with CAS present with a broad range of speech deficits, which results in long term detrimental impact on both academic and social outcomes (Lewis, Freebairn, & Taylor, 2000). Although there is no generally accepted diagnostic criteria, children with CAS are predominantly diagnosed using a feature-based approach, focusing on features such as a limited phoneme repertoire, inconsistent speech errors and sequencing problems, that signal apraxic-type deficits. However, most of these features are not specific to CAS and can occur in idiopathic SSD, such as PD. Furthermore, the higher-level phonological-linguistic deficits seen in children with CAS are similar between these two disorders

It has been long observed, therefore, that there is little empirical evidence that differentiates between CAS and PD, questioning whether in fact CAS and PD are distinct disorders with different underlying etiologies. This thesis tackles this question by investigating markers of developmental constraints in the speech and language system of children with CAS and PD. In particular, our focus was to determine whether the speech and phonological deficits observed in CAS and PD arise as a result of different processing constraints suggestive of distinct causes with the developing speech and language system. Consistent with the dynamic nature of speech and language development, we hypothesized that if CAS arises from a core deficit in the development speech motor system, that also constrains higher order linguistic development, and PD arises from an underlying deficit in phonological processing, then measures of speech motor development would predict measures of phonological competence in children with CAS to a great degree than children with PD. Such a difference in this predictive relationship would indicate different constraints on development consistent with different causal origins.

In Study 1 tasks were designed to evaluate different aspects of phonological competence and speech motor ability and then piloted on 23 younger ($M = 67.1$ months) and 24 older children ($M = 96.7$ months) with typical development (TD), to ascertain if the methodology was suitable and whether the measures were valid

indicators of developmental change in speech motor and/or phonological competence. The findings showed that measures from a speech discrimination task and nonword repetition task were valid indicators of developmental change in phonological competence. A simple verbal reaction time task, targeting execution of speech motor plans, was found to be sensitive to developmental change and was the only measures of speech motor ability that predicted phonological competence, specifically the development of input phonological, while controlling for other factors such as vocabulary size.

Prior to undertaking Study 2 with children with CAS, PD and TD, a systematic review of the protocols used by researchers to classify children with CAS was undertaken. This review showed researchers across two decades from 1992 applied a wide range of CAS related features, although a small number of those features were highly prevalent. In general, these CAS related features were poorly operationalized. A classification protocol for CAS was subsequently developed that operationalized the most prevalent features identified in the review. Using clinically ascertained children with CAS, PD and TD (the same participants in Study 2), an exploratory factor analysis examined the underlying latent structure of the target CAS related features. This resulted in a one-factor solution, with loadings that clearly separated the three groups, consistent with CAS being a unidimensional praxis type deficit. This classification protocol was used with discriminant function analysis to arrive at a final allocation of the children into the CAS and PD groupings for Study 2.

In Study 2, the children with CAS ($n = 14$) and PD ($n = 22$) shared a number of deficits at the level of phonological competence compared to the TD children ($n = 18$), although unexpectedly there were little differences between the groups in the revised measure of output phonology using phonological priming during picture naming. As expected, the children with CAS had more severe speech motor deficits. Hierarchical moderator regression analysis was then used to determine if speech motor measures predicted measures of phonological competence. Overall the regression analyses failed to show differences between groups in the degree to which the measures of speech motor ability predicted the measures of phonological competence, while controlling for other factors such as age and vocabulary. A delayed picture naming reaction time task targeting execution of speech motor plans in Study 2 showed a significant interaction with group for picture naming reaction

time. This was supported by the positive correlation between these measures for the children with PD and TD, but not for the children with CAS, indicating a clear association for the children with PD and TD and a clear dissociation for the children with CAS. There were a number of additional associations that emerged from the correlation analysis that differentiated between the children with CAS and the children with PD and TD. Consistent with the dynamic and interactive nature of speech and language development, the present study provides some preliminary evidence that different patterns or associations can emerge in the developing system, despite overlapping symptoms, and be indicative of different underlying etiologies. The present study shows further research is warranted to investigate the associations and dissociations that potentially differentiate between the causal origins in CAS and PD.

LIST OF TABLES

Table 1	<i>Age and Test Scores for Younger and Older Children (N = 47)</i>	50
Table 2	<i>Nonword Repetition Mean Percent Accuracy Scores and Standard Error of the Mean for Younger and Older Children for Low and High Phonotactic Frequency Sequences for 2 and 3 Syllable Nonwords (N = 47)</i>	60
Table 3	<i>Picture-naming Mean Reaction Time (PNrt) in Milliseconds and Percent Correct (%) with Standard Errors for Younger and Older Children for Unrelated, Related and Silence Priming Conditions (N = 47)</i>	62
Table 4	<i>Speech Discrimination Ability as Measured by d-prime for Younger and Older Children for Word Pairs, Cat/cap and Tap/tack in the Whole Word, Short Gate and Long Gate Conditions (with Percent Correct in Brackets) (N = 47)</i>	64
Table 5	<i>Mean Reaction Time in Milliseconds (ms) with Standard Error of the Mean (SEM) for the Simple Verbal Reaction Time for Younger and Older Children (N = 47)</i>	65
Table 6	<i>Mean Reaction Time in Milliseconds (ms) with Standard Error (SEM) of the Mean for Choice Verbal Reaction Time for Younger and Older Children (N = 47)</i>	66
Table 7	<i>Pearson Correlations Between Phonological Measures and Speech Motor Measures for All Participants (N = 47).</i>	71
Table 8	<i>Pearson Correlations Between Outcome Measures and Measures of Expressive and Receptive Vocabulary for the Whole Group (N = 47)</i>	72

Table 9	<i>Unstandardized (B) and Standardized (β) Regression Coefficients, and Squared Semi-partial Correlations (sr^2) for a Hierarchical Multiple Regression Model Predicting the Mean Nonword Repetition Accuracy for the Low Frequency Sequences from GFTA Raw Scores, Chronological Age, and the Chronological Age x GFTA Interaction (N =47)</i>	73
Table 10	<i>Unstandardized (B) and Standardized (β) Regression Coefficients, and Squared Semi-partial Correlations (sr^2) for a Hierarchical Multiple Regression Model Predicting the Phonotactic Frequency Effect from GFTA Raw Scores, Chronological Age (ChronAge), and the Chronological Age x GFTA Interaction (N = 47)</i>	75
Table 11	<i>Unstandardized (B) and Standardized (β) Regression Coefficients, and Squared Semi-partial Correlations (sr^2) for a Hierarchical Multiple Regression Model Predicting Picture-naming Reaction Time from GFTA Raw Scores, Chronological Age, and the Chronological Age x GFTA Interaction (N = 47)</i>	77
Table 12	<i>Unstandardized (B) and Standardized (β) Regression Coefficients, and Squared Semi-partial Correlations (sr^2) for a Hierarchical Multiple Regression Model Predicting Picture-naming Reaction Time from Simple Verbal Reaction Time, Chronological Age (ChronAge), and the Chronological Age x Simple Verbal Reaction Time Interaction (N = 47)</i>	78
Table 13	<i>Unstandardized (B) and Standardized (β) Regression Coefficients, and Squared Semi-partial Correlations (sr^2) for a Hierarchical Multiple Regression Model Predicting</i>	79

Picture-naming Reaction Time (PNrt) from Choice Verbal Reaction Time, Chronological Age (ChronAge), and the Chronological Age x Choice Verbal Reaction Time Interaction (N = 47)

Table 14	<i>Unstandardized (B) and Standardized (β) Regression Coefficients, and Squared Part Correlations (sr^2) for a Hierarchical Multiple Regression Model Predicting d-prime from Simple Verbal Reaction Time, Chronological Age, and the Chronological Age x SVRT Interaction (N = 47).</i>	81
Table 15	<i>Unstandardized (B) and Standardized (β) Regression Coefficients, and Squared Semi-partial Correlations (sr^2) for a Hierarchical Multiple Regression Model Predicting d-prime from Choice Verbal Reaction Time, Chronological Age (ChronAge), and the ChronAge x Choice Verbal Reaction Time Interaction (N = 47)</i>	82
Table 16	<i>Description of Final 18 CAS Features Based on Systematic Review of CAS Research Publications from 1993-2013</i>	103
Table 17	<i>Features Used to Classify Children as CAS and Number and Percentage of Papers Citing Each Feature</i>	104
Table 18	<i>Final CAS Features that included Operational Definitions</i>	106
Table 19	<i>Features Consistent with CAS Used in Research Literature During the First Period (1993 to 1999), Second Period (2000 to 2006) and Third Period (2007 to 2013).</i>	108
Table 20	<i>Demographic Data and Test Scores for Children with CAS, PD and TD with Standard Deviations (SD) (53)</i>	120

Table 21	<i>Related Features of CAS for Children with CAS, PD and TD with Standard Deviations (SD) (N = 53)</i>	126
Table 22	<i>Related Features of CAS in Children with CAS and PD (N = 35)</i>	131
Table 23	<i>Correlation Matrix Between Feature Variables of CAS (N = 53).</i>	132
Table 24	<i>Factor Loading for Features of CAS Loading Onto a One-Factor Solution (N = 53).</i>	133
Table 25	<i>Factor Loading for Features of CAS Loading Onto a One-Factor Solution Excluding ICD (N = 53).</i>	134
Table 26	<i>Discriminate Function Analysis Resulting in Reassigned Group Membership (N = 53)</i>	135
Table 27	<i>Sensitivity and Specificity Values for Operationalized Features of CAS (N= 35)</i>	136
Table 28	<i>Demographic Data and Test Scores for Children with CAS, PD and TD with Standard Deviations (SD) (N = 53)</i>	152
Table 29	<i>Summary of Outcome Measures for Tasks Evaluating Phonological Competency and Speech Motor Ability</i>	157
Table 30	<i>NWR Mean Accuracy Scores (%) and Standard Error of the Mean for Children with CAS, PD and TD for Low and High Frequency Sequences for 2 Syllable and 3 Syllable Nonwords (N = 53)</i>	159
Table 31	<i>Picture-naming Mean Reaction Time for Children with CAS,</i>	160

*PD and TD for Unrelated, Related and Silence Prime
Conditions for Short and Long SOAs (N = 53)*

Table 32	<i>Picture-naming Mean Accuracy (%) for Children with CAS, PD and TD for Related, Unrelated and Silence Priming Conditions for Short and Long SOAs (N = 53)</i>	161
Table 33	<i>Speech Discrimination Ability as Measured by D-prime for Children with CAS, PD and TD for Whole Word, Short Gate and Long Gate and for Word Pairs, Cat/Cap and Tap/Tack (N = 47)</i>	163
Table 34	<i>Delayed Picture-naming Mean Reaction Time for Children with CAS, PD and TD for Low and High Frequency Words and Short and Long ISI's (N = 47)</i>	165
Table 35	<i>Pearson Correlations Between Measures of Phonological Competency for All Participants (N = 53).</i>	169
Table 36	<i>Pearson Correlations Between Measures of Phonological Competency for the Children with CAS (n = 14), PD (n = 21) and TD (n = 18).</i>	170
Table 37	<i>Correlations Between Outcome Measures, Measures of Vocabulary and Articulation and Predictors for the Whole Group (N = 53).</i>	171
Table 38	<i>Correlations Between Outcome Measures, Measures of Vocabulary and Articulation and Predictors for Children with CAS (N = 14)</i>	172

Table 39	<i>Correlations Between Outcome Measures, Measures of Vocabulary and Articulation and Predictors for Children with PD (N = 21)</i>	173
Table 40	<i>Correlations Between Outcome Measures, Measures of Vocabulary and Articulation and Predictors for Children with TD (N = 18)</i>	174
Table 41	<i>Unstandardized (B) and Standardized (β) Regression Coefficients, and Squared Part Correlations (sr^2) for a Hierarchical Multiple Regression Model Predicting NWR Mean Accuracy for Low Frequency Sequences from GFTA Raw Scores, Group, and the Group x GFTA Interaction (N = 53)</i>	176
Table 42	<i>Unstandardized (B) and Standardized (β) Regression Coefficients, and Squared Part Correlations (sr^2) for a Hierarchical Multiple Regression Model Predicting NWR Mean Accuracy for Low Frequency Sequences from CAS Trait Score (CASts), Group, and the Group x CASts Interaction (N = 53)</i>	177
Table 43	<i>Unstandardized (B) and Standardized (β) Regression Coefficients, and Squared Part Correlations (sr^2) for a Hierarchical Multiple Regression Model Predicting NWR Mean Accuracy for Low Frequency Sequences from DPNrt (DPNrt), Group, and the Group x DPNrt Interaction (N = 53)</i>	179

Table 44	<i>Unstandardized (B) and Standardized (β) Regression Coefficients, and Squared Part Correlations (sr^2) for a Hierarchical Multiple Regression Model Predicting the Phonotactic Frequency Effect from GFTA Raw Scores, Group, and the Group x GFTA Interaction (N = 53)</i>	180
Table 45	<i>Unstandardized (B) and Standardized (β) Regression Coefficients, and Squared Part Correlations (sr^2) for a Hierarchical Multiple Regression Model Predicting the Phonotactic Frequency Effect from CAS Trait Scores, Group, and the Group x CAS Trait Score Interaction (N = 53).</i>	181
Table 46	<i>Unstandardized (B) and Standardized (β) Regression Coefficients, and Squared Part Correlations (sr^2) for a Hierarchical Multiple Regression Model Predicting Picture-naming Reaction Time from CAS Trait Score (CAS_{Ts}), Group, and the Group x CAS_{Ts} Interaction (N = 53)</i>	183
Table 47	<i>Unstandardized (B) and Standardized (β) Regression Coefficients, and Squared Part Correlations (sr^2) for a Hierarchical Multiple Regression Model Predicting Picture-naming Reaction Time from DPN_{rt}, Group, and the Group x DPN_{rt} Interaction (N = 53)</i>	184
Table 48	<i>Unstandardized (B) and Standardized (β) Regression Coefficients, and Squared Part Correlations (sr^2) for a Hierarchical Multiple Regression Model Predicting Speech Discrimination Ability (D-prime) from GFTA Raw Scores, Group, and the Group x GFTA Interaction (N = 53)</i>	186
Table 49	<i>Unstandardized (B) and Standardized (β) Regression Coefficients, and Squared Part Correlations (sr^2) for a</i>	187

*Hierarchical Multiple Regression Model Predicting Speech
Discrimination Ability (D' -prime) from CAS Trait Score
(CAS_Ts), Group, and the Group \times CAS_Ts Interaction ($N = 53$)*

TABLE OF CONTENTS

DECLARATION	2
ACKNOWLEDGEMENTS.....	3
ABSTRACT	5
LIST OF TABLES.....	8
Chapter 1	20
Introduction	21
Speech Sound Disorders	23
Childhood Apraxia of Speech	24
Phonological Disorder	27
Models of Speech Production.....	28
WEAVER Model	29
Deficits in CAS	32
Speech Motor Deficits	34
Linguistic/Phonological Deficits.....	41
Speech Input and Output	45
Limitations in Research in CAS	50
Speech Development.....	51
A Developmental Account of CAS.....	54
Computational Modelling of CAS.....	56
Research Aims and Rationale	58
Chapter 2	60
Introduction	61
Evaluating Phonological Competence	61
Evaluating Speech Motor Ability	63
Hypotheses	66
Method	68
Participants	68
Standardized Assessments	70
Stimulus Materials.....	71
Procedure.....	73
Scoring of Dependent Measures	76

Data Analysis	77
Results	78
Nonword Repetition	78
Picture-naming	80
Speech Discrimination	82
SVRT and CVRT	83
Regression Analysis	85
Predicting Nonword Repetition Accuracy	88
Predicting the Phonotactic Frequency Effect	92
Predicting Picture-naming Reaction Time	94
Predicting Speech Discrimination Ability	98
Discussion	101
Developmental Differences in Phonological Competency	102
Developmental Differences in Speech Motor Ability	105
Relationship Between Speech Motor Measures and Phonological Competence..	107
Limitations and Future Directions.....	111
Conclusion.....	112
Chapter 3	114
Introduction	115
Aims and Rationale	116
Method	117
Types of Studies.....	118
Data Extraction and Management.....	118
Classification Protocols	119
Results	121
Features of CAS.....	121
Prevalence of CAS Classification Features.....	121
Number of Features	123
Operational Definitions	124
Single Subject Methodology.....	125
Alternative Approach For Classification	126
Changes to the features of CAS used in research over time	126
Discussion	128
Limitations and Future Directions.....	130
Conclusion.....	131

Chapter 4	133
Introduction	134
Aims and Rationale	135
Features of CAS.....	136
Method	137
Participants	137
Operationalized Features	140
Procedure.....	143
Inter-rater Reliability.....	143
Data Analysis	144
Results	144
Prevalence of Features Within Sample	146
Exploratory Factor Analysis	148
Final Classification Protocol.....	152
Sensitivity and Specificity.....	154
Discussion	155
Limitations and Future Directions.....	160
Conclusion.....	161
Chapter 5	162
Introduction	163
Phonological Competence	165
Speech Motor Ability	166
Hypotheses	167
Method	170
Participants	170
Stimulus Materials.....	172
Procedure.....	173
Scoring of Dependent Measures	175
Data Analysis	175
Results	176
Nonword Repetition	176
Picture-Naming.....	178
Speech Discrimination.....	181
Delayed Picture-naming.....	183
Regression Analysis	185
Predicting Nonword Repetition Accuracy.....	195

Predicting the Phonotactic Frequency Effect.....	198
Predicting Picture-Naming Reaction Time.....	203
Predicting Speech Discrimination Ability.....	206
Discussion.....	208
Phonological Deficits in CAS and PD	209
Relationship Between Speech Motor Measures and Phonological Competence..	214
Regression Analysis	216
Limitations and Future Directions.....	221
Summary and Conclusions.....	223
Chapter 6	225
Overview	226
General Summary of Findings	227
CAS and PD	229
Limitations and Future Directions.....	233
Conclusion.....	235
REFERENCES.....	240
Appendix A	257
Appendix B	267
APPENDIX C	343
APPENDIX D.....	345

Chapter 1

Introduction

Childhood apraxia of speech (CAS) is a speech sound disorder of unknown etiology. It impacts severely on speech intelligibility resulting in a long-term detrimental impact on both academic and social outcomes for children diagnosed as such. Children with CAS present with a broad range of deficits that vary dependent upon severity and stage of development. Diagnosis of CAS is predominantly based on the presence of features consistent with these deficits. Deficits include inconsistent speech errors, sequencing deficits, vowel errors, and prosodic disturbances, to name a few. However, the features that researchers and clinicians use for classification purposes are not only varied, but the number of features considered sufficient to warrant classification as CAS also differ. Consequently, the “feature based approach” as it is currently implemented to classify children as CAS has a number of limitations in terms of its reliability and consistency across clinicians and researchers. The American Speech and Hearing (ASHA) (2007) technical report on CAS highlights this diagnostic problem as the largest impediment to advancement in our knowledge and understanding of CAS. Consequently, the debate continues as to whether CAS is an independent speech sound disorder and researchers continue to search for specific markers that have the potential to differentiate CAS from other forms of speech sound disorders (SSD). CAS is predominantly perceived as a speech motor planning/programming disorder, giving rise to a broad range of speech deficits. However, some of these speech deficits are not specific to CAS and occur in idiopathic SSD, such as phonological disorder (PD). Idiopathic SSD relates to speech disorders that have an unknown etiology, in contrast to speech disorders with a known cause, such as speech disorders that arise as a result of fluctuating hearing loss due to repetitive ear infections. In contrast to children with CAS, children with PD are believed to have an underlying deficit at the linguistic level of processing, despite the number of shared deficits with CAS.

A developmental perspective is presented in this thesis where speech and language development is perceived as a dynamic process that involves the integration of speech input processing pathways and output processing pathways. From this perspective, the key focus for this thesis is whether the speech related deficits observed in CAS and PD, although similar, arise as a result of distinct processing constraints within the developing speech and language system. Our initial goal was to directly compare children with CAS and PD in relation to their

shared deficits. An argument will be presented that if the causal origins of CAS and PD are different, involving different constraints on the emergence of the speech production system, then evidence for differences in development can potentially be found in the relationships between measures targeting those interacting components. In a regression model, for example, different patterns of covariance between measures of interacting components should be reflective of the degree of constraint of one component over another. Consequently, our main goal was to determine whether patterns of covariance between the different levels of processing vary in children with CAS and PD, and children with typical development (TD). The regression analysis reported in this thesis tested specifically whether measures of speech motor control (assumed to be the core deficit affecting speech and language development for children with CAS, but not PD) predicted the development of higher-level/phonological processes to different degrees in children with SSD depending on their diagnosis (Study 2). It is argued that such differences in prediction are associated with different forms of constraint on the developing speech and language system, and, in view of this, has the potential to provide evidence that CAS and PD are distinct disorders with respect to their causal origin.

The primary purpose of Study 1 was to pilot the tasks planned for Study 2 with children of a similar age to develop the methodology for our main objective. Because the covariant relationships between some of the target measures have not been previously investigated, a developmental perspective was valuable to pilot the planned methodology. We also wanted to assess the validity of outcome measures as sensitive indicators of the specific levels of processing targeted, and developmental changes in those processes, by undertaking a comparison of younger and older children with TD within the target age range. Study 1, therefore, enabled us to explore the relationships between the different levels of processing during typical speech and language development and also pilot the methodology for Study 2.

In the process of undertaking the literature review on CAS it became apparent that there were a number of inconsistencies with regard to how children were classified as having CAS. This motivated a systematic review of classification protocols used in the research literature in CAS over a 20-year period from 1993 to 2013, presented in Chapter 3. This review resulted in a comprehensive analysis of features used for classification purposes, the number of features required to warrant classification as CAS and the identification of operationally defined features. The

most prevalent features were selected and a method of operationalizing those features developed and validated using the clinically diagnosed participant groups for Study 2 (see Chapter 4). The underlying dimensionality of those operationalized features was analysed and validated using an exploratory factor analysis, presented in Chapter 4. Based on these findings a protocol for classifying children with CAS, as distinct from PD, was developed to refine the participant groupings for Study 2, presented in Chapter 5.

The following sections of this introductory chapter examine speech sound disorders and the broad range of deficits observed in children with CAS, including the overlap of these deficits in children with PD. An overview of theoretical perspectives on speech production and speech development is provided and a theoretical framework is provided to assist in identifying the multiple levels of processing implicated in CAS and PD. In particular, theoretical accounts that acknowledge the interactive nature of speech and language development and highlight the complexity of speech development for children with speech sound disorders, such as CAS and PD are discussed. Finally, this chapter concludes with the specific aims and rationale for the research presented in Chapters 2 to 5.

Speech Sound Disorders

Children with speech sound disorders (SSD) of unknown etiology make up approximately 10% to 15% of preschoolers and 6% school age children (McLeod, 2009). These children are classified as such based on their inability to accurately produce the sounds of their native language. These children do not form a homogenous group, they differ in terms of severity, the types of errors they produce and their ability to resolve these difficulties (Bowen, 2009; Dodd, 2005; Stackhouse & Wells, 1997). There is agreement that the heterogeneity of this population is a problem with regard to accurate diagnosis and treatment management (Bowen, 2009; Dodd, 2005; Shriberg, Austin, Lewis, McSweeney, & Wilson, 1997; Stackhouse & Wells, 1997). Various classification systems have been developed with the aim to subtype children with SSD and in doing so provide more accurate diagnosis and treatment protocols. However, to date there is no consensus on a universally agreed upon classification system (Waring & Knight, 2013). Consequently, children with SSD include children with speech deficits that encompass mild to severe deficits, some of which are highly persistent and require on-going treatment. The majority of

cases of SSD are attributable to an unknown origin and children are typically diagnosed between 2 and 4 years of age (Gierut, 1998b).

Childhood Apraxia of Speech

CAS is a developmental SSD of unknown origin which can have genetic, neurological or idiopathic causes (Murray, McCabe, & Ballard, 2012). It is regarded as a highly heritable condition (Hall, Jordon, & Robin, 1993; Lewis, Freebairn, Hansen, Taylor, et al., 2004; Thoonen, Maassen, Gabreels, Schreuder, & deSwart, 1997). CAS can occur in isolation with unknown origin but it can also occur as a result of known neurological damage or impairment. CAS is usually contrasted with childhood dysarthria, although there is evidence that both can co-occur (Shriberg et al., 2006). Childhood dysarthria is a motor speech disorder in children arising from impairments to the speech related neuromuscular subsystems controlling articulation, phonation, respiration and nasal response. Consequently, the speech characteristics displayed by children with dysarthria result in slow speech rate, unclear speech and difficulty or inability to produce rapid sound sequences and for this reason can be confused with characteristics seen in children with CAS (Bradford, Murdoch, Thompson, & Stokes, 1997; Caruso & Strand, 1999; Kent, 2000; Kent & Kim, 2003; McNeil, Robin, & Schmidt, 1997), which highlights the complexity of diagnosis of CAS.

Children with CAS are also predisposed to language deficits and are at risk for persistent reading disorder (Lewis, Freebairn, Hansen, Iyengar, & Taylor, 2004; Moriarty & Gillon, 2006; Stackhouse & Snowling, 1992). The functional impact of CAS on children with this disorder results in life long deficits with research showing children with CAS demonstrate poorer academic outcomes as a result of poorer spelling and reading ability compared to typically developing peers (Lewis et al., 2000; Mc Neil, Gillon, & Dodd, 2009). Furthermore, some adults with a history of speech disorders, including CAS, have been reported to have lower socio-economic status compared to their non-speech disordered peers (Lewis, Freebairn, Hansen, Iyengar, et al., 2004).

Like most children with SSD of unknown etiology, children with CAS also form a highly heterogeneous group, differing in severity and speech error characteristics (Lewis et al., 2011; Waring & Knight, 2013). Moreover, the specific characteristics associated with the disorder vary with age and stage of development,

further confounding diagnosis and subsequent treatment efficacy (Pennington & Bishop, 2009). Furthermore, children with CAS often remain in therapy for extensive periods and make slow progress (Lewis, Freebairn, Hansen, Iyengar, et al., 2004). In addition, speech pathologists frequently report that once treatment stops these children show deterioration and/or loss of skills, such as loss of articulation ability (Forrest, 2003).

Children with CAS were originally classified as such based on our existing knowledge of acquired apraxia of speech (AOS). AOS is an acquired speech disorder resulting from lesions in the left hemisphere of the brain, which are often the result of an infarction of the left middle cerebral artery (Ziegler, Staiger, & Aichert, 2010). Apraxia of speech represents a disruption in the generation and production of speech plans (Jacks & Robin, 2013). It is characterized by phonetic sound distortions, dysfluency and dysprosody, which are assumed to reflect a motor planning deficit (Ziegler et al., 2010). The perceived similarity in the characteristics of AOS and CAS resulted in the early assumption that CAS is a developmental version of the same type of planning deficit (Hall et al., 1993). However, in contrast to the comprehensive research in AOS research, interest only began in CAS in the early 70's, gaining real momentum in the 90's, with the bulk of research in this area taking place in the last two decades (Shriberg, Lohmeier, Strand, & Jakielski, 2012). Consistent with the research on CAS, the most recognized impediment to theoretical and clinical advancement in AOS is the lack of a comprehensive and clear definition of this disorder (McNeil, Pratt, & Fossett, 2004). Despite the similarities between AOS and CAS there is a fundamental problem comparing the nature of an acquired disorder with that of a developmental disorder. With an acquired disorder, assuming it occurs in adulthood, the speech and language system is already established, unlike that of a developing system. Deficits that occur in acquired disorders, such as AOS are more likely to be localized, whereas the deficits in developmental disorders, such as CAS, are more widespread and can impact on the entire speech and language system (Bishop, 1997).

Historically, CAS has been a controversial disorder due to the longstanding debate over its existence as a separate entity. Guyette and Diedrich (1981, p. 39) summed up CAS in the early 80's by defining it as a "label in search of a population" based on the view that deficits observed in children with CAS were also observed in children with SSD and therefore not unique to a specific group. Some researchers

proposed that the speech deficits present in children with CAS were related to a higher-level linguistic deficit (Marion, Sussman, & Marquardt, 1993; Marquardt, Jacks, & Davis, 2004; Marquardt, Sussman, Snow, & Jacks, 2002), while others argued for a lower level speech motor deficit (Lewis, Freebairn, Hansen, Iyengar, et al., 2004; Maassen, Nijland, & Van der Meulen, 2001; Nijland, Maassen, & van der Meulen, 2003; Nijland, Terband, & Maassen, 2015; Thoonen, Maassen, Gabreels, & Schreuder, 1999). CAS remains controversial with regard to etiology, clinical manifestations and treatment and the proposition that higher level linguistic deficits are more consistent with idiopathic SSD, such as PD, than a speech motor disorder (Davis, Jakielski, & Marquardt, 1998). Despite the numerous attempts to find reliable diagnostic markers that differentiate CAS from other developmental speech sound disorders the debate continues with regard to the clinical characteristics associated with this disorder and the precise etiology of CAS (Nijland et al., 2015).

This diagnostic uncertainty and variability in classification protocols used for identifying children with CAS has impeded our advancement in understanding this complex disorder. Although there is some agreement with regard to the most pertinent features consistent with CAS, there is ongoing variability between clinicians and researchers alike regarding these features and the number of features that are considered necessary to warrant diagnosis as such. Significant inconsistencies have been identified regarding the features used for classification purposes (Forrest, 2003; McCabe, Rosenthal, & McLeod, 1998). McCabe et al. (1998) identified 30 features consistent with a diagnosis of CAS in a cohort of 50 children (mean age of 5.10) with articulation and/or phonological impairment, with the number of features varying from 4 to 23 per child. Similarly, Forrest (2003) identified 50 features considered consistent with a diagnosis of CAS in a survey of 75 clinicians, the application of which were highly variable with regard to the features considered most pertinent and the number of features regarded necessary to warrant diagnosis. Given the weight of clinical judgment in the diagnosis of CAS, both clinically and in research, this uncertainty poses a significant problem. Sufficiently so that the technical report on CAS by American Speech-Language and Hearing Association (ASHA) has highlighted the uncertainty with regard to diagnosis among clinicians and researchers alike as the primary barrier to research in CAS (ASHA, 2007).

Despite these limitations a number of advancements have been made with regard to our understanding of CAS with ASHA acknowledging that CAS is a symptom complex, rather than a unitary disorder, consistent with the proposal that a number of core features are required for diagnosis (ASHA, 2007). ASHA also recognizes that symptoms of CAS can change over time and can vary across children and within the same child (2007). ASHA (2007) therefore emphasizes the importance of a more stringent methodological protocol in research and highlights the importance of the inclusion of additional experimental groups that include, not only children with typical development, but also more importantly children with non-apraxic speech sound disorders (ASHA, 2007).

Phonological Disorder

Phonological disorder (PD) is another speech sound disorder of unknown etiology (Gierut, 1998b). PD was the generic term initially used to classify children with a speech sound disorder, consequently, anyone that presented with speech errors could be a candidate for having a phonological disorder (Shriberg & Kwiatkowski, 1982). Children with PD have multiple speech errors despite having normal hearing, intelligence, social and emotional social skills. Traditionally speech errors observed in children with PD were identified using linguistic descriptions such as omissions, substitutions, distortions or additions. Substitutions and omissions were interpreted as a deficit at the phonological level. In contrast, phonetic errors, such as distortions and additions, were interpreted as a deficit at the level of speech motor planning, consistent with a speech motor deficit. This resulted in the general acceptance of two classifications of children with speech sound production errors, phonological disorder, consistent with a higher level phonological deficit and articulation disorder, consistent with a lower level speech motor deficit, which included children with CAS.

Children with PD are typically diagnosed using a standardized test, such as the GFTA, indicating a significant gap in speech sound development for the child's age, with no known cause (such as hearing impairment) and where there is no evidence of motoric involvement which might suggest a speech motor disorder. Children with PD are often able to produce sounds with instruction, suggesting that the articulatory abilities are in place for these sounds (Gibbon, 2002; Miccio, Elbert, & Forrest, 1999; Rvachew, 2005). This is consistent with the view that the sounds in error are

not related to the speech motor skill but how speech sounds are stored in the child's phonological system, a component of their language ability. However, despite the view that the main underlying deficit in children with PD is phonological in nature, a number of studies have also revealed speech motor deficits in some children with PD (Bradford & Dodd, 1994, 1996; Dodd, Holm, Crosbie, & McCormack, 2005; Forrest & Elbert, 2001; Gibbon, 1999; Gierut, 1998a).

Similar to CAS, severity can vary in PD ranging from a mild disorder to a severe speech disorder. Children with a severe speech sound disorder tend to present with additional speech and language deficits, whereas children with a mild SSD present with fewer speech and language deficits (Stackhouse & Wells, 1993). Furthermore, children with PD have demonstrated deficits in phonological awareness (Gierut, 1998b) and decreased perceptual knowledge of phonological structure of language relative to their typically developing peers (Edwards, Fox, & Rogers, 2002; Jamieson & Rvachew, 1992; Locke, 1980; Rvachew & Jamieson, 1989). Consequently, these children are predisposed to problems with literacy acquisition, which in turn can impact on academic success (Hesketh, Adams, Nightingale, & Hall, 2000; Larrivee & Catts, 1999).

Models of Speech Production

Theoretical models, such as those proposed by Dell et al (1997) and Levelt and colleagues (Levelt, 1989; Roelofs, 1997b), have been developed as models of the adult or fully developed speech production system. However, they also provide a useful theoretical network to understand where breakdowns or deficits may occur in various speech and language disorders, including speech sound disorders, and have been used for this purpose in the research literature (Maassen, 2002; Ziegler & Maassen, 2004). These models propose that the speech production system is composed of separate levels of processing with specific tasks assigned to each of the levels. Levelt (1989; Levelt et al., 1999) proposes that speech is generated predominantly in a serial order, with tasks completed at higher levels before processing can begin at lower levels. The core principle of this model is that at each level a unit is selected only when it reaches activation threshold, with the most activated unit being selected for use. When units are selected at the level of phonological encoding a phonetic plan is then produced, which is then transferred to the articulatory module for execution. Van der Merwe's (1997) model, distinguishes

three components at the level of speech output; speech motor planning, speech motor programming and speech motor execution, which occur in a cascading serial order. The DIVA model (Guenther, 1994; Guenther & Perkell, 2004; Tourville & Guenther, 2011), a neural network model and Ozanne's (1995, 2005) psycholinguistic model, both specifically designed to explore speech deficits, in contrast to the WEAVER model, will be implemented as a means of identifying differences between deficits observed in CAS and PD. Maassen's (2002) developmental model will also be included to explore speech deficits observed in CAS from a developmental perspective.

WEAVER Model

The WEAVER (Word-Form Encoding by Activation and VERification) is an extension on Levelt's (1989) model. The WEAVER model (Roelofs, 1997b) was conceived as a network of nodes whereby unit selection was based on spreading activation during lexical selection and phonetic encoding. The verification process ensures the accuracy of the unit selected based on whether it is linked to the preceding level of representation (Roelofs, 1997b). An adaptation of the model is depicted in Figure 1. According to the WEAVER model there are three distinct stages in the network; conceptualisation, formulation and articulation. Conceptualization involves selecting the appropriate lexical concepts to convey the intended meaning and entails lemma activation and selection. When this stage has been completed, the next stage begins. Formulation comprises two steps; morphological encoding and phonological encoding, which occur in tandem. Morphological encoding involves retrieval of the morphological codes of the selected lemma, whereby the relevant morphemes are retrieved, whereas phonological encoding is the spelling out of the sounds of a morpheme including retrieving, ordering and organizing phonemes. Phonological encoding also encompasses metrical spellout and segmental spellout, which occur simultaneously. During metrical spellout successive syllables are allocated to a given position within a prosodic frame and stress is assigned to particular syllables within that frame, following the rules of the speaker's native language (Levelt et. al, 1999). Segmental spellout involves the retrieval of the individual segments that make up the target word form stored within the lexicon and these segments are slotted into the metrical syllabic frame, resulting in the formation of phonological syllables. The

phonological encoding in the WEAVER model consists of inserting phonemes into syllabic frames to create the phonological representation of the word (Levelt, Roelofs, & Meyer, 1999). Deficits that occur at this stage of processing are indicative of a higher-level impairment in speech production, consistent with a phonological deficit.

Phonetic encoding translates the phonological representation into a context dependent phonetic representation, a phonetic plan (Roelofs, 1997a). This involves retrieval of syllable programs from the mental syllabary. The mental syllabary is a composite store of high frequency or established syllable programs that have become learned over the course of speech and language development (Levelt et al., 1999). These ‘ready made’ motor programs are consistent with Browman and Goldstein’s (1992) gestural scores, which are defined as “abstract characterizations of articulatory events, each with an intrinsic time or duration” (Browman & Goldstein, 1988; 1992, p. 155; 1997). Problems that occur during this stage of processing are considered to reflect a lower level, speech motor deficit specifically at the level of phonetic planning, consistent with some deficits observed in CAS.

The final stage, articulation, takes the phonetic motor plan, which according to Levelt et al. (1999), specifies the articulatory gestures including segmental and prosodic information, and translates that into muscle commands for execution. Prior to execution, the phonetic plan is temporarily stored in the articulatory buffer as gestures may not be retrieved at the same rate as execution, and this temporary storage permits a fluent and constant rate of speaking. To ensure the articulatory goal is achieved the speech system must combine segmental and suprasegmental features and auditory feedback guides the adjustments for loudness, prosody and rate, in addition to other cues, such as speaker familiarity with the subject matter, ambient noise, etc. (Roelofs, 1997b). However, the mechanisms that alter suprasegmental features can have a simultaneous effect on segmental features. For example, the mechanism that alters loudness or amplitude can have a simultaneous impact on presence or absence of voicing thus altering the sound produced (Perkell et al., 2000). Therefore, the speech motor control system must take into consideration this interaction when adjusting parameter settings so the integration of both sensory and motor information is critical to ensure articulatory goals are maintained (Perkell et al., 2000). Problems that occur at this stage of processing, in the absence of any physical deficit such as poor muscular control or weakness (e.g.,

flaccid dysarthria) are indicative of a speech motor control deficit. This specifically relates to the level of speech motor planning and programming that precede the neuromuscular system producing overt speech related movements.

Speech deficits in CAS have been attributed to the level of both speech planning and speech programming for execution (Lewis, Freebairn, Hansen, Iyengar, et al., 2004; Maassen et al., 2001; Nijland, Maassen, & van der Meulen, 2003; Nijland et al., 2015; Shriberg, Campbell, et al., 2003; Shriberg, Green, Campbell, McSweeney, & Scheer, 2003; Thoonen et al., 1999). To clarify, speech motor planning, refers to the formulation of the specific articulatory movements or actions required to produce sequences of speech sounds, with each sound having a core motor plan that includes a number of motor goals arranged in a specific order (Van der Merwe, 1997). Motor planning is consistent with phonetic encoding and assembly of the phonetic plan, depicted on the WEAVER model. Motor programming, on the other hand, refers to the specifying the exact muscle tone, movement velocity, force and range of the articulators, which works in conjunction with the motor goals already established and which is facilitated by sensorimotor integration (Van der Merwe, 1997). This stage of processing is consistent with articulation in the WEAVER model and encompasses muscle command preparation and parameter setting (refer to Figure 1) (Levelt et al., 1999). These lower levels of speech processing, relating to speech motor planning (i.e., phonetic encoding) and speech motor programming, are highly pertinent to CAS as deficits observed in CAS are assumed to be due to a difficulty at this level of processing (Lewis et al., 2004, Maassen et al., 2001, Nijland et al., 2003, Shriberg et al., 2003). However, deficits have also been identified at the level of phonological encoding, consistent with a higher-level phonological deficit (Marion et al., 1993; Marquardt et al., 2004; Marquardt et al., 2002). In contrast, deficits observed in PD in phonological organization or learning phonological-linguistic rules are in the context of the WEAVER model predominately associated with a higher level deficit at the level of segmental spell-out, during which sound segments are slotted into phonological frame (Gierut, 1998b; Larrivee & Catts, 1999; Miccio et al., 1999; Rvachew, 2005). Deficits that emerge at this level of processing indicate poorly established or inaccurate phonological representations, including how phonological forms are organized in the mental lexicon. Evidence supporting this proposal is presented in the next section.

Deficits in CAS

Various theoretical frameworks have been developed to differentiate between deficits seen in CAS and PD. Ozanne (1995) developed a psycholinguistic model of speech output planning and programming, which denotes three different deficits that can occur in children with SSD: (a) deficits with the phonological plan, (b) deficits in assembly of phonetic program/plan, and (c) deficits in implementation or execution of motor plan, which according to Ozanne (1995) includes dysarthria. The first level relates to a phonological linguistic deficit, consistent with PD, whereas the remaining two levels relate to a speech motor deficit, including CAS. Consequently, children with CAS can exhibit difficulties at all levels of breakdown but in order to have a diagnosis of CAS they must have a deficit at the motor level relating to the assembly of the phonetic program/ plan (level b in the model). Children who only display deficits at the phonological level are not considered to have CAS and are consequently classified as having PD.

Stackhouse and Well's (1997) also developed a framework to profile speech-processing deficits. It was originally developed to explore deficits observed in children with a reading disability but also has the potential to profile the underlying speech processing abilities and deficits in SSD. This model enables categorization of children with SSD but has the potential to discriminate between children with phonological problems and motor-based problems. However, despite the various models and frameworks developed to differentiate between speech motor and phonological deficits in children with SSD a universally agreed framework is yet to be developed that permits a better understanding of these disorders. For the purpose of this PhD we adopted the WEAVER model, however, the terms speech planning and programming will be used throughout this thesis. These terms are dominant in the CAS literature and in the context of the WEAVER model they refer to phonetic encoding processes and articulation mechanisms.

The following section will reveal the broad range of deficits observed in CAS and the overlap of these deficits with PD. However, the majority of these research papers included either children with CAS and TD, or children with PD and TD, in their experimental groups, with relatively few studies combining CAS with PD and

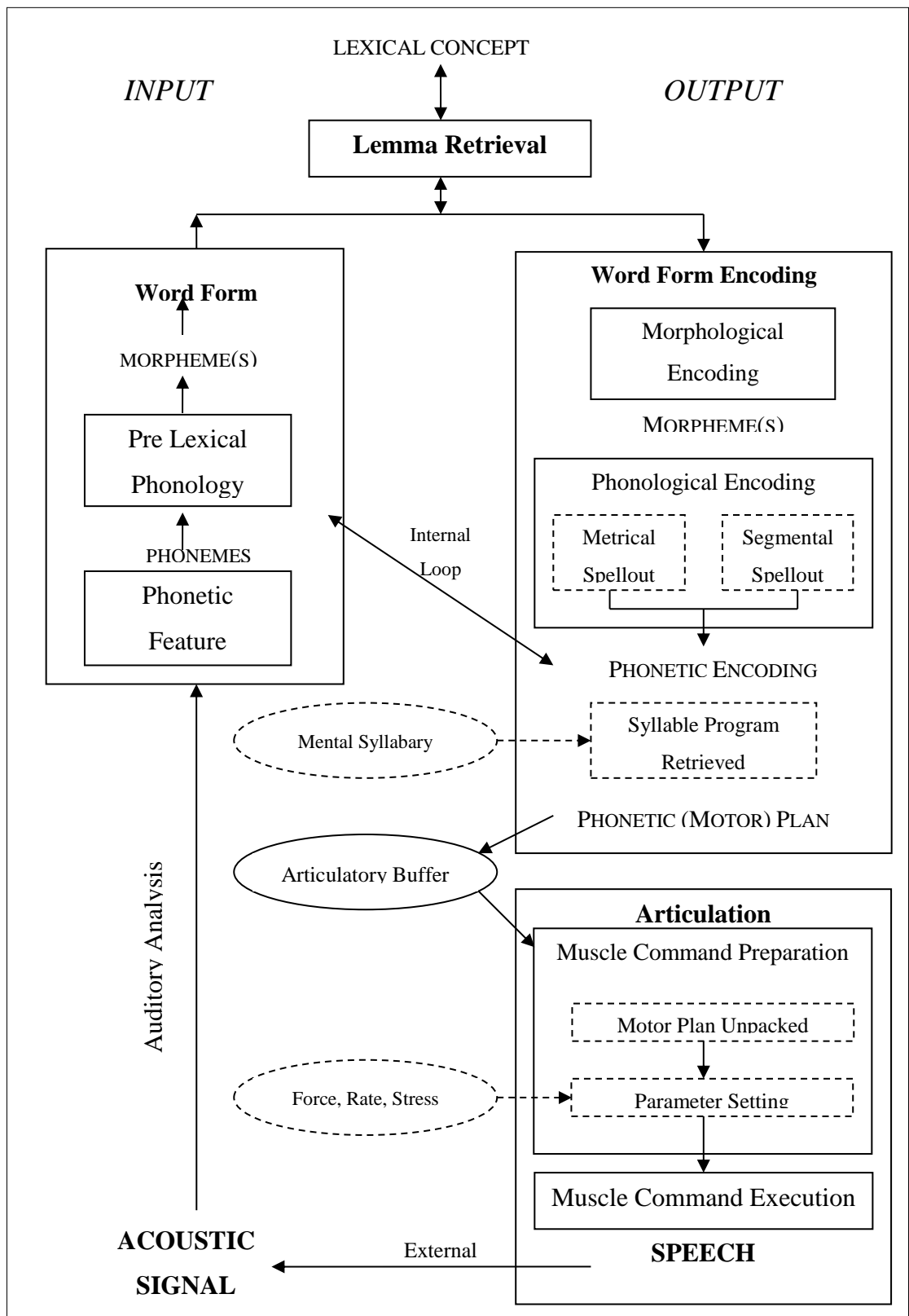


Figure 1

A Framework for Speech Recognition and Speech Production Adapted from the WEAVER Model (Roelofs, 1997)

TD children. Consequently, it is difficult to ascertain the true extent of shared deficits in CAS and PD. Our review of deficits in CAS and PD will start with evidence of speech motor deficits in children with CAS, which includes coarticulation studies, studies exploring lexical stress, timing deficits and compensatory speech motor abilities. We will then review evidence of phonological deficits in CAS, reviewing studies investigating speech perception deficits in CAS and the quality of underlying phonological representations.

Speech Motor Deficits

A number of experimental paradigms have been used to ascertain if children with CAS have a speech motor deficit that sets them apart from children with TD. Studies have explored coarticulation, lexical stress and timing deficits. Compensatory speech motor abilities have also been explored to determine if children with CAS have diminished ability to adapt to different speaking contexts, by altering speech parameters to compensate for these different conditions. These speech characteristics are explored from different approaches, such as acoustic measures that reflect the different components of the speech system.

Coarticulation. Coarticulation refers to the influence adjacent sounds and syllables have on one another in continuous speech. (Nijland, Maassen, van der Meulen, Gabreels, Kraaimaat, & Schrueder, 2002). Consistent with this view, Browman and Goldstein (1992) proposed that speech is comprised of successive articulatory gestures, the production of which are highly dependent on surrounding sounds and syllables. Consequently, the smooth transition between sounds and syllables is indicative of a well-established or more refined speech motor control system as a child's speech develops (Nittrouer, 2002). For this reason, a number of studies have investigated coarticulation in children with CAS to investigate potential problems with syllable planning and programming. This is consistent with the view that difficulties that emerge in coarticulation are indicative of problems with syllable structure and cohesion, indicative of a speech motor planning deficit (Nijland, Maassen, van der Meulen, Gabreels, Kraaimaat, & Schrueder, 2002).

Children with CAS have been shown to have difficulties in coarticulation (Maassen et al., 2001; Nijland & Maassen, 2005; Nijland, Maassen, & van der Meulen, 2003; Nijland, Maassen, van der Meulen, Gabreels, Kraaimaat, & Schreuder, 2002; Nijland, Maassen, van der Meulen, et al., 2003). Maassen, Nijland

and Van Der Meulen (2001) compared coarticulatory patterns in children with CAS ($n = 6$) and TD ($n = 6$) by exploring speech production errors in utterances that varied in complexity. The children with CAS revealed a greater percentage of errors for the more complicated speech motor programs (e.g., multi-syllabic words) and demonstrated greater overall variability in coarticulatory effects, indicative of unstable speech motor plans and consistent with a speech motor planning deficit. In the same study, the children with TD demonstrated a syllable boundary effect, demonstrated by durational differences between syllable initial and syllable final segments. The children with CAS did not show a syllable boundary effect, nor did they reveal any systematic durational changes. The authors concluded that, either this metrical information is deficient in children with CAS, or the motor system is not capable of planning and executing these temporal changes (Maassen et al., 2001).

Additional studies, also exploring coarticulatory effects in children with CAS, have revealed findings consistent with Maassen et al. (2001). Nijland et al. (2002) employed acoustic analysis to explore different properties of coarticulation. Second formant (F_2) trajectories were extracted from repetitions of disyllabic nonsense utterances to compare F_2 values across 9 children with CAS, children with TD, and adult women (Nijland, Maassen, van der Meulen, Gabreels, Kraaimaat, & Schrueder, 2002). Findings revealed that the children with CAS demonstrated greater variability overall in F_2 values when compared to the TD children and adult women. In addition, the average F_2 measures for the children with CAS was greater and F_2 ratios (calculated for each child separately to account for the anatomical differences between subjects) revealed that the children with CAS had a smaller [i/u] ratio, indicating less distinction between vowels, compared to the TD children and adult women. The children with CAS also showed idiosyncratic F_2 ratio patterns indicative of a less developed speech motor system. Overall, the high variability observed in the children with CAS was taken to be consistent with a less developed speech system and interpreted as immature automation of speech motor control (Nijland, Maassen, van der Meulen, Gabreels, Kraaimaat, & Schreuder, 2002). These findings were echoed in a later study that compared syllable planning in children with CAS and TD (Nijland, Maassen, van der Meulen, et al., 2003). This study also showed that the children with CAS had greater variability in coarticulation patterns when compared to the TD children. The children with CAS also displayed significantly longer total and segment durations. Furthermore, the children with TD

produced shorter vowel durations for vowels in prosodically weaker positions, something that was not demonstrated by the children with CAS. This is consistent with the view that children with CAS have difficulty retrieving syllable programs or have difficulty implementing subtle temporal differences required to reflect changes in duration and prosody (Nijland, Maassen, van der Meulen, et al., 2003). Overall, these findings indicate that the speech of children with CAS show deviant coarticulation patterns, with overall greater variability and less acoustic distinction between vowels, consistent with a speech motor planning/programming deficit. However, these deviations from typical coarticulation patterns could arise from impaired phonological representations or deviant phonological encoding. For example, if phonological representations are inaccurate or unstable then this could result in greater variability or less precision in the subsequent stages of speech motor planning and programming. The findings, therefore, could also be more consistent with a phonological deficit, rather than just a speech motor deficit, as proposed.

Bahr (2005) assessed the articulatory skills of five children with CAS, PD and TD using a gesture articulation test developed to ascertain if children with CAS had difficulties transitioning from one point of articulation to another. The test assessed individual consonant vowel (CV) gestures, vowel consonant (VC) combinations, and multiple repetitions of CV and VC syllables and use of various stress patterns in multisyllabic words. The children with CAS and PD did not differ in relation to accuracy, however, acoustic analysis revealed that the groups differed in relation to duration, with the children with CAS having significantly longer word durations than the children with PD and TD for all gesture patterns (Bahr, 2005). Although the findings were equivocal in relation to phoneme accuracy, the authors proposed that the longer word durations for the children with CAS suggest that these children have more difficulty coordinating gestures, as well as movement between gestures, indicative of a speech motor programming deficit (Bahr, Velleman, & Ziegler, 1999).

Lexical stress. Children with CAS are often reported as having difficulty with stress, intonational contour and other suprasegmental characteristics of speech. The findings are not straightforward in terms of their implication, with lexical stress errors being attributed to different levels of processing within the speech production system. According to Roelofs (1997b), stress deficits are indicative of a problem at the level of speech motor programming, during which acoustic parameters for stress

are assigned. These acoustic parameters govern how the speech is produced with regard to intonation, pitch and loudness. However, deficits assigning lexical stress have also been attributed to the linguistic level of processing (Levelt et al., 1999) whereby incorrect stress assignment was claimed to be a result of poorly defined phonological representations, presumably at the level of metrical spellout in the WEAVER model (Roelofs, 1997b).

A number of studies have explored the perceptual correlates of stress in children with CAS (Munson, Bjorum, & Windsor, 2003; Skinder, Connaghan, Strand, & Betz, 2000; Skinder, Strand, & Mignerey, 1999). Skinder et al. (2000) evaluated the relationship between perceptual codes and acoustic measures in five children with CAS and TD to determine if stress could be accurately judged perceptually. Bi-syllabic words varying in stress patterns (i.e., trochaic and iambic stress patterns) were elicited in isolation and phrase final position. Acoustic measures extracted for analysis included vowel duration, peak fundamental frequency, fundamental frequency excursion (i.e., the difference between the highest and lowest point on fundamental frequency contour), and peak amplitude. For the children with TD acoustic measures and perceptual ratings were consistent, whereas, for the children with CAS mean values for the acoustic correlates of stress were not consistent with some of the perceptual ratings. The acoustic measures for peak fundamental frequency and amplitude were accurately produced to mark stress but perceptually were coded as incorrect. In addition, for the children with CAS speech that was coded as accurate had more consistent acoustic measures compared to speech that was coded as incorrect, suggesting that segmental accuracy may play a role in the perception of the appropriate stress assignment. One interpretation could be that the accuracy of these segments may in fact influence the production of lexical stress errors, that is, stress is assigned incorrectly because the phonetic plans are incorrect. Alternatively, the lack of an established motor plan could also adversely affect stress assignment (Skinder et al., 2000).

Munson, Bjorum and Windsor (2003) also explored the perceptual correlates of stress in five children CAS and PD using nonwords. Measures relating to stress production were examined, these included; vowel duration, fundamental frequency (F_0) at vowel midpoint, timing of F_0 peak and intensity. There were no group differences in the production of stress in relation to the acoustic parameters but listeners judged that the children with CAS did not match the stress contour to the

same extent as the children with PD. These findings were consistent with Skinder et al.'s findings (1999, 2000) indicating that children with CAS can alter acoustic parameters to mark stress, despite being perceived as not marking stress appropriately.

Shriberg et al. (2003) used a lexical stress task to assess stress assignment in 15 children with CAS and PD. The lexical stress task required the participants to imitate 24 bi-syllabic words in isolation with varying stress patterns. Findings revealed that on a number of acoustic measures (i.e., fundamental frequency, amplitude, duration) the children with CAS had the most extreme scores indicating that some of these children demonstrated excessive stress, whereas others demonstrated the weakest stress patterns. Shriberg and colleagues (2003) interpreted this variability to be indicative of a praxis deficit in speech motor programming.

Timing deficits. Deficits observed in relation to timing are indicative of a problem in speech motor control based on the assumption that parameters for timing are assigned at the level of speech motor programming (Roelofs, 1997b). Shriberg et al. (2003) explored temporal variation in conversational speech in 11 children with CAS, PD and TD. The objective being to determine if temporal variation in speech could be used as a diagnostic marker for CAS by investigating speech and pause events in children who had been described as having isochronous and segregated syllables, consistent with children with CAS. Acoustic analysis was used to calculate a coefficient of variation ratio to compare the variations between speech events and pause events. The children with CAS differed from the children with PD and TD, demonstrated by reduced variation in the duration of speech events and increased variation in the duration of pause events. The findings from pause events were ambiguous; given pause events could reflect other aspects of the speech production system related to tasks demands, such as sentence formulation and/or word retrieval. However, the reduced variation in speech events was interpreted as a deficit at the level of speech motor control, consistent with the view that temporal variation in speech events is assigned at this level of processing (Shriberg, Green, et al., 2003).

Peter and Stoel-Gammon (2005, 2008) also explored speech and other timing abilities of two children with CAS to investigate prosodic errors in relation to timing. They compared children with CAS and TD on three speech tasks (e.g., sentence imitation, nonword repetition and picture-naming) and three music tasks (e.g., paced

repetitive tapping, clapping and singing). For the speech tasks the authors looked specifically at vowel duration and syllable omissions. Overall, the children with CAS were less accurate than the children with TD on the speech imitation tasks. The children with CAS also had a less accurate syllable count than the TD peers, omitting weak syllables. In relation to the music tasks, the children with CAS were less accurate for all three tasks, however, the singing task yielded the greatest accuracy discrepancy between the children with CAS and TD, demonstrated by the children with CAS having difficulty in producing a coherent rhythmic structure. Overall, these results support the proposal that children with CAS have a central timing deficit, which the authors contended was consistent with a speech motor deficit (Peter & Stoel-Gammon, 2005, 2008). However, given the sample did not include children with PD findings cannot be interpreted as specific to CAS and therefore may occur in SSD in general.

Compensatory speech motor deficit. Compensatory speech motor ability is the ability to adjust online speech production parameters to reflect the immediate speaking environment (Crary, 1995). The bite block paradigm has been used to assess these compensatory skills in children with CAS and PD (Edwards 1992; Nijland et al., 2003). The bite block task requires a participant to hold a bite block between their teeth during a speech task, during which measures are obtained during a speech task. This task assesses the compensatory speech motor ability of the speaker to adjust online parameters to ensure articulatory goals are achieved, therefore directly targeting speech motor programming efficiency and the speaker's ability to adjust on-line parameters relating to muscle tone, rate, direction and movements (Perkell et al., 2000; Perkell et al., 1997; Van der Merwe, 1997). Nijland et al. (2003) used a bite block paradigm to assess speech motor compensatory abilities in six children with CAS, TD children and adult women (AW). Second formant (F_2) values were extracted from two-word utterances with simple CV shapes that each participant produced with and without the bite block. The F_2 trajectory was measured throughout the utterance and a ratio was calculated for each participant to correct systematic differences among speakers for both speaking conditions. F_2 values were then used to assess the effect of the bite block for each speaker to determine the articulatory compensation and the impact the bite block had on speech production. Findings revealed that the adult women compensated fully for the bite block (as expected), demonstrated by no significant change to F_2

values. The children with CAS and the TD children were unable to compensate completely, consistent with a less developed speech motor system. Remarkably, the children with CAS appeared to benefit from the bite block, demonstrated by a reduction in F_2 values, bringing F_2 values closer to TD children. The authors proposed that the bite block provided greater stability for the children with CAS resulting in enhanced speech production due to the reduction in potential movement parameters. However, the variability analysis revealed that the adult women showed the smallest within-subject variability and the children with CAS showed the largest, in both the normal and bite block conditions. The variability demonstrated by the children with CAS was taken to be indicative of deficient speech motor programming. A similar paradigm was used to assess speech compensatory abilities in children with PD and TD (Edwards, 1992). Findings revealed that the TD children were unable to compensate for the bite block condition to the same degree as the adult women, however, they did reveal a consistent pattern of compensation. In contrast, the children with PD revealed high within group variability, consistent with poor compensation by their speech motor system. Furthermore, two of the four children with PD most closely resembled the TD children whereas the other two revealed idiosyncratic problems with and without the bite block, suggesting that some children with PD may have speech motor control deficits (Edwards, 1992). It is plausible that these two children may have been incorrectly classified as PD or alternatively their phonological deficit has resulted in instability within the speech motor system, resulting in poorer speech motor skills. These findings highlight the complexity of diagnosing developmental speech disorders and emphasize the need for future research to examine different speech disorders simultaneously to ascertain what differences or similarities emerge.

In summary, there is a range of evidence indicating an underlying deficit at the level of speech motor control in children with CAS. Studies exploring coarticulation and syllable planning have demonstrated that children with CAS have difficulty with phonetic planning, although these studies typically do not include children with PD, for example. In addition, poor compensatory speech motor abilities have been demonstrated in children with CAS, indicative of a deficit in speech motor programming. However, children with PD have also been shown to have poorer speech compensatory motor abilities compared to TD (Edwards, 1992). Although, as previously stated, the source of the deficit is unclear, and it is feasible that in the case

of the children with PD, a phonological deficit could affect phonetic planning and articulation, as discussed in the following section.

Linguistic/Phonological Deficits

The general consensus is that PD is a disorder at the level of phonological encoding, reflecting an underlying deficit in forming and accessing accurate and well-specified phonological representations or in developing the phonological rules that govern the patterns of sounds within the child's language (Dodd, 2005; Dodd et al., 2005). A number of studies using speech perception and discrimination tasks have demonstrated that children with PD have deficits at the phonological level of processing (Edwards et al., 2002; Jamieson & Rvachew, 1992; Locke, 1980; Rvachew & Jamieson, 1989). However, phonological deficits have also been demonstrated in children with CAS, resulting in the proposition that the underlying deficit in CAS is impoverished phonological representations and the motor output problems are a consequence of this higher level linguistic deficit (Marion et al., 1993; Marquardt et al., 2004; Marquardt et al., 2002; Mc Neil et al., 2009; Sussman, Marquardt, & Doyle, 2000).

Deficits at the phonological level have been explored in relation to both speech input and output phonological representations. However, there are different theoretical perspectives with regard to the nature of the phonological representations for speech input and output processes. One account is that there are separate phonological representations for speech input and output (Monsell, 1987) and the alternative is a shared phonology between input and output (Dell et al., 1997). Although this is unresolved there is some argument to support the separate view. In the context of the WEAVER model speech input and output do not share the same phonological representations (Roelofs et al., 2007) although, they do interact through direct connections, with the pathways only merging at the higher level of syntactic coding (lemmas) and meaning (lexical concepts). Consequently, the input phonological representations in the WEAVER model are different from those that play a role in phonological encoding.

Speech perception. Groenen et al. (1996) explored the relationship between speech perception and production of speech in 17 children with CAS and TD (mean age 8:9, years: months). Children were assessed on identification and discrimination of monosyllabic words combining natural and synthesized speech. The identification

task required phonemic judgment of the presented items based on phonetic properties of the speech signal (e.g., place of articulation) and assessed auditory processing of the speech signal and the quality of input phonological representations. The discrimination task, consisted of same or different judgments, and assessed the phonetic properties of the speech signal and auditory memory (Groenen, Maassen, Crul, & Thoonen, 1996). The children with CAS demonstrated similar phonetic processing skills to the children with TD for the identification task, indicating that they did not have a problem with phonetic categorization. However, they demonstrated poorer discrimination skills, indicating diminished input phonological representations and diminished access to information in auditory memory than children with TD (Groenen, et al., 1996).

Maassen et al. (2003) used a similar paradigm to explore the auditory/phonetic perception of vowels in 11 children with CAS (aged 6:11 to 9:6 years) and 12 children with TD. They found that the children with CAS performed more poorly than the children with TD on both the discrimination task and the identification task. They also found that the children with CAS also exhibited greater variability in both identification and discrimination abilities. Overall, the results were interpreted as evidence that children with CAS have difficulty with phonetic and auditory processing of vowels (Maassen et al., 2003).

More recently, Nijland (2009) investigated the possibility that perception problems might underlie speech production problems observed in children with CAS and PD. This was based on previous findings that had revealed speech perception deficits in children with CAS and a specific relationship between production and perception errors (Groenen et al., 1996; Maassen et al., 2003). They implemented a number of tasks that assessed both higher-level speech perception and lower-level speech perception to ascertain if children with CAS only presented with lower-level speech perception deficits and children with PD presented with higher level deficits. The higher-level speech perception tasks included a rhyming task and categorical classification, and the lower-level speech perception tasks included a nonword auditory discrimination task and categorical discrimination task. A frequency pattern task was also implemented to assess auditory temporal processing. Findings revealed that the children with CAS and PD performed more poorly overall than the children with TD, but there was no distinction between the children with CAS and PD (Nijland, 2009). The children with PD demonstrated lower scores on the

rhyming task but not on the discrimination task (performing more similar to the children with TD), consistent with the view that higher-level perception problems are linked to higher-level speech production problems in these children. However, the children with CAS demonstrated lower scores on the rhyming task and discrimination task, indicating both higher and lower-level production problems. Although the findings were ambiguous for the children with CAS and the deficits were not isolated to lower-level of processing, as hypothesized, they are consistent with earlier studies (Groenen et al., 1996; Maassen et al., 2003). These findings highlight the need for further research to explore the causal relationship between speech perception and production and its potential impact on speech development.

Phonological representations. Consistent with the view that children with CAS have a speech motor deficit resulting in difficulty coordinating movement sequences responsible for speech output, it has been proposed that the underlying deficit is phonological/linguistic in nature. Marion, Sussman and Marquardt (1993), proposed that children with CAS do not have well-formed output phonological representations required to guide articulatory goals. They investigated rhyming ability in children with CAS and TD using a number of tasks that included; rhyme production, rhyme detection (forced choice of two words that rhyme with target, e.g., target: ball, options: bought or fall), serial rhyme detection (child had to identify which words rhymed with target word presented at the outset of the ten-trials), and vowel comparisons (child asked to judge acoustic similarities between vowels). Their hypothesis was based on the view that speech production and perception are interdependent and consequently speech motor output would be compromised if phonological targets were incomplete. Similarly, they proposed that perceptual processes would also be compromised, because the auditory processing of the speech signal could not be mapped to well-formed phonological representations. Findings revealed that the children with CAS had a severe inability to recognize and produce rhymes. In addition, they were significantly less accurate judging vowel similarities (children with CAS achieved a score of 35.6% compared to 80% for the children with TD). The authors, whilst acknowledging that children with CAS have a severe inability to execute articulatory actions, proposed that the speech output deficit of these children is a consequence of a higher level phonological deficit that affects both perception and production of rhyme (Marion et al., 1993). Their concluding remarks were that the vowel is integral for rhyme detection and production and given

the children with CAS demonstrated such diminished accuracy in vowel comparisons, they suggested that an impoverished vowel system could also account for the rhyming deficits observed in these children (Marion et al., 1993).

Sussman, Marquardt and Doyle (2000) later investigated the phonemic integrity and distinction of phonological representations in children with CAS using acoustic analysis. Locus equations are an acoustic measure of articulatory synergy (Iskarous, Fowler & Whalen, 2010) and, therefore, can be used to measure the degree of anticipatory coarticulation. Sussman, Marquardt and Doyle (2000) used locus equations to capture the strength of the vowels influence on the preceding consonant with the aim of measuring the acoustic distinction between different consonants. Findings revealed that the children with TD had a similar acoustic contrast between the consonants to the adults, whereas the children with CAS were unable to refine coarticulation to maximally distinguish between the consonants. The authors proposed that the children with CAS lacked the underlying phonological prerequisites that permitted maximal articulatory goals and auditory processing distinctiveness (Sussman et al., 2000).

Marquardt, Sussman, Snow and Jacks (2002), whilst recognizing that there was a consensus that CAS was a motor speech disorder, they wanted to ascertain if deficits in CAS could be more related to linguistic mechanisms. They proposed that an underlying linguistic deficit at the level of phonological representations could account for the wide diversity of clinical deficits observed in CAS (Marquardt et al., 2002). Based on preliminary findings from Marion et al.'s (1993) rhyming study and the concluding remarks that the vowel plays an integral part in rhyme detection and production, Marquardt et al. (2002) investigated the integrity of the syllable from a metalinguistic perspective. The tasks they implemented included syllable awareness, demonstrated by the ability to detect syllables, intra-syllabic position, demonstrated by ability to judge intra-syllabic sounds (i.e., whether the different sound was first, middle or last of three CVC words presented), and intra-syllabic structure, demonstrated by ability to judge single versus consonant clusters. The children with CAS had difficulty segmenting syllables, judging intra-syllabic sound positions and constructing single and consonant clusters, compared to children with TD. Marquardt, Sussman and Jacks (2002) proposed that these findings provided a clear indication that the underlying deficit in CAS is greater than an articulatory deficit and that deficits are more consistent with an impoverished phonological

representational system that also impacts their phonological awareness of spoken language.

Davis, Jacks and Marquardt (2005) also explored the integrity of vowel patterns in children with CAS. They did this over a three year period to ascertain a longitudinal description of vowel inventory and error persistency, despite ongoing treatment (Davis et al., 2005). In typical development the vowel inventory of infants is complete by 24 months, with some studies suggesting that the inventory is complete by 14 months (Davis et al., 2005). Davis et al. (2005) revealed the all three children had impaired vowel accuracy despite having complete vowel inventories. However, no length or syllable complexity effects were found, which the authors concluded were more consistent with a phonological representational deficit rather than a motor one (Davis et al., 2005).

The proposal that the deficits listed above are more consistent with a higher level of processing needs to be interpreted with caution. Sussman et al. (2000) suggest that problems with coarticulation are consistent with a higher-level phonological deficit, although coarticulation deficits have been largely interpreted as pertaining to a deficit at the level of speech motor programming (Nijland, Maassen, van der Meulen, Gabreels, Kraaimaat, & Schreuder, 2002). Marquardt, Sussman, Snow and Jacks (2002) also interpreted a syllable awareness deficit as indicative of a higher level linguistic deficit, although deficits relating to the syllable have been interpreted as an immature mental syllabary (Maassen, 2002) and more consistent with a motor programming deficit (Nijland, Maassen, van der Meulen, et al., 2003). Davis, Jacks and Marquardt (2005) interpretation that vowel errors in CAS related were more consistent with a phonological deficit rather than a motor deficit because they also found that no consistent error pattern emerged for the children with CAS, despite the children all having complete vowel inventories.

Speech Input and Output

A number of studies have investigated the relationship between speech input and output in the context of speech and language development. From a developmental perspective, the different levels of processing are highly interdependent, and as the child develops a speech and language system more consistent with that of an adult, the levels achieve autonomy. This view is in keeping with developmental models, such as Maassen's (2002), Westerman and

Miranda's sensori integration model (2004) and Kulh's (1992) Native Language Magnet theory. It also reiterates the view that in speech and language development it is the associations between deficits that are informative not the dissociations (Bishop, 1997) and validates the importance of exploring speech impairments in a developmental context.

Edwards, Fourakis, Beckman and Fox (1999) investigated the relationship between speech input and output processes in children with PD and TD (age range 3 to 5 years). Their aim was to ascertain the degree of interaction between phonological knowledge, perceptual knowledge and the motor component of speech production. They investigated different levels of phonological representations by looking at three components of phonological development, acoustic/perceptual space, articulatory/production space and the inverse mapping between the two. The acoustic perceptual space relates to how the acoustic signal is processed in terms of the speech input representation, which enables the child to detect different consonants and vowels in their native language. The second component, the articulatory/production space, relates to the child's internal model for articulatory goals. To evaluate perceptual knowledge a backward gating task and a noise-centered identification task were implemented. For these tasks participants were required to identify a word with incomplete acoustic information. For the backward gating task, the final stop consonant of CVC words (e.g., /p/ in *tap*) was removed to varying extents (e.g., just the release burst portion along with the preceding stop-gap, or the formant transition in the vowel prior to the stop gap and the release burst). For the noise-centered task the vowel (of CVC target words) was degraded with added noise. To evaluate articulatory/production space, measures of vowel formant dynamics were abstracted from phrases that contained CV transitions with voiced stop consonants (e.g., /b/, /d/ and /g/). Findings revealed that the children with PD were significantly less accurate than children with TD in identifying both backward gated and the noise centered CVC words. For the articulatory/production space evaluation the children with PD and TD showed considerable variability, within and between participants, typical of children of the age range investigated. However, the children with PD showed poorer control over speech rate, which was most obvious for the vowel durations. Overall the children with PD were less able to vary duration when instructed to vary rate, indicating that they were already speaking as quickly as they could. The authors contended that the sample size was small and was

heterogeneous, however, the children with PD differed from their peers in relation to perception, production and the inverse mapping between the two (Edwards et al., 1999). Overall these findings were interpreted as an indication that some children with PD have a weak cognitive representation required for perception and motor control necessary for producing coordinating gestures. These results are consistent with the view that phonological contrasts emerge as a result of the incremental acquisition of robust representations at the different levels of processing and in doing so highlights the interdependency between the different levels of processing during development (Edwards, 1999).

A later study implementing a similar paradigm, evaluated preschool aged children with PD on their ability to discriminate CVC words that differed only in the identity of the final consonant (Edwards et al., 2002). Performance was also assessed, as a comparison, in three groups of children with typical development (age range; 3-4 years, 5-6 years and 7-8 years) and adults. In the first experiment, with the typically developing children and adults, findings revealed that the younger children needed more acoustic information to accurately discriminate between the words that differed only in final consonant, consistent with previous findings (Edwards et al., 1999; Munson, 2001b). In the ungated condition, the performances of the two older groups were similar to the adults but for the gated conditions there were significant differences in performance across all age groups indicating that there is a gradual improvement with age in word discrimination under difficult listening conditions (Edwards et al., 2002). In a second experiment, with 35 preschoolers with PD and 35 age matched TD peers (mean age 56 months) the children with the PD performed more poorly than the TD children indicating that children with PD have difficulty attending to fine phonetic detail.

Munson et al. (2005b) implemented the same auditory discrimination task in conjunction with a nonword repetition task in children with PD and TD to investigate the relationship between phonological development and speech discrimination ability. A nonword repetition task was used to ascertain the degree of phonological development, consistent with the view that ability to repeat nonwords accurately is a reflection of the quality and abstractness of underlying (sub-lexical) phonological representations. Furthermore, by including phoneme sequences within the nonwords that varied in phonotactic frequency (i.e., the frequency sequences as attested in real words) permitted a greater insight into the stage of phonological development.

Phonotactic frequency refers to the how frequently a sequence of phonemes occur in the mental lexicon (i.e., the frequency a sequence appears in a real word, for example /*mp*/ occurs in many words and therefore is considered a high frequency sequence, whereas /*fk*/ occurs in few words and therefore is considered a low frequency sequence). A growing body of evidence suggests that phonotactic frequency effect (i.e., the difference in repetition accuracy between high and low frequency sequences) changes as a function of speech and language development (Beckman & Edwards, 2000; Coady & Aslin, 2004). Phonotactic frequency has been shown to influence naming latencies in picture-naming and accuracy of nonword repetition resulting in faster naming and higher accuracy (Newman & German, 2002). Furthermore, the phonotactic frequency effect in children has been shown to decline with age in relation to duration and accuracy and this decline is predicted by vocabulary size after controlling for speech discrimination and articulation ability (Munson, 2001a). These findings are consistent with the view that the phonotactic frequency effect reflects the development of phonological knowledge and is not just linked, for example, to developmental changes in speech motor control. In particular, as the child becomes a more competent speaker the underlying phonological representations emerge as autonomous units. This results in the phonological representations becoming more segmented from the words in which they occur, regardless of their frequency, and this process has been shown to occur as a result of vocabulary growth (Metsala, 1999; Walley, 1993). This is based on the view that as the lexicon expands words cannot be stored holistically and processed efficiently, consequently, the representations need to be broken down into more manageable parts to enable differentiation between many lexical items that overlap in their phonological structure (Munson, Edwards, & Beckman, 2005b). These independent phonemic units can therefore be productively assembled in novel ways, independent of how frequently the sequences occur in the mental lexicon, consistent with the phonotactic frequency effect declining with age and vocabulary growth. The nonword repetition task with high and low phonotactic frequency sequences potentially targets components of phonological knowledge relating specifically to the formation of abstract phonemic categories that are available for explicit manipulation, as in nonword repetition.

Munson et al. (2005b) found that the children with PD were overall less accurate than the children with TD, however they were not disadvantaged when

repeating the low frequency sequences, not what they had predicted. Munson et al. (2005b) then examined predictors of overall accuracy and the phonotactic frequency effect, using regression analysis. Potential predictors in their study included age, speech discrimination ability, measures of expressive and receptive vocabulary (as measured by the EVT and PPVT, respectively) and articulatory ability (as measured by the GFTA). They found that measures of vocabulary predicted the difference in repetition accuracy between low and high frequency sequences (i.e., the phonotactic frequency effect) over and above that predicted by age (Munson, et al., 2005). They also found that the magnitude of the phonotactic frequency effect was independent of speech discrimination ability, indicating that developmental decreases in the frequency effect were due to vocabulary growth and not development in speech perception or production (Munson et al., 2005). Munson et al. (2005b) concluded that the deficits in PD are not associated with forming robust phonological representations and are more likely to be associated with difficulties forming poor acoustic-auditory maps relating to more peripheral processes. The findings from regression analysis support the proposal that phonological development is driven by the gradual acquisition of lexical items that help establish links between acoustic and articulatory maps required, which in turn permit children to produce novel gestures accurately and fluently, but phonological knowledge may also be influenced by other levels of development, such as speech perception and production skills (Munson et al., 2005b). More importantly, the research paradigm suggests that different patterns of predictive relationship or covariance between levels of development, such as vocabulary, speech input and speech output abilities, on the emergence of phonological skills can be informative regarding sources of constraint on development and therefore location of underlying deficits. The research paradigm used by Munson and colleagues has been applied to other types of disorder where phonological processing deficits have been implicated, such as specific language impairment (Munson, Kurtz, & Windsor, 2005), but to date research examining these aspects of phonological development has not yet been investigated in children with CAS. This paradigm provides a platform to extend this research and include children with CAS and PD and may be a useful strategy to compare differences in developmental constraint on speech and language development in children with CAS and PD.

Limitations in Research in CAS

There are a number of limitations in the research in children with CAS. The sample sizes in these studies were predominantly small, with the most common sample size range from two children to 15 children. Furthermore, the classification criteria for a number of these studies did not provide a clear classification protocol, clearly defining how participants were assigned to their respective groups. Consequently, it is difficult to ascertain if all participants were true cases of CAS given the problem with over diagnosis (ASHA, 2007). A major limitation is that research in CAS has rarely included children with another speech sound disorder, such as PD, in the same studies. Consequently, it is not clear that the findings are specific to CAS. The few studies that have included children with CAS and PD failed to reveal significant differences between these groups that could be used for differentiation purposes. The majority of these studies have shown that children with CAS and PD share a number of speech deficits, implicating deficits at multiple level of processing with deficits revealed in the speech output pathways, at the level of phonological encoding and speech motor control, and speech input pathways.

There are a number of discrepancies in relation to interpretation of findings regarding the level of processing implicated. For example, difficulties with coarticulation have been interpreted as a deficit at the level of speech motor programming (Maassen et al., 2001; Nijland, Maassen, & van der Meulen, 2003; Nijland, Maassen, van der Meulen, Gabreels, Kraaimaat, & Schreuder, 2002) and higher-level phonological deficit (Sussman et al., 2000). In addition, deficits assigning lexical stress have been attributed to a deficit at the level of phonological encoding (Shriberg, Aram, & Kwiatkowski, 1997b, 1997c), yet according to the WEAVER model (Roelofs, 1997b) the parameters that determine lexical stress are assigned during muscle command preparation stage, just prior to execution.

Although both speech perception and production have been examined in children with CAS the causal relationship between these processes has not been sufficiently explored. Both Groenen et al. (1996) and Maassen et al. (2003) examined the speech perception and production in children with CAS and demonstrated poorer speech discrimination and identification skills when compared to children with TD. Nijland (2009) explored the basis for these speech perception deficits by comparing children with CAS and PD on higher-level speech perception tasks presumed to affect children with PD (e.g., rhyming and categorical perception)

and lower-level speech perception tasks presumed to affect children with CAS (e.g., nonword auditory discrimination and categorical discrimination). Findings revealed that the children with PD had higher-level deficits but the children with CAS had both higher and lower-level deficits. Nijland (2009) contended that despite the findings being inconclusive further research looking at the interdependencies between the different levels of processing was needed. Edwards et al. (1999; 2002) explored the interactions between speech input and output in children with PD and revealed findings consistent with the view that constraints can emerge in a developing system resulting in deficits at a number of levels of the speech and language system. More recent research has revealed a sensorimotor influence on speech perception skills in infants, indicating that articulatory configurations can influence auditory speech perception consistent with the proposal that speech motor ability can to some extent determine the quality of phonological representations (Bruderer, Danielson, Kandhadai, & Werker, 2015a). Consequently, lower level/speech motor deficits can affect higher order processes and vice versa and can explain why children with CAS present with difficulties at multiple levels of processing (Bruderer et al., 2015). The interdependency between the different levels of processing can explain why children with CAS present with speech perception deficits and can explain why deficits that appear to indicate a higher-level phonological deficit could in fact be due to a speech motor deficit. The following section will address speech development in the context of Maassen's (2002) developmental model and computational models to elaborate on this point. We will then look at studies that have investigated the interaction between the different levels of processing during development.

Speech Development

The WEAVER (Roelofs, 1997b) model depicts the processes involved in speech in a mature speech system. However, a more relevant approach for developmental disorders would be the use of a developmental model, such as Maassen's (2002) model, which is an extension of the WEAVER (Roelofs, 1997b) model incorporating a developmental perspective. A developmental perspective permits a greater understanding of potential outcomes should a child's speech and language not advance consistent with our understanding of typical speech and language milestones. In addition, this perspective emphasizes how the different

levels of processing interact during speech and language development.

Computational models also permit a broader understanding of the interactions between the different levels of processing by demonstrating possible outcomes when specific processes are purposely manipulated or impaired. The following section gives an overview of speech development and explores these different theoretical perspectives.

According to Terband and Maassen (2010), there are two clear developmental stages in the acquisition of speech. The first is the perceptual motor stage, which involves the development of the articulatory motor system, during which the child develops internal models (abstract representations) to meet articulatory goals. These representations are initially very basic but with practice they become more refined and more consistent with the adult model. The second stage is the phonological stage, which involves the acquisition of the phonological system and involves the child establishing abstract phonological representations that relate to sounds of the child's native language. Terband and Maassen (2010) propose that during the second stage perceptual motor skills are further shaped and refined due to the ongoing feedback resulting in the first meaningful utterances (Terband & Maassen, 2010). This theoretical perspective is in keeping with earlier models of speech and language development, such as Kuhl's (1992) Native Language Magnet (NML) theory. Kuhl's (1992) NML theory proposes that perceptual prototypes are developed through early linguistic experience and these prototypes enable the infant to detect phonetic differences in the speech signal. This in turn helps the child refine production patterns consistent with the phonological structure of the child's native language. These early language experiences significantly impact on the child's speech and language development. Movement patterns, which the child has acquired through vocal play and babbling, are favored when acquiring first words and development proceeds with the gradual implementation of new movement patterns (Lindblom, 2000). The onset of the phonological organization is superimposed on the child's ongoing phonetic learning and first phonology, which is based on familiar patterns (or articulatory motor plans/gestures) the child has already acquired (Vihman & Velleman, 2000). Furthermore, as new patterns are acquired, already established patterns are modified, so that acquisition of more accurate and complex representations can be established (Piske, 1997). These articulatory motor plans/gestures help establish phonetic features required for speech perception and

production and the changes that occur over the course of development reflect the child's ability to integrate these gestures to produce words of increasing complexity (Fowler, 1991).

Westerman and Miranda's (2004) model of sensorimotor integration reiterates this view of speech and language development. This model proposes that speech output (e.g., babbling) allows the infant to develop the link between articulation and auditory feedback permitting the coupling between perception and production. Westerman and Miranda's (2004) model comprises two maps, one for auditory stimuli and the other for motor commands. Westerman and Miranda (2004) propose that connections develop between these two maps and these connections change over the course of development resulting in developmental changes in both speech perception and production. This is consistent with Perkell's (1980; 1997) concept of an internal model. Perkell (19880; 1997) proposes that segmental speech movements are regions in auditory-perceptual space, defined by oro-sensory patterns, which are developed due to the integration of auditory, proprioceptive and tactile feedback. These segmental speech movements develop over the course of language development and result in an internal model that corresponds to the production of a specific sound (Perkell, 1980). The internal model is a learned model that correlates with the configuration of the vocal tract when producing that particular sound and auditory feedback is used to train and maintain the internal model, in accordance with somatosensory information (Perkell et al., 1997). Consequently, the more vocal the child is during this early stage of speech development the greater the opportunity to establish the feedback mechanisms that enables the child to monitor and refine their own speech output (Menn & Stoel-Gammon, 1995). These first syllables form the protosyllabary, as described by Maassen (2002), and the repertoire of sounds and syllables increase as frequently used syllables are added to the mental syllabary.

The first words produced by children usually simplify the adult model and this is generally interpreted as immaturity in neuromotor control, or immature linguistic representations (Piske, 1997). The child, therefore, initially operates on a very limited inventory of articulatory patterns and these patterns function as the building blocks for phonological development (McCune & Vihman, 2001). However, through successive attempts, words become more refined resulting in more accurate and consistent productions, which occurs as the child attempt to match what they hear (Kuhl & Meltzoff, 1996). As development proceeds, the units in the child's

speech output repertoire develops from the syllable to the phoneme (Nittrouer, 1993; Nittrouer, Studdert-Kennedy, & Neely, 1996). Early phonological representations are holistic and lexical growth influences the child's ability to segment words into individual phonemes and produce and perceive these phonemes independent of the segments in which they occur (Edwards et al., 2004; Metsala, 1999; Storkel, 2002). This enables the child to distinguish among the ever increasing number of lexical items in their vocabulary (McCune & Vihman, 2001). The phonemic units that emerge as a result of the vocabulary spurt are more abstract categorical representations that link to the acoustic and articulatory representations that the child accrues during speech and language development (Munson et. al, 200b).

A Developmental Account of CAS

Maassen (2002) adapted Levelt and colleagues (Levelt et al., 1999) model to provide a developmental account of CAS (see Figure 2). This model depicts the infant's speech production pathways at the transition from pre-linguistic/nonverbal to early linguistic stage producing their first words or verbal utterances with communicative intent. According to Maassen (2002), there are two major differences between the child's system and the adult's system. The first difference relates to the fewer processing stages resulting in a more simplified model. Figure 2 depicts the adult system on the left and the child's system on the right and the large arrow signifies how the child's speech and language system has not yet attained the different components depicted in the adult model. Consequently, the child is operating on a limited set of components (e.g., conceptualization and articulation). According to Maassen (2002) the child has a restricted set of articulatory forms the size of a syllable at their disposal, which form the protosyllabary. Consequently, early speech attempts are approximations of adult speech that rely on direct mapping between meaning and articulatory routines (McCune & Vihman, 2001), stored as part of the protosyllabary. As development proceeds in TD, due to communicative pressure, the phonological system expands and the word form lexicon and phonological encoding system develop (Maassen, 2002). However, in relation to CAS, these articulatory routines, which in the context of the WEAVER model form the basis of the mental syllabary, consistent with Maassen's (2002) protosyllabary, do not develop in keeping with typical development and consequently speech output is compromised.

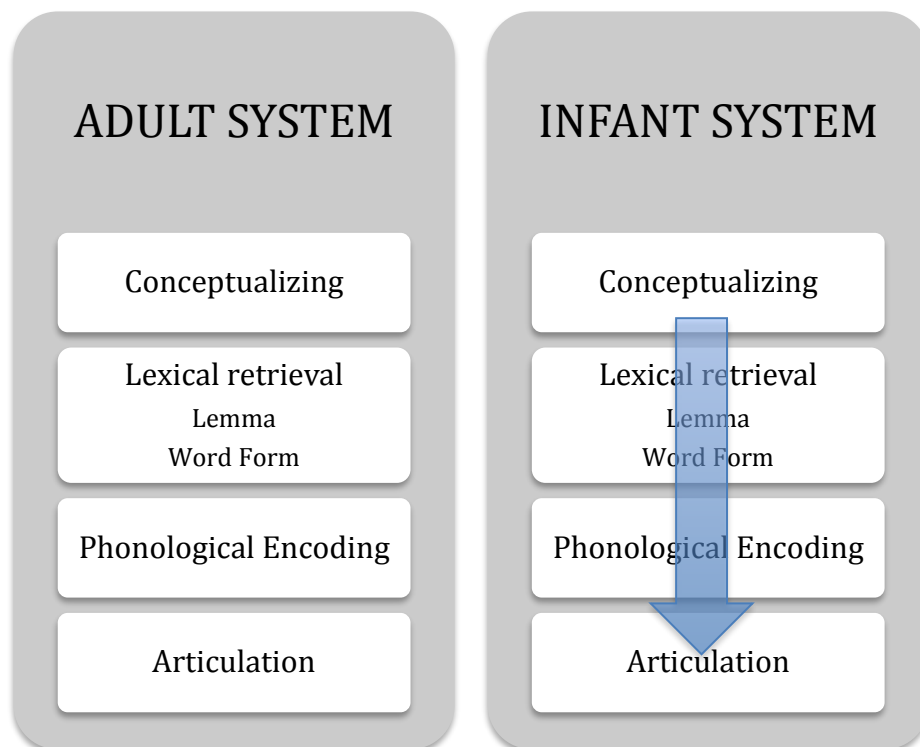


Figure 2

Simplified Model of Speech Production Proposed by Maassen (2002) and Ziegler and Maassen (2004).

The second major difference between the adult and child system relates to association and dissociation. Maassen's (2002) model recognizes that during development the different levels of processing interact resulting in a number of associations between the different levels of processing during this developmental phase. Consequently, a deficit at any one specific level of processing can have a detrimental impact on other levels of processing as a result of this interdependency. The core deficit in CAS is assumed to be at the level of the speech motor system, which in the context of this model is at the level of articulation. There is no assumption that deficits originate in other parts of the system, for example, the input pathway has the potential to develop normally up until the child is transitioning to the linguistic stage of development. The evidence that children with CAS present with receptive language within normal limits and, often stronger, receptive language compared to TD peers is consistent with this view (Lewis, Freebairn, Hansen, Iyengar, et al., 2004).

Computational Modelling of CAS

Computational models have been used to explore both typical and atypical speech production in children and adults (Ballard, Robin, & Folkins, 2003; MacWhinney, 1998; Plunkett, Karmiloff-Smith, Bates, Elman, & Johnson, 1997; H. Terband & Maassen, 2010a; H. Terband, Maassen, Guenther, & Brumberg, 2009; Van der Merwe, 1997). Computational networks learn to perform information processing tasks on the basis of exposure to a training set of items and gradual changes occur to weights on connections between units within the network as a result of this learning (Seidenberg, 1997). These weights control patterns of activation and the resulting behaviors are a reflection of the architecture of the network. Consequently, these models provide a concrete computational basis to interpret the possible underlying causes of a specific deficit during speech development.

The DIVA (Directions Into Velocities of Articulators) model is a neural network computational model that was designed to simulate how infants acquire the speech motor skills required for speech production (Tourville & Guenther, 2011). This model focuses on the sensorimotor transformations that underpin the control of articulatory movements and highlights the importance of the interactions that occur between the different levels of processing during development. Originally described by Guenther (1994), and since modified, the model consists of feedforward and feedback control loops that interact to optimise learning auditory targets (Tourville & Guenther, 2011). The feedforward commands are stored in feedforward projections that result in articulatory trajectories that produce specific auditory targets (Tourville & Guenther, 2011). Once trained, the model takes a phonological code, such as a syllable, as input and generates the desired sound by varying a sequence of articulator positions that command movements of a simulated vocal tract, similar to movements of the jaw, lips and tongue. If the realised auditory signal does not match the target then feedforward control system updates the feedforward command to be more accurate for the next production. The term mapping refers to the transformation from one neural representation to another resulting in audible output, which resembles infant babbling. The articulatory adjustments made by the simulated vocal tract to improve auditory targets are similar to the theoretical maps between acoustic input and articulatory targets that occur in speech development.

Terband, Maassen, Guenther and Brumberg (2009) utilised the DIVA model to explore the underlying deficit in CAS. Four key characteristics of CAS were

evaluated; these included *deviant coarticulation*, *speech sound distortions*, *searching articulatory behaviours (such as groping)*, and *increased variability*. The series of simulations varied the ratio between feed-forward and feedback control to test the prediction that speech motor control was biased in children with CAS towards feedback control. In the typical development simulation the feed-forward/feedback ratio was high, set at 90/10 with slightly lower ratios still corresponding to typical speech development. However, to simulate CAS the ratio was set much lower (e.g., 70/30) and as the ratio decreased severity of the disorder increased. Overall, the CAS simulation results revealed an increase in coarticulation, speech sound distortions, and searching articulatory behaviour (determined by comparing the formant frequencies at the beginning, middle and end of each speech sound). These findings suggest that the symptoms of CAS could be due to over reliance on feedback mechanisms due to deficient feed-forward commands. More specifically the findings demonstrate how a single type of underlying deficit within the speech motor system can give rise to a broad range of symptoms that are typically associated with CAS.

Terband and Maassen (2010c) went on to explore the cause of the degraded feedforward commands by testing two possible hypotheses. The first hypothesis sought to determine if the underlying deficit was due to impaired somatosensory feedback, consistent with the belief that children with CAS have reduced or degraded oral sensitivity. The second hypothesis sought to determine if the deficits observed in CAS could be explained by increased neural noise. The simulation tested two stages of speech development, babbling and imitation learning. In the DIVA model degraded sensitivity, depicted in the model by lack of somatosensory information, would have a detrimental impact on the babbling stage due to underspecified synaptic projections that permit the mapping between motor commands and articulatory goals. For the babbling stage, both deficits resulted in decreased feedforward performance, with greater impairment for the neural noise than reduced/degraded somatosensory information. In contrast, for the imitation learning stage, the effect was larger for the reduced/degraded somatosensory information, indicating that the acquired motor commands were unstable, whereas the neural noise did not lead to unstable motor commands. These findings suggest that deficits observed in children with CAS are more likely to reflect a deficit in somatosensory impairment rather than a neural noise component (Terband & Maassen, 2010).

The DIVA model provides an explanation of CAS as a speech motor deficit, which affects speech output, but which could also give rise to a variety of speech behaviours seen in children with CAS. Maassen (2002) reiterates this view and his developmental model also provides an explanation why higher level linguistic deficits are seen in children with CAS, despite the general consensus being that the underlying deficit is at the level of speech motor programming. Although Maassen's (2002) developmental account of CAS and Terband and Maassen's (2010) implementation of the DIVA model in simulating deficits seen in CAS underscore a speech motor deficit as the core deficit in CAS, distinctly different from a phonological deficit in PD, evidence supporting this point of difference is still limited.

Research Aims and Rationale

The overall goal of this PhD is to examine the deficits observed in CAS and PD from a developmental perspective. Maassen (2002) highlighted the issues comparing a developmental disorder with theoretical accounts derived from acquired disorders, stressing that it is the associations in developmental disorders that are informative about the deficits and the underlying disorder, not just the dissociations. The majority of the research undertaken to date that has compared CAS with other speech sound disorders have neglected to look at the shared deficits from this developmental context. Groenen et al. (1996) looked at the relationship between input and output deficits in children with CAS and TD, however, their findings were inconclusive and the exact nature of the relationship between speech perception and production remains ambiguous. Nijland (2009) also looked at speech perception deficits in children with output speech disorders, which included children with CAS and PD, however, findings were also inconclusive, and the authors contended that further research was needed that addressed the complex relationship between the different levels of processing. To date, no research has looked at the predictive relationships between speech motor and phonological levels of processing in children with CAS or compared these relationships with another speech sound disorder, such as PD.

This thesis examined whether the relationship between the speech motor ability and phonological competence differed in children with CAS and children with PD. We planned to do this by initially developing and validating measures of

phonological competence and speech motor ability, with a view to profiling the shared deficits in CAS and PD at both levels. We then planned to investigate the relationships between speech motor ability and phonological competence in children with CAS and PD using regression analysis, based on the similar research paradigm implemented by Munson et al. (2005b). The assumption in the regression models is that the covariance or strength of relationship between measures that target different domains or levels of development, where one constrains the development of the other, will be proportional to the degree of constraint (Portney & Watkins, 2011). Therefore, this paradigm has the potential to reveal differences in the source of the deficit in CAS and PD by comparing the two groups in terms of the strength of the relationships between speech motor measures, the hypothesized source of deficit in CAS specifically, and measures of phonological competence, the level of processing affected through developmental constraint. In particular, if speech motor deficits constrain the development of higher-level linguistic/phonological units in CAS but not PD, with the phonological deficits in PD originating at a higher linguistic-phonological level (see Pennington & Bishop, 2009, for similar discussion), then it is predicted that measures of speech motor ability will predict outcome measures of phonological competence in children with CAS to a greater extent (i.e., account for more variance) than in children with PD and TD. The influences on development are potentially bi-directional, and a phonological deficit could arguably affect the development of the speech motor control system to some degree. However, it is argued that a higher level phonological deficit will not cause an underlying impairment to the speech motor control system and that the degree of constraint between development of speech motor ability and phonological competence will be less for PD (assuming the deficit originates at the phonological level) compared to CAS.

Prior to looking at the interdependency between speech motor measures and measures of phonological competence in children with CAS and PD, the tasks were used to evaluate different aspects of phonological competence and speech motor development in children with TD of a similar age. This was the focus of Study 1, which is described in the next chapter.

Chapter 2

Developmental Changes in Phonological Competence and Speech Motor

Abilities in Children with Typical Development

Introduction

The purpose of Study 1 was to develop and pilot a set of tasks designed to target the different components of phonological competency and speech motor ability in children with TD. Our goal was, firstly, to develop tasks that examine input phonology, the abstraction of phoneme categories, output phonology, as well as tasks that could be used to evaluate the efficiency of on-line processes of speech motor planning and execution. Secondly, to establish whether these measures were valid indicators of developmental change in younger and older children with TD. Thirdly, utilizing Munson et al.'s (2005) paradigm, regression analysis was used to determine if (a) if vocabulary predicted phonological competence, over and above that predicted by age, consistent with previous findings (Edwards, Beckman, & Munson, 2004; Munson, Edwards, et al., 2005b), and (b) whether speech motor measures predicted measures of phonological competence independent of the contribution made by vocabulary and if this differed for the younger and older children. Our primary goal was to examine the relationship between speech motor competence and phonological ability as a function of normal development and in doing so provide a useful benchmark for understanding the patterns that might emerge using the same paradigm in children with SSD.

Evaluating Phonological Competence

The tasks implemented targeted a broader view of phonological knowledge by targeting both input and output phonology, consistent with the view that input and output pathways have separate phonological representations (Monsell, 1987; Roelofs, 1997b). To do this we implemented a speech discrimination task and a picture-naming task with phonological primes. We also assessed the quality of underlying abstract phonological representations using a nonword repetition task.

The ability to repeat nonwords accurately has been used to evaluate phonological development in children with typical development and children with PD (Edwards et al., 2004; Gathercole, 2006; Metsala, 1999; Munson, 2001a; Munson, Swenson, & Manthei, 2005). A nonword repetition task was used to establish the degree of phonological development, consistent with the view that ability to repeat nonwords accurately is a reflection of the quality and abstractness of underlying (sub-lexical) phonological representations. In addition, by manipulating the phonotactic frequency of sequences embedded within the nonwords and

comparing performance on high and low frequency sequences, provided further insight into the abstractness of these underlying phonological representations, consistent with the view that high frequency sequences are stored earlier than low frequency sequences. The nonword repetition task implemented in Study 1 replicated Edwards et al. (2004).

Picture-naming tasks have been used extensively to evaluate lexical retrieval processes during speech production (Cutting, Ferreira, Damian, & Martin, 1999; Damian & Martin, 1999; Jerger, Martin, & Damian, 2002; Levelt et al., 1999; Newman & German, 2002; Swan & Goswami, 1997; Truman & Hennessey, 2006). Picture-naming includes three main processing stages, prior to articulation: object identification, lexical access and phonological encoding (Brooks & Mac Whinney, 2000). In the context of the WEAVER model these stages relate to identification of the lexical concept, lemma retrieval and word form encoding. To specifically target phonological output representations a phonological priming paradigm was implemented. Previous research has shown that phonologically related auditory primes (i.e., primes that share phonemes with the picture to be named) enhance picture-naming resulting in a facilitation effect (i.e., a faster reaction time), relative to unrelated phonological prime (Brooks & Mac Whinney, 2000; Jerger et al., 2002; Meyer, 1991; Schriefers, Meyer, & Levelt, 1990; Truman & Hennessey, 2006). The phonological facilitation effect is presumed to be located at the stage of retrieving phonological representations during phonological encoding (Levelt et al., 1999). This is consistent with Dell's (1986) theory of spreading activation, whereby the phonemes that make up the phonological syllable are selected based on activation levels within the network. Consequently, when an auditory prime is heard that matches the phoneme(s) of the target word, this results in heightened activation for the target phoneme and faster retrieval and encoding of the target word. The magnitude of the phonological priming effect in the picture-naming task has the potential to assess the quality of output phonological representations and phonological encoding efficiency. In particular, phonological encoding that is less efficient is likely to benefit to a greater degree from the additional activation of phonologically related primes than a more efficient phonological encoding system. This assumption has been confirmed, for example in children with dyslexia who have phonological deficits (Truman & Hennessey, 2006). The picture-naming task

used in Study 1 replicated Jerger et al.'s (2002) picture-naming task with auditory primes.

Speech discrimination relates to the ability to attend to fine phonetic detail of the speech signal and has been shown to be directly linked to speech and language acquisition (Tsao, Liu, & Kuhl, 2004). Speech discrimination deficits have been demonstrated in children with TD revealing that younger children perform more poorly than older children and adults when the speech signal is degraded in some way (Munson, 2001b; Walley, Michela, & Wood, 1995). Walley, Michela and Wood (1995) found that children with typical development (kindergarten and grade 1) needed more acoustic information than adults to recognize words accurately. Similar deficits have been observed in children with PD when compared to children with TD (Edwards et al., 1999; Edwards et al., 2002). Edwards et al. (1999; 2002) found that children with PD had greater difficulty identifying words that differed only in the final consonant (e.g. "cap" and "cat") when the final consonant was deleted. The younger children with TD in Edwards et al.'s (2002) study also performed more poorly, compared to the older children and adults. This suggests that children with PD perform in a similar way to younger children with TD and indicating that the input phonological representations of children with PD and younger children with TD are not as well defined as the representations of older children with TD (Walley, 1993). Furthermore, vocabulary size has been shown to predict speech discrimination performance using the same task as used in the thesis after controlling for age (Edwards, et al., 2002; Munson et al., 2005). This suggests that speech discrimination ability is sensitive to higher level phonological development and not just differences in auditory or phonetic processing. A speech discrimination task replicating Edwards et al. (2002) was implemented to assess the quality of input phonological representations in younger and older children with TD.

Evaluating Speech Motor Ability

The Goldman Fristoe Test of Articulation (Goldman & Fristoe, 2000) was used to measure articulatory accuracy. Measures based on assessing accuracy of articulation, while sensitive to speech motor disorders, can be confounded by deficits at the level of output phonological knowledge. Therefore to directly assess phonetic planning and the execution of these speech motor plans we implemented additional tasks that targeted speech motor ability.

A simple verbal reaction time (SVRT) and a choice verbal reaction time (CVRT) paradigm were used to target the efficiency of on-line processes of speech motor planning and execution processes (Klapp, 1995, 2003; Sternberg, Knoll, Monsell, & Wright, 1988). The paradigm was based on previous research, for example by Klapp (1995), and involves naming the same pictures under two conditions. In the SVRT condition each individual picture is repeatedly presented for speeded naming in a separate block of trials, therefore the response is known in advance. In contrast to the SVRT, in the CVRT condition there are two alternative responses in each block of trials, consequently, the response is not known in advance of the target stimulus. Studies have shown that the SVRT for single word responses are not affected by the complexity of the response, such as word length, consistent with the premise that the verbal response is already programmed and planned in advance and temporarily stored in the articulation buffer prior to execution (Levelt, 1999; Roelofs, 1997b). Simple response time, therefore, excludes all preceding levels of linguistic processing (e.g., phonological and phonetic planning), but will be sensitive to late stage muscle command preparation and execution processes needed to translate the speech motor program into overt movements of speech (see, also Sternberg et al., 1988).

In the CVRT task, the target response cannot be preplanned because it is not known in advance (i.e., two fully encoded responses cannot be stored simultaneously in the articulatory buffer (Levelt, 1999; Roelofs, 1997b). Consequently, reaction time for the CVRT condition includes the phonetic planning processes. This assumption is supported by finding CVRT is affected by response complexity of the target response, such as length, consistent with the view that planning time for longer responses takes longer thereby increasing CVRT (Klapp, 1995, 2003; Sternberg et al., 1988). It is also assumed that because there are only two available responses during a single block of trials, the phonological codes for each response will be pre-activated and maintained in a phonological short-term or working memory store ready for phonetic encoding depending on the stimulus presented. This is based on the assumption that two items will be within the phonological working memory span of children (Baddeley, 1986, 1993, 2003). Therefore the CVRT interval will not be confounded by the time taken to retrieve lexical phonology from the output lexicon because the phonological codes will have already been retrieved. To verify the CVRT was sensitive to phonetic planning, and not sensitive to phonological

retrieval, we manipulated word length by comparing one and two syllable words, and word frequency by comparing words with low and high frequency of occurrence. We predicted longer words would reveal a longer reaction time for the CVRT because more phonetic planning time was required. Given evidence that word frequency effects are associated with accessing lexical phonology (Levelt et al., 1999), we predicted that if the CVRT included phonological retrieval then a word frequency effect should be observed. Otherwise, an absent word frequency effect would help verify the assumption of minimal involvement of phonological retrieval processes during the CVRT task. We also predicted that reaction time would not differ for the SVRT as a function of word length or word frequency. More importantly however, we expected these measures to be sensitive to individual differences in the efficiency of speech motor planning and execution processes, such as between children who are at different stages in the development of their speech motor control system. Younger children with less developed articulatory systems should demonstrate slower SVRT if their speech plan to execution processes are less efficient, and slower CVRT and a larger length effect if their speech motor planning or phonetic encoding processes are less efficient.

In summary, the goal of Study 1 was to compare younger (preschool) and older (early primary school-aged) children with TD. The age range of the children in Study 1 spans the stage of development during which ongoing refinement of phonological encoding and speech motor skills occur that permitted a direct comparison of potential changes in processing efficiency during this period. We had a number of research questions that we wanted to address; do younger and older children with TD differ in processing efficiency at the different levels targeted by the tasks implemented; does the interdependency between these measures change over the course of development; does vocabulary predict nonword repetition accuracy and the phonotactic frequency effect, over and above that predicted by age, replicating Edwards, Beckman and Munson (2004); does vocabulary predict picture-naming reaction time and the phonological facilitation effect, over and above that predicted by age, not yet explored in the research literature; do younger children need more acoustic information to discriminate between two words that differ only in relation to the final consonant; and does articulatory ability and on-line measures of speech motor control predict phonological competence in children with TD and does this relationship differ with age.

Hypotheses

There were a number of different hypothesis derived from the research questions. These are listed in relation to the experimental tasks, each of which have a number of hypotheses. The hypotheses relating to the regression analysis are listed.

Nonword repetition:

- For the nonword repetition task we hypothesized that the younger children would be less accurate than the older in overall nonword repetition accuracy.
- We also hypothesized that the phonotactic frequency effect (i.e., difference in repetition accuracy between the high and low frequency sequences) would be larger in the younger children than the older children, indicative of less abstracted phonological representations in younger children and consistent with previous findings (Edwards et al., 2004).
- We also hypothesized that the nonwords that contained high frequency sequences would be repeated more accurately than the nonwords that contained the low frequency sequences.

Picture Naming:

- For the picture-naming task we hypothesized that the younger children would be slower at overall picture-naming than the older children.
- The younger children would demonstrate a larger phonological facilitation effect compared to the older children, based on the premise that younger children have less efficient phonological encoding. This is consistent with the expectation that the younger children would benefit to a greater extent than the older children from hearing a related prime, resulting in enhanced facilitation (Brooks & Mac Whinney, 2000).

Speech Discrimination:

- For the speech discrimination task we hypothesized the younger children would reveal poorer speech discrimination accuracy than the older children, consistent with previous research that has shown that younger children require more acoustic information than older children to accurately discriminate between words when the acoustic signal is degraded (Edwards et al., 2002).

Speech Motor Measures

- For the speech motor measures we hypothesized that the younger children would have a slower SVRT than the older children, given the preplanned

response needs to be unpacked for execution and these skills are likely to be less developed in younger children, consistent with the view that speech motor programming skills continue to develop throughout childhood and into adolescence (Kent, 2000).

- We hypothesized that the younger children would demonstrate a slower CVRT than the older children, indicative of less developed speech motor skills and there would be an effect of word length for the CVRT that was larger for younger children. This hypothesis was based on the view that the entire word needs to be phonetically encoded prior to execution, with longer words taking longer to encode and therefore resulting in a slower reaction time and that this process would take longer in younger children with less efficient phonetic encoding (Markus, Bowers, Stadthagen-Gonzalez, & Spalek, 2010).
- Given each of the picture names for the CVRT had to be programmed prior to articulation for each trial, we expected that the longer words would take longer to program resulting in a word length effect, consistent with previous research and validating this task as an accurate measure of speech motor programming efficiency (Klapp, 2003; Sternberg et al., 1988).
- Assuming CVRT is not confounded by lexical phonological retrieval processes prior to phonetic encoding, we expected that there would be no difference in CVRT between low and high frequency words.

Regression Analysis:

There were a number of additional hypotheses in relation to the regression analysis, however, the degree to which speech motor development predicts phonological development in TD is not known and given the exploratory nature of a number of our research questions in this PhD, some of these hypotheses are speculative. Developmental changes in the degree of interaction between the different levels of processing become more encapsulated or modularized over the course of development, suggesting levels of processing are more interdependent earlier in development (Kamiloff-Smith, 1998). In keeping with this perspective, our overall hypothesis was that speech motor measures would predict phonological competence in the younger children to a greater extent than the older children.

- We hypothesized that vocabulary would predict nonword repetition accuracy and the phonotactic frequency effect over and above that predicted by age (Edwards et al., 2002; Munson et al., 2005b).
- We also hypothesized that vocabulary would predict the picture-naming reaction time and the phonological facilitation effect, over and above that predicted by age, consistent with the understanding that vocabulary is the driving force behind phonological development and this may extend to phonological encoding processes during speech production (Edwards et al., 2004; Metsala, 1999).
- Furthermore, we hypothesized that vocabulary would predict speech discrimination ability over and above that predicted by age, also consistent with previous findings (Edwards et al., 2002).
- We hypothesized that articulation ability, as measured by the GFTA (Goldman & Fristoe, 2000), would predict nonword repetition accuracy, and speech discrimination ability, but not the phonotactic frequency effect, consistent with previous findings (Edwards, 2004).
- We also hypothesized that the GFTA would predict picture-naming reaction time and the phonological facilitation effect, and this would to a greater extent in the younger children than the older children.

Method

Participants

The participants were 47 children with typical development, ranging in age from 5 to 8 years. The children were divided into two groups; the younger group comprised 13 girls and 10 boys enrolled in preprimary and year 1 ($M = 5.7$ years, $SD = 6$ months; age range = 5.1 to 7.2 years) and the older group comprised 19 girls and 5 boys enrolled in year 2 and year 3 ($M = 8.1$ years, $SD = 7$ months; age range 7.2 to 9.2 years). All children were recruited from a Perth metropolitan public primary school. The Expressive Vocabulary Test (EVT; Williams, 1997) and the Peabody Picture Vocabulary Test-4 (PPVT-4; Dunn, 2007) were used to measure expressive and receptive vocabulary and the Sounds in Words Subtest of the Goldman Fristoe Test of Articulation 2 (GFTA-2; Goldman & Fristoe, 2000) was used to assess articulation ability. All children received a standard score greater than 85 on the

EVT, PPVT and GFTA indicating performance within normal limits (refer to Table 1). None of the children had a history of speech, language or hearing impairment, according to parent report. All children that were recruited were identified as progressing normally through school, according to feedback from each child's teacher. All participants had English as a first language.

An independent samples *t*-test revealed that groups differed significantly on chronological age, $t(45) = 16.36, p < .001, \eta^2 = .85$. Although the groups differed significantly on standard scores for the EVT, $t(45) = 2.72, p = .009, \eta^2 = .14$, with the younger children having a significantly higher standard score compared to the older children (see Table 1), the mean raw scores indicated the expected developmental difference between these two groups. Likewise, the younger children had a higher mean GFTA raw score than the older children, $t(45) = 6.03, p < .001, \eta^2 = .45$, indicating more articulation errors on average for the younger children compared to the older children. However, it appears as though the younger children had better articulation skills relative to the older children for their age, indicated by the higher standard score. Groups did not differ on PPVT standard score, $t(45) = 0.031, p = .975, \eta^2 = .00$.

Table 1

Age and Test Scores for Younger and Older Children

		Younger (<i>n</i> = 23)		Older (<i>n</i> = 24)	
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Age in months		67.1	5.7	96.7	6.6
EVT	Standard Score	105.0	5.8	100.0	6.5
	Raw Score	63.0	6.8	79.0	9.3
PPVT	Standard Score	106.0	9.1	106.0	9.3
	Raw Score	107.0	16.1	135.0	14.2
GFTA	Percentile Rank	64.0	16.7	49.0	5.5
	Standard Score	110.0	4.3	104.0	1.4
	Raw Score	1.2	2.1	>1	>1

Note. GFTA 2= Goldman-Fristoe Test of Articulation 2; EVT = Expressive Vocabulary Test; PPVT-4 = Peabody Picture Vocabulary Test – 4.

Standardized Assessments

Expressive Vocabulary Test (EVT). The EVT (Williams, 1997) is a well-used individually administered instrument that measures expressive vocabulary and word retrieval in children and adults. It is norm referenced on participants age 2 through to 90 years and co-normed with the PPVT to enable comparisons between expressive and receptive vocabulary. The EVT measures expressive vocabulary knowledge including labeling and synonym knowledge. The examiner presents a picture to be named and a stimulus word or words within a carrier phrase and the examinee responds with a one-word answer that is a noun, verb, adjective or adverb. Two unscored answers are presented before the test items are presented. The EVT has been shown to have high internal consistency and test-retest reliability with corrected coefficient of 0.77 for ages 2.6 to 5.11 years and corrected coefficient of .84 for ages 6.0 to 10.11 (Williams, 1997). EVT results can be reported as standard scores (with a mean of 100 and a standard deviation of 15).

Peabody Picture Vocabulary Test 4 (PPVT-4). The PPVT-4 (Dunn, 2007) evaluates comprehension of spoken words in Standard English and is therefore a measurement of receptive vocabulary of children and adults. The 228 items are grouped into 19 sets of 12 items each, which are arranged in order of increasing difficulty. This permits the examiner to administer only sets appropriate for the examinees vocabulary level using basal and ceiling rules. The PPVT is norm referenced on participants age 2 through to 90. The PPVT is a well-used standardized assessment and has been shown to have high internal consistency and test-retest reliability with corrected reliability coefficients of .93 for ages 2.0 to 4.0 years and 7.0 to 10 years and .92 for ages 5.0 to 6.0 years (Dunn, 2007). PPVT results can be reported as standard scores (with a mean of 100 and a standard deviation of 15). The PPVT also provides an estimate of verbal intelligence, thereby giving an indication of language ability and verbal IQ (Dunn, 2007).

Goldman-Fristoe Test of Articulation-2 (GFTA-2). The GFTA-2 assesses children's articulation abilities. The GFTA-2 is normed on participant's aged 2 through 21. The GFTA-2 Sounds in Words subtest uses 34 picture plates and 53 target words to elicit the articulation of 61 consonant sounds in the initial, medial and final position and 16 consonant clusters (blends) in the initial position. The number of errors on Sounds-in-Words can be converted to a standard score using separate normative tables for females and males. The standard score has a mean of 100 and a

standard deviation of 15. The GFTA-2 has undergone extensive reliability and validity tests. It has been shown to have high internal reliability, inter-rater reliability and test-retest reliability.

Stimulus Materials

Nonword repetition. The stimuli for the nonword repetition task consisted of the 22 2-syllable nonwords and 22 3-syllable nonwords used by Munson et al. (2005b). At each syllable length, one half of the nonwords contained a low phonotactic probability sequence and the other half contained a high phonotactic probability sequence. The sequences were embedded in the same position within larger word shapes and had the same stress and syllable structure. An example of the 2-syllable low and high frequency nonword pairs is /jugoin/ and /bogib/, with the target sequences underlined. An example of the 3-syllable low and high frequency nonword pairs, respectively, is /dugnəted/ and /tagnədɪt/. Practice items were included to permit participants to practice repeating nonwords prior to the test trials. Practice stimuli were four 2-syllable and four 3-syllable nonwords, similar in complexity to the target nonwords. A spoken version of each nonword to be used as production prompts were recorded by an Australian female speaker and trained speech pathologist, following the phonetic transcriptions provided by Munson et al. (2005b). Each nonword was recorded at a sampling frequency of 44100 Hz using PRAAT (Boersma & Weenink, 1995) software and intensity set at 65 dB with a lead in time of 15 ms. For a complete list of the stimuli refer to Appendix A1.

Picture-naming task. The picture-naming task stimuli consisted of 18 digitized black and white photographs of everyday objects. Pictures were all black and white line drawings, very similar, but not matched for visual complexity. The name of all pictured objects were concrete nouns (e.g., *goat*, *hammer*) that were chosen to be familiar to children within the range of age of the children in this study (5-9 years) based on age of acquisition data from MRC Psycholinguistic Database (Coltheart, 1981). Items included nine 1 and 2 syllable words with a maximum word frequency of 9 ($M = 5.3$, $SD = 2.3$) occurrences per million (Kučera & Francis, 1967). Words with a low frequency of occurrence were selected since words with lower frequency are retrieved more slowly than words with high frequency and therefore differences in reaction times are more likely to be detected between groups (Newman & German, 2002). Phonological primes consisted of two priming words

formatted for each picture name target word, one related and one unrelated. Related auditory priming words shared the onset consonant and vowel with the target (e.g., target = *cage*, related prime = *case*) whereas unrelated auditory primes did not share onset consonant or vowel with the target (e.g., target = *sock*, unrelated prime = *pet*). Related and unrelated primes were matched in relation to word frequency and length. Independent groups t-test showed no difference in Kucera and Francis (1967) word frequency and number of phonemes and syllables between the two sets of items. Two practice items, each with a related and unrelated auditory prime, were included to permit participants to practice naming pictures under each of the conditions prior to the test trials. All priming words (36 in total), including practice primes (four in total) were digitally recorded at a sampling frequency of 44100 Hz in 16 bit by a female Australian adult speaker using PRAAT. Each sound file was edited to ensure a lead-time of 15 ms prior to onset and finishing at word offset. For a complete list of the stimuli refer to Appendix A2.

Speech discrimination task. Stimuli for this task were taken from Edwards et al. (2002). These comprised two pairs of minimally contrastive words *cap* and *cat* and *tap* and *tack*, which were expected to be familiar to young children and also able to be represented by pictures. The low front vowel /æ/ was chosen to ensure substantial formant transitions from the medial vowel to the final consonant (Edwards et al., 2002). Ten repetitions of each word were digitally recorded using PRAAT. The speaker was a female speaker of Australian English and was instructed to release the final stop consonant of each repetition. Two tokens were selected that were most similar in relation to acoustic duration and overall quality as determined by visual inspection using PRAAT. Two backward-gated versions of these tokens were then created by removing a portion of the end of the speech signal. For the short backward-gated version the stop-gap and release burst was removed from the end of the word, and for the long gated version the formant transition (post vowel steady state), stop-gap and release burst were removed. The start of the formant transition was identified by visual inspection of a wide-band spectrogram and formant analysis. A total number of 24 items were used as stimuli: three versions (i.e., whole word, short gate and long gate) of two tokens of each of the four words.

Simple and choice verbal reaction time task. The same stimuli were used for the SVRT and CVRT tasks. Stimuli consisted of eight matched pairs of 1 and 2 syllable words that included four low frequency pairs, less than nine occurrences per

million, and four high frequency pairs, greater than 20 occurrences per million (Kučera & Francis, 1967). Word frequency was included as a control variable. In each word pair the 1 syllable and 2 syllable words shared the same first syllable (e.g., *pig/piglet*, *cart/cartwheel*). Words were all concrete nouns represented by a picture. For a complete list of the stimuli refer to Appendix A3.

Procedure

Each child was tested individually in a quiet room at school to minimize distractions. Participation took place over two 1 hour sessions. The standardized assessments were administered in the first session, in a fixed sequence designed to vary the cognitive demands across the successive tasks: GFTA-2, EVT and PPVT-4. The experimental tasks were administered in the second session in the same order for each participant; SVRT, CVRT, speech discrimination task, picture-naming task and nonword repetition task. All test items that required naming of pictures were presented prior to administration of each individual task to ensure children could name all the pictures presented. For tasks that required picture presentation and measurement of picture naming RT, pictures were presented on a Sony VAIO laptop computer screen using DMDX software (Forster, 1997) and the internal voice key in the DMDX was used to measure verbal RT. Participants wore a Logitech head mounted microphone connected to the computer for software to detect voice onsets and the entire verbal responses was automatically recorded to the hard disk for later error analysis and, in the case of the nonword repetition task, for scoring. Participants were given a break as needed during both sessions to maximize performance and minimize fatigue. Participants were also given general feedback and verbal encouragement during test trials as required (e.g., “great job”, “keep going”, “almost finished”). Encouragement was also provided on reaction time tasks for the child to respond as fast as they could. Each child was awarded with participation stickers and a lucky dip prize on completion of both sessions.

Nonword repetition task. The two syllable words were always presented in a separate block of trials directly prior to the three syllable words for all participants. The sound files were presented in a different random order to each participant using DMDX software (Forster, 1997). Stimuli were played at a comfortable volume from the hard drive of a laptop computer through the headset connected to the computer were recorded. On each test trial playing the sound file also started the digital audio

capture via the headset microphone and the recording continued for a further 3500 ms after each word was presented. There were four practice trials prior to the test trials. When the practice trials were completed the instructions were repeated as needed prior to commencement of the test trials.

Picture-naming task. The pictures to be named were presented in the middle of the computer screen on a white background and were approximately 8cm by 6cm in size. The computer screen was mounted directly in front of the child at a distance of approximately 60 cm. The 18 test items were randomly split into three sets of six, with each set allocated to a different priming condition: phonological related prime, unrelated prime and no prime (i.e., silence). The items were presented in three cycles with the set rotated across conditions so that each item was presented once in each condition and each item appeared only once in a cycle. The items within each cycle were randomly presented each time it was run, therefore controlling for any order effect of those cycles across children. Each test item was presented in each condition, once with phonologically related prime, an unrelated prime and with no prime (i.e., silence). Within each test trial auditory primes were presented at 116 ms after picture presentation. This stimulus onset asynchrony (SOA) was chosen based on previous findings that suggested that the largest effect of phonological facilitation were found between 0 and +150 ms for children of a similar age (Brooks & MacWhinney, 2000). Only one SOA was implemented to limit the number of trials and length of the task. The picture disappeared from the screen when triggered by voice input detected via the headset microphone, or following a time out period, set at 3500 ms. The wait time after each picture disappeared from the screen was set at 4000 ms to ensure the vocal response was recorded before the next trial started.

Each participant named all test items prior to the commencement of the test trials. A practice phase with auditory primes was undertaken to familiarize the participants with the task and further instruction was provided as needed, prior to the commencement of the test trials. The practice phase included phonologically related primes, phonologically unrelated primes and the silence condition to ensure the participant understood the task. Participants were instructed to name the picture on the computer screen as quickly as possible and ignore what they heard via the headphones. They were also told that sometimes they would not hear anything at all.

During the test trials the picture disappeared from the screen when triggered by voice input, in the absence of a voice input it disappeared from the screen at a time

out set at 3500 ms. The audio input used to trigger the voice key was automatically recorded on each trial, which was later checked for accuracy. Reaction time was measured from onset of the verbal response. The wait time after each picture disappeared from the screen was set at 4000 ms to ensure the response was recorded in its entirety.

Speech discrimination task. Stimuli were presented in two blocks of trials, one for the *cap* and *cat* pair, and the other for the *tap* and *tack* pair. The order of presentation of the two blocks of trials was counterbalanced within each age group with one half of the participants receiving the tap and tack pair first. For each test trial within each block the two pictures were presented on the computer screen (e.g., the picture of a *cap* and *cat* or *tap* and *tack*). Each participant had 48 tokens in total to identify, 24 for each word pair. Auditory stimuli were presented via Logitech headphones. Practice trials consisted of 12 items that comprised of each word presented for each condition.

Participants were instructed to point to the picture on the computer that they heard via the headphones. They were instructed that sometimes the end of the word would be missing and this would make it difficult to hear what word had been said, however the child was instructed to make a choice between the two pictures. To counterbalance for participants preference pointing to a picture on one side of the screen each of the tokens were presented on both sides and equal number of times. When the child indicated which picture they heard by pointing the experimenter entered the child's response by clicking either the left or right mouse button.

SVRT. For the SVRT task each participant had to name eight pictures comprising four 1-syllable words (e.g. *foot*, *news*, *cart* and *pig*) and four 2-syllable words (e.g. *football*, *newspaper*, *cartwheel* and *piglet*). The 1 and 2-syllable words shared the same onset syllable (e.g., *foot* and *football*, *news* and *newspaper*). Stimuli can be seen in Table A3. A practice phase of two test items was presented prior to commencement of the test trials. After the short practice phase each test item was presented six times. This number was chosen after piloting the task showed that the task was too long for the younger children and performance was compromised with more trials. Half of the participants received the one-syllable test items first and the other half received the two syllable test items first. Each picture was presented six times so that each participant had to name 24 one syllable words and 24 two syllable words, 48 test items in total. Each participant was instructed that they would see the

same picture six times and were instructed to name the picture as fast as possible. The participants knew the response in advance, however, inter-stimulus interval (ISI) was varied to ensure the onset of the stimulus was not predictable, therefore preventing the participant preempting response initiation. The picture disappeared from the screen when triggered by voice input, alternatively in the absence of a voice input it disappeared from the screen at a time out, set to 2,500 ms.

CVRT. For the CVRT two pictures were randomly presented within each test trial (e.g. *foot* or *news*) so that each participant did not know which test item would appear. The test items were the same test items used for the SVRT and can be seen in Appendix A3. The one-syllable words were presented in the one block and two syllable words were presented in the another block. Each test item was presented six times so that a block for the CVRT consisted of 12 test items (six trials for each word). Each participant was instructed that they would see one of two pictures and they were instructed to name the picture as fast and accurately as possible. A set of practice trials comprising four test items (two of each test item) was presented prior to the test trials.

Scoring of Dependent Measures

Accuracy of nonword repetition was scored following Edwards et al (2004). Segmental accuracy was calculated for each of the two phonemes in the target sequence (CC, VC, or CV). For consonants, one point was awarded for correct place, correct manner and correct voicing. For vowels, one point was awarded for correct production in terms of tongue advancement (i.e., front, central or back), one point for correct height (i.e., high, mid or low) and one point was awarded for correct length (i.e. short/long vowel, diphthong). Therefore a maximum of three points could be awarded for each vowel and consonant, with the total possible accuracy score being six points for each target sequence within each word. Consistent and identifiable errors were treated as errors. The outcome measures were mean nonword repetition accuracy within each condition and the phonotactic frequency effect, that is, the difference in accuracy between high and low frequency sequences, averaged across two and three syllable words.

For the picture-naming task only correctly named targets were included in the analysis. We trimmed for outliers removing responses that were two standard deviations above or below mean reaction time for each participant for each

condition. This is consistent with common practice in reaction time research and recommended by Ratcliff (1993). Outcome measures from the picture-naming task included mean picture reaction time in the unprimed condition in milliseconds (ms), the phonological facilitation effect (i.e., the difference in mean reaction time between related and unrelated auditory priming conditions) and mean picture-naming accuracy (percent correct). For the speech discrimination task d-prime values were used as the dependent variable for all the statistical analyses undertaken, replicating Edwards et al. (2002). d-prime is a measure of how much the participant is responding to the stimulus versus using a response strategy that does not relate to the stimulus. A d-prime of 0.0 indicates that the participant is performing at chance level and a d-prime of 1.0 indicates that the participant is performing one standard deviation above chance (Edwards et al., 2002). For the SVRT and the CVRT the outcome measures were mean reaction time for pictures named correctly. Outliers were defined in the same way as before, responses were excluded that were two standard deviations above or below the mean reaction time for each participant for each condition. Errors were not analysed, as there were very few.

Data Analysis

Data analysis was undertaken in two stages. The first stage used General Linear Mixed Modeling (GLMM) to explore age group differences on each of the tasks that assess phonological competence (nonword repetition, picture-naming and speech discrimination) and speech motor ability (SVRT and CVRT tasks). We then examined the predictive relationship on outcome measures of phonological competence obtained from each of the tasks by measures of articulatory ability (GFTA-2 raw score) and SVRT and CVRT, and whether they differed as a function of age group. This second stage used a series of hierarchical regression analyses, also with GLMM, to test predictors of phonological competence.

For each task, GLMM was used to test within and between group fixed effects and interactions, similar to factorial design in ANOVA. GLMM were implemented through SPSS's (Version 22) GENLINMIXED procedure. The GLMM represents a special class of regression model that is 'generalized' in the sense that it can accommodate a variety of outcome variables including those with markedly non-normal distributions. In addition, it also has the advantage that the analysis can handle missing data without excluding participants or imputing missing values. The

GLMM is ‘mixed’ in the sense that it includes both random and fixed effects. For the hierarchical regression GLMMs, there was one nominal random effect (participant), one scale fixed effect (the Primary Predictor, i.e., motor measures), one nominal fixed effect (age), and the Age x Primary Predictor interaction. In order to reduce colinearity problems caused by strong associations between Primary Predictor and the Age x Primary Predictor interaction, the Primary Predictor was standardized (centered on zero) before multiplying it by the age variable to create the interaction term. Motor measures included GFTA raw score, SVRT and CVRT. Potential control variables were the EVT and PPVT raw scores. If any of the control variables were correlated with the outcome measure, and therefore had the potential to confound the relationship between the Primary Predictor and the outcome, they were included as fixed effects in the GLMM. The GLMM ‘robust statistics’ option was invoked to accommodate violations of normality, since normality of the Primary Predictor is not a requirement of the GLMM.

Compared to the traditional least squares regression approach, the GLMM maximum likelihood regression approach adjusts standard errors and *p*-values to account for model violations. The two approaches, however, converge on the same values for the regression parameters. Parameters omitted from the maximum likelihood output (namely, the standardized regression coefficient, the part-correlation, and the multiple correlation coefficient) were therefore taken from the least squares output.

Results

Nonword Repetition

Age differences in the mean accuracy of high and low frequency sequences from the nonword repetition task were examined using a three-way mixed design GLMM with group as a between groups independent variable with two levels (younger vs. older groups) and frequency of diphone sequence with two levels (low vs. high phonotactic frequency sequences) and length of nonword with two levels (2 vs. 3 syllable nonwords) as repeated measures independent variables. As can be seen in Table 3 the younger children ($M = 94\%$, $SEM = 1\%$) showed poorer performance overall on nonword repetition accuracy when compared to the older children ($M = 96\%$, $SEM = 1\%$). The low frequency sequences ($M = 93\%$, $SEM = 1\%$) were repeatedly less accurately than the high frequency sequences ($M = 97\%$,

$SEM = 1\%$). The two syllable nonwords were repeated at the same accuracy as the three syllable nonwords ($M = 95\%$, $SEM = 1\%$). Refer to Table 2.

The three-way GLMM model revealed a significant main effect for group, $F(1, 180) = 6.62$, $p = .011$, partial $\eta^2 = .13$, and a main effect for frequency, $F(1, 180) = 64.51$, $p < .001$, partial $\eta^2 = .58$. The effect of length was not significant, $F(1, 180) = 0.77$, $p = .380$, partial $\eta^2 = .017$. The interaction between group and frequency just failed to reach significance, $F(1, 180) = 3.86$, $p = .051$, partial $\eta^2 = .077$. The remaining interactions were all non-significant with small effect sizes: group by length interaction, $F(1, 180) = 0.65$, $p = .421$, partial $\eta^2 = .014$, frequency by length interaction, $F(1, 180) = 0.04$, $p = .846$, partial $\eta^2 = .001$, and the three-way interaction between group, frequency by length, $F(1, 180) = 1.49$, $p = .224$, partial $\eta^2 = .031$.

Table 2

Nonword Repetition Mean Percent Accuracy Scores and Standard Error of the Mean for Younger and Older Children for Low and High Phonotactic Frequency Sequences for 2 and 3 Syllable Nonwords ($N = 47$)

		Low Frequency		High Frequency	
		2 Syll	3 Syll	2 Syll	3 Syll
Younger	<i>M</i>	91.4	91.6	95.9	97.1
	<i>SEM</i>	1.1	1.3	0.9	0.7
Older	<i>M</i>	94.3	95.1	98.0	97.4
	<i>SEM</i>	0.8	0.7	0.5	0.5

Simple effect contrasts revealed that the difference in nonword repetition accuracy between the younger and older children was significant for the low frequency sequences, $t(180) = 2.72$, $p = .007$, partial $\eta^2 = .14$, with the younger children repeating low phonotactic frequency sequences ($M = 92\%$, $SEM = 1\%$) less accurately than the older children ($M = 95\%$, $SEM = 1\%$). The younger children ($M = 97\%$, $SEM = 1\%$) did not differ significantly compared to the older children ($M = 98\%$, $SEM = 0.5\%$) for the high frequency items, $t(180) = 1.58$, $p = .116$, partial $\eta^2 = .05$.

Inspection of the difference in accuracy between the high and low frequency sequences averaged across syllable length for each participant revealed that the

younger children had a numerically a larger phonotactic frequency effect ($M = 5\%$, $SEM = 1\%$) compared to the older children ($M = 3\%$, $SEM = 0.5\%$), although the difference was not significant, $F(1,45) = 3.76$, $p = .059$, partial $\eta^2 = .077$.

Picture-naming

The analysis for the picture-naming task examined age group differences in mean reaction time for each level of priming. Mean reaction time for each participant for each condition was calculated after excluding naming errors (5%), microphone errors, caused by child accidentally knocking microphone or coughing, (5.8%), dysfluency errors (1.1%) and timed out errors (6.9%), totaling 18.8% errors. A total of 2.6% of data were excluded as outliers after errors were excluded.

A two-way mixed design GLMM was used for the analysis with age group as a between groups independent variable with 2 levels (younger and older) and priming with 3 levels (related, unrelated and silence condition) as a repeated measures independent variable. The GLMM revealed a significant main effect for group, $F(1, 135) = 47.18$, $p < .001$, partial $\eta^2 = .50$, with the younger children ($M = 1658$ ms, $SEM = 50$ ms) being significantly slower at naming pictures than the older children ($M = 1244$ ms, $SEM = 38$ ms). The main effect of priming was significant, $F(2, 135) = 257.09$, $p < .001$, $d = 4.782$, partial $\eta^2 = .85$, with the silence condition ($M = 1078$ ms, $SEM = 35$ ms) having the fastest reaction time, followed by the related condition ($M = 1496$ ms, $SEM = 54$ ms) and then the unrelated condition ($M = 1778$ ms, $SEM = 56$ ms). The interaction between group and priming was significant, $F(2, 134) = 7.38$, $p < .001$, partial $\eta^2 = .12$. Refer to Table 3.

Further inspection of the main effect for priming for picture-naming reaction time revealed that the related priming condition was significantly faster than the unrelated condition, $t(134) = 8.66$, $p < .001$, partial $\eta^2 = 2.61$, showing a 282 ms phonological facilitation effect. The unrelated condition was significantly slower than the silence condition, $t(134) = 21.47$, $p < .001$, partial $\eta^2 = .89$, showing a robust 700 ms inhibition effect of the unrelated primes compared to having no prime. The mean reaction time for the related condition was also significantly slower than the silence condition by 418 ms, $t(134) = 15.54$, $p < .001$, partial $\eta^2 = .81$.

Table 3

Picture-naming Mean Reaction Time (PNrt) in milliseconds and Percent Correct (%) with Standard Errors for Younger and Older Children for Unrelated, Related and Silence Priming Conditions (N = 47)

		Priming					
		Unrelated		Related		Silence	
		<i>RT</i>	%	<i>RT</i>	%	<i>RT</i>	%
Younger	<i>M</i>	2022	93	1730	92	1221	94
	<i>SEM</i>	50	2	64	2	41	1
Older	<i>M</i>	1434	95	1262	98	935	97
	<i>SEM</i>	61	1	43	1	30	1

For all group contrasts the younger children were significantly slower than the older children, $t(135) = 5.65$, $p < .001$, partial $\eta^2 = .41$, $t(135) = 6.20$, $p < .001$, partial $\eta^2 = .45$, $t(135) = 6.08$, $p < .001$, partial $\eta^2 = .44$, for silence, related and unrelated prime conditions, respectively. Further inspection of the group by priming interaction revealed the group difference between younger and older children was smallest for the silence condition (286 ms), and largest for the unrelated condition (488 ms). The difference between the younger and older children for the related condition (467 ms) was only marginally smaller than the unrelated condition. Separate analysis of facilitation (related vs. unrelated prime) and inhibition (silence vs. unrelated prime) were undertaken. The phonological facilitation effect was only marginally larger for the younger children (292 ms) compared to the older children (272 ms). A two-way mixed design GLMM, restricting the analysis to the related and unrelated priming conditions, failed to show a group by priming interaction, $F(1, 90) = 0.10$, $p = .752$, partial $\eta^2 = .002$, indicating that the group by priming interaction in the main analysis is not explained by a group difference in phonological facilitation. However, the younger children (801 ms) did show a larger inhibition effect, the difference in accuracy between the silence and unrelated priming condition, when compared to the older children (599 ms). A mixed model GLMM comparing just the silence and unrelated priming condition revealed a significant group by priming interaction, $F(1, 90) = 9.55$, $p = .003$, partial $\eta^2 = .20$, explaining the two way interaction in the main analysis.

For the analysis of picture-naming accuracy, the GLMM revealed a significant main effect for group, $F(1, 135) = 8.39, p = .004$, partial $\eta^2 = .16$, with the older children having a higher mean accuracy ($M = 97\%$, $SEM = 1\%$) than the younger children ($M = 93\%$, $SEM = 2\%$). The main effect of priming was not significant, $F(2, 135) = 0.89, p = .413$, partial $\eta^2 = .04$. The interaction between group and priming was not significant, $F(2, 135) = 1.44, p = .241$, partial $\eta^2 = .06$.

Speech Discrimination

d-prime values were analysed with word pair (i.e., *cap/cat* and *tap/tack*) and gating condition (i.e., whole word, short gate and long gate) as repeated measures independent variables, and group (i.e., younger and older) as between subjects independent variable. The GLMM revealed that groups did not differ significantly, $F(1, 270) = 0.83, p = .362$, partial $\eta^2 = .02$, with a small difference in detection accuracy between the younger ($M = 1.38$, $SEM = 0.14$) and older ($M = 1.42$, $SEM = 0.11$) children. The main effect for gate was significant, $F(2, 270) = 142.07, p < .001$, partial $\eta^2 = .74$, with performance for the ungated, that is, whole word condition, having a significantly higher d-prime ($M = 2.18$, $SEM = 0.04$) when compared to the short-gate condition ($M = 1.55$, $SEM = 0.15$), $t(270) = 7.93, p < .001$, partial $\eta^2 = .57$, and the long-gate condition ($M = 0.54$, $SEM = 0.11$), $t(270) = 16.07, p < .001$, partial $\eta^2 = .85$. A mean level of accuracy (e.g., $M = 0.54$) in terms of percentage indicated children are performing at chance level in the long-gate condition. The short and long gate conditions were also significantly different, $t(270) = 8.52, p < .001$, partial $\eta^2 = .60$. Refer to Table 4.

The main effect for word pair was not significant, $F(1, 270) = 3.00, p = .084$, partial $\eta^2 = .06$, although, there was a significant interaction between group and word pair, $F(1, 270) = 7.95, p = .005$, partial $\eta^2 = .15$. Further inspection of this interaction revealed that the groups were significantly different on the *tap/tack* word pair, $F(1, 278) = 4.93, p = .027$, partial $\eta^2 = .63$, with the younger children ($M = 1.21$, $SEM = 0.16$) having a lower d-prime than the older children ($M = 1.51$, $SEM = 0.12$). Groups were not significantly different for the *cap/cat* word pair, $F(1, 278) = 0.10, p = .319$, partial $\eta^2 = .77$, in this instance the younger children had a higher d-prime ($M = 1.55$, $SEM = 0.13$) than the older children ($M = 1.43$, $SEM = 0.10$).

The interaction between gate and group was not significant, $F(2, 270) = 1.04, p = .356$, partial $\eta^2 = .04$, however the younger children had a lower d-prime than the

older children for the whole word condition ($M = 2.07$ vs. $M = 2.29$) and the short gate condition ($M = 1.4$ vs. $M = 1.61$). The difference was marginal for the long gate condition with the younger children having a nominally higher d -prime than the older children ($M = 0.58$ vs. $M = 0.51$). Planned comparisons between groups for each of the gating conditions were undertaken. The whole word condition was the only condition that revealed a significant difference between the younger and older children, $F(1, 270) = 9.03$, $p = .003$, partial $\eta^2 = .17$. The groups did not differ significantly for the short gate condition or long gate condition, $F(1, 270) = .61$, $p = .437$, partial $\eta^2 = .01$ and $F(1, 270) = 0.13$, $p = .723$, partial $\eta^2 = .003$, respectively. The three-way interaction between group, gate and pair was not significant, $F(2, 270) = 0.97$, $p = .379$, partial $\eta^2 = .03$.

Table 4

Speech Discrimination Ability as Measured by d -prime for Younger and Older Children for Word Pairs, Cat/cap and Tap/tack in the Whole Word, Short Gate and Long Gate Conditions (with Percent Correct in Brackets) ($N = 47$)

		Whole		Short		Long	
		Cat/cap	Tap/tack	Cat/cap	Tap/tack	Cat/cap	Tap/tack
Younger	M	2.24	1.91	1.49	1.47	0.91	.024
		(98%)	(91%)	(82%)	(81%)	(69%)	(55%)
	SEM	0.03	0.14	0.15	0.16	0.20	0.17
Older	M	2.03	2.28	1.52	1.70	0.46	0.55
		(100%)	(100%)	(82%)	(86%)	(59%)	(62%)
	SEM	0.00	0.00	0.14	0.16	0.16	0.19

Note. Whole = whole word, Short = final stop-gap and release burst removed, Long = formant transition, stop gap and release burst removed.

SVRT and CVRT

The SVRT and CVRT were analysed separately. The independent variables used in the GLMM analysis were age group, as a between groups effect, and length (i.e., 1 and 2 syllables words) and word frequency (i.e., high and low frequency words) as repeated measures. Mean reaction time in each condition for each task was calculated after excluding errors and outliers. Errors included targets named incorrectly (2.3% for SVRT, 8.2% for CVRT) or where the trial had timed out at

2500 ms (2.6% for SVRT, 2.6% for CVRT). Outliers were set at 2 standard deviations above or below the mean for each condition for each participant and made up 2.7% of total responses. There were fewer errors for the SVRT compared to the CVRT for both the younger children (1.9% vs. 9.4%, respectively) and older children (2.7% vs. 7.0%, respectively). Errors were not analysed as the focus was on reaction time.

SVRT. GLMM analysis showed a main effect for group, $F(1, 180) = 4.48, p = .036$, partial $\eta^2 = .088$, with the younger children demonstrating a slower reaction time ($M = 575$ ms, $SEM = 20$ ms) when compared to the older children ($M = 520$ ms, $SEM = 16$ ms), with a difference of 55 ms. The main effect of word frequency was significant, $F(1, 180) = 11.92, p = .001$, partial $\eta^2 = .204$, with the low frequency items being named slower ($M = 565$ ms, $SEM = 14$ ms) than the high frequency items ($M = 529$ ms, $SEM = 14$ ms). The main effect for length was not significant, $F(1, 180) = 0.04, p = .844$, partial $\eta^2 = .001$. There were no significant interaction effects; group by word frequency, $F(1, 180) = 2.66, p = .105$, partial $\eta^2 = .054$, group by word length, $F(1, 180) = 0.01, p = .913$, partial $\eta^2 < .001$; length by word frequency, $F(1, 180) = 0.62, p = .434$, partial $\eta^2 = .013$; group by word frequency by length, $F(1, 180) = 0.96, p = .328$, partial $\eta^2 = .020$. Refer to Table 5.

Table 5

Mean Reaction Time in Milliseconds (ms) with Standard Error of the Mean (SEM) for the Simple Verbal Reaction Time for Younger and Older Children (N = 47)

		Low Word Frequency		High Word Frequency	
		1 Syllable	2 Syllable	1 Syllable	2 Syllable
Younger	<i>M</i>	586	582	565	566
	<i>SEM</i>	23	23	30	24
Older	<i>M</i>	539	552	505	482
	<i>SEM</i>	21	28	23	16

CVRT. The GLMM analysis showed that the main effect for group was significant for the CVT, $F(1, 180) = 27.98, p < .001$, partial $\eta^2 = .373$, with the younger children demonstrating a slower reaction time ($M = 855$ ms, $SEM = 25$) than the older children ($M = 675$ ms, $SEM = 23$) with a mean difference of 180 ms. The

main effect for word frequency was not significant, $F(1, 180) = 1.85, p = .175$, partial $\eta^2 = .038$, although the low frequency items were named slower ($M = 773$ ms, $SEM = 19$) than the high frequency items ($M = 757$ ms, $SEM = 18$). The main effect for length was not significant, $F(1, 180) = 0.44, p = .509$, partial $\eta^2 = .009$, however, the mean reaction time for the 1 syllable items was faster ($M = 758$ ms, $SEM = 17$) than the 2 syllable items ($M = 771$ ms, $SEM = 21$). Refer to Table 6.

The interaction effects were all non-significant; group by word frequency, $F(1, 180) = 0.00, p = .981$, partial $\eta^2 < .001$, group by word length, $F(1, 180) = 0.43, p = .513$, partial $\eta^2 = .009$; length by word frequency, $F(1, 180) = 0.39, p = .535$, partial $\eta^2 = .008$. The three-way interaction between group, word frequency and length just failed to reach significance with a medium effect size, $F(1, 180) = 3.44, p = .065$, partial $\eta^2 = .068$. Inspection of the means in Table 6 shows that the 2 syllable CVRTs were longer than the 1 syllable for the low frequency and high frequency items for the older children, in the expected direction, but for the younger children the difference was in the opposite direction for the low frequency items, with the 1 syllable words being named slower than the 2 syllable words.

Table 6

Mean Reaction Time in Milliseconds (ms) with Standard Error (SEM) of the Mean for Choice Verbal Reaction Time for Younger and Older Children (N = 47)

		Low Word Frequency		High Word Frequency	
		1 Syllable	2 Syllable	1 Syllable	2 Syllable
Younger	<i>M</i>	879	847	831	862
	<i>SEM</i>	31	33	29	31
Older	<i>M</i>	662	703	662	671
	<i>SEM</i>	24	38	25	30

Regression Analysis

A variety of regressions analyses were used to examine predictors of the various outcome measures of phonological competence. Table 7 contains simple correlations across all participants (Pearson's product-moment correlations) between the different outcome measures and speech motor control measures. Table 8 shows the correlations between the primary predictor measures (i.e., speech motor control

measures) and outcome measures of phonological competence with measures of age and vocabulary. Age was treated as a continuous variable, rather than a binary variable. Nonword repetition for the low frequency sequences (NWRepLow) was included as an additional measure of phonological competence because this measure was more sensitive than the overall nonword repetition accuracy to developmental differences. Phonological inhibition (PhonInhib) was also included as a variable of interest in the correlational analysis because the two age groups differed significantly on this measure. However, PhonInhib was not used in the regression analysis because the effect could be related to higher level cognitive processes, such as developmental differences in attentional control and therefore it could not be interpreted as a reliable measure of phonological competence.

Inspection of the correlation tables indicates that there were a number of significant correlations between the various outcome measures. Nonword repetition mean accuracy (NWRepPC) correlated with the nonword repetition PhonFreq effect and the PhonFac effect correlated with PhonInhib effect from the picture-naming task. NWRepLow also correlated with PhonInhib. Mean picture-naming reaction time (PNrt) correlated with d-prime. PhonInhib also correlated with PNrt, indicating that as PNrt decreased PhonInhib effects decreased. Picture naming PhonFac did not correlate with any of the measures and d-prime did not correlate with nonword PhonFreq effect. The PNrt, PhonInhib and d-prime were the only measures of phonological encoding ability that correlated with any of the speech motor measures. d-prime was negatively correlated with the SVRT and CVRT, indicating the higher the d-prime value (i.e., the better the speech discrimination accuracy) the lower the reaction time for both the SVRTs and CVRTs. The correlation between PNrt and GFTA-2 raw score was significant, indicating as the GFTA-2 raw score increased so did reaction time. The SVRT and CVRT correlated with one another, however, they did not correlate with GFTA-2 raw score.

In keeping with Munson et al. (2005b) we explored correlations between the outcome measures and measures of expressive and receptive vocabulary and age. For the nonword repetition task the correlations for the low phonotactic frequency sequences and the high phonotactic frequency sequences were analysed separately. As can be seen in Table 9, the mean accuracy scores for the low frequency sequences were positively correlated with age, PPVT raw score and the EVT raw score. The mean accuracy for the high frequency sequences did not significantly

correlate with any of the measures of vocabulary. The PhonFreq correlated negatively with age and PPVT raw score, indicating that as receptive vocabulary increased the PhonFreq effect decreased. The correlation between the phonotactic frequency effect and EVT raw was not significant but was in the expected direction, indicating that as the EVT raw score increased the phonotactic frequency effect decreased.

The mean PNrt was negatively correlated with age, PPVT and EVT raw score, consistent with the view that as age and vocabulary increase, reaction time decreases. The correlation was positive between d-prime and age, PPVT and EVT raw scores, showing that as age and vocabulary increase, speech discrimination ability also increases. Both the SVRT and CVRT were negatively correlated with age and PPVT and the CVRT was also negatively correlated with EVT, indicating that as age and vocabulary increase, reaction time decreases. GFTA raw score was negatively correlated with age and PPVT raw score indicating that the children with more articulation errors tended to be younger and had smaller receptive vocabularies.

The hierarchical regression analysis used GLMM to test for significant predictors of phonological competency while controlling for any relationship with vocabulary. After examining the inter-correlations reported in the previous section (see Table 7 and Table 8) the following phonological measures were included: NWRep for the low frequency sequences, the PhonFreq, mean PNrt, picture naming accuracy and d-prime. The phonological facilitation effect was intended to be a measure of efficiency of phonological encoding during speech production but since it did not correlate with age or any other measures we excluded it from the regression analysis. The inhibition effect did show some association with age, vocabulary and speech motor ability. However, as noted above, it is unclear whether the reaction time difference between the unrelated and silence conditions was related to phonological processes or higher level cognitive processes in managing the distracting effect of the auditory prime (e.g., attentional control).

The purpose of the series of hierarchical regressions was to test for independent contribution the speech motor measures (the primary predictors) and vocabulary in predicting phonological competency measures and whether these predictive relationships varied with age. In particular, we wanted to ascertain if measures of articulatory ability, as measured by the GFTA and supplementary measures of speech motor ability, the SVRT and CVRT, accounted for any

additional variance in the measures of phonological competence and to determine if these relationships differed for the younger and older children. There were separate regression analyses for each primary predictor for each outcome measure of phonological competence.

The analysis proceeded in two stages. Stage 1 tested if the PPVT raw score, and the EVT raw score were significant predictors of each of the outcome measure of phonological competence prior to the regression. For this stage of the analysis the sr^2 value was obtained from the conventional least squares analysis of variance model using SPSS GLM procedure. Stage 2 comprised 3 steps, each of which included the primary predictor; step 1 included potential covariates identified in Stage 1; step 2 included covariates and age; and step 3 included covariates, age and the interaction between the primary predictor and age. The regressions that produced non-significant results for all steps at Stage 2 of the analysis are reported in the Appendix.

Predicting Nonword Repetition Accuracy

Predictors for nonword repetition accuracy for the low phonotactic frequency sequences revealed that the PPVT raw score and EVT raw score were significant predictors, $t(45) = 3.41, p = .001, sr^2 = .537$, $t(45) = 2.43, p = .019, sr^2 = .372$, respectively. Stage 2 of the regression analysis revealed that the SVRT and CVRT were not significant primary predictors of the nonword repetition accuracy for the low phonotactic frequency sequences, consistent with the correlations in Table 7. The results of these hierarchical regressions are reported in Table A4 and A5 in the Appendix.

Primary predictor: GFTA raw score. At step 1 of the regression analysis the GFTA raw score was not a significant predictor of mean nonword repetition accuracy for the low phonotactic frequency sequences, after controlling for the PPVT and EVT raw scores ($p = .443$). The GFTA was a significant predictor ($p < .001$) at step 3 of the regression, however, the marked difference in the regression coefficients (from small and positive to large and negative), refer to Table 9, suggests instability in estimating the coefficient leaving this outcome difficult to

interpret and indicating it may be the result of suppression effects¹. This may be due to the high correlation between the PPVT raw score and age ($r = .78$), refer Table 8. Therefore, the following analysis should be interpreted with caution. Step 3 showed a significant interaction between the GTA raw score and age ($p < .001$). To explore this interaction separate regressions for subgroups based on age were firstly conducted. However, the GFTA raw score was not a significant predictor of nonword repetition accuracy for the low phonotactic frequency sequences for either age group, $p > .05$, leaving the interaction unexplained.

Following this, subgroups were formed based on the GFTA raw score. The first group consisted children with no articulation errors (raw score = 0, $n = 30$) and second group consisted of children with one or more errors (raw score > 0 , $n = 17$). Chronological age was approaching significance as a predictor for children with no articulation errors, $B = 0.121$, $F(1, 26) = 3.72$, $p = .065$, whereas for the children with one or more errors chronological age was not significant, $B = -0.051$, $F(1, 13) = 0.177$, $p = .680$. The positive B coefficient for age for children with no articulation errors shows a trend for older compared to younger children with no articulation errors having higher repetition accuracy. This relationship was not evident for children with one or more articulation errors. However, this may be due to the subgroup with speech errors being restricted to younger children with few older children, given it is the younger children that predominantly have poorer articulatory skills and therefore produce more errors.

The PPVT raw score uniquely predicted accuracy of nonword repetition for low phoneme frequency sequences on each of the three steps (Step 1: $p = .001$; Step 2: $p = .010$; Step 3: $p = .002$), independent of age and other predictors; as the PPVT score increased, accuracy increased. Furthermore, there was a significant increase in the variance (ΔR^2) explained when the interaction was included in the analysis in Step 3.

¹ Suppression effects can occur in regressions when the relationship between the IVs is stronger than the relationship between the IV and the DV, producing a significant result (Cohen & Cohen, 1983).

Table 7

Pearson Correlations Between Phonological Measures and Speech Motor Measures for All Participants (N = 47).

	Phonological Competence							Speech Motor Ability		
	PhonFreq	NWrepPC	NWrep Low	PhonFac	Phon Inhib	PNrt	d-prime	SVrt	CVrt	GFTA Raw
PhonFreq	-									
NWrepPC	-.494**	-								
NWrepLow	-.773**	.933**	-							
PhonFac	.088	.027	-.017	-						
PhonInhib	.372*	-.272	-.352*	.220	-					
PNrt	.125	-.245	-.231	-.094	.294*	-				
d-prime	-.127	.271	.250	.040	-.118	-.358*	-			
SVrt	.166	-.226	-.233	.049	.406**	.479**	-.358*	-		
CVrt	.173	-.281	-.276	-.063	.500**	.586**	-	.647**	-	
							.441**			
GFTA Raw	.166	-.072	-.121	-.174	.179	.324*	-.142	.145	.157	-

Note. PhonFreq = phonotactic frequency effect from picture-naming task; NWRepPC = nonword repetition accuracy percent correct; NWRepLow = nonword repetition accuracy for low phonotactic frequency sequences; PhonFac = phonological facilitation effect from picture-naming task; PhonInhib = phonological inhibition effect from picture-naming task; PNrt = picture-naming reaction time; d-prime = speech discrimination ability for whole word condition; SVrt = simple reaction time; CVrt = choice reaction time; GFTArw = GFTA-2 raw score.

Table 8

Pearson Correlations Between Outcome Measures and Measures of Expressive and Receptive Vocabulary for the Whole Group (N = 47)

	Phonological Competence							Speech Motor Competence		
	NWRep Low	NWRep High	PhonFreq	PhonFac	PhonInhib	PNrt	d-prime	SVrt	CVrt	GFTArw
Age	.462**	.231	-.382**	-.011	-	-	.394**	-.338*	-.633**	-.313*
					.533**	.724**				
PPVT										
Raw	.537**	.279	-.436**	.069	-	-	.364*	-.435**	-.436**	-.376**
					.477**	.611**				
Standard	.127	.151	-.038	.106	-.029	-.122	-.028	-.098	.183	-.232
EVT										
Raw	.372*	.232	-.272	.095	-	-	.320*	-.281	-.444**	-.226
					.418**	.571**				
Standard	-.432**	-.036	.496**	.094	.324*	.261	-.349*	.342*	.447**	.183

Note. NWRepLow = nonword repetition accuracy for low phonotactic frequency sequences; NWRepHigh = nonword repetition accuracy for high phonotactic frequency sequences; PhonFreq = phonotactic frequency effect from picture-naming task; PhonFac = phonological facilitation effect from picture-naming task; PhonInhib = phonological inhibition effect from picture-naming task; PNrt = picture-naming reaction time; d-prime = speech discrimination ability for whole word condition; SVrt = simple reaction time; CVrt = choice reaction time; GFTArw = GFTA raw score.

Table 9

Unstandardized (B) and Standardized (β) Regression Coefficients, and Squared Semi-partial Correlations (sr^2) for a Hierarchical Multiple Regression Model Predicting the Mean Nonword Repetition Accuracy for the Low Frequency Sequences from GFTA Raw Scores, Chronological Age, and the Chronological Age x GFTA Interaction (N = 47)

Predictors (IVs)	B	95% CI ¹	B	sr^2	p-value ¹
DV: NWrepLow					
Step 1					
EVT raw	-0.054	-0.167, 0.058	-.144	.008	.335
PPVT raw	0.142	0.061, 0.222	.689	.164	.001**
GFTA raw	0.293	-0.471, 1.057	.106	.009	.443
$R^2 = .304, p = .001^{**}$					
Step 2					
EVT raw	-0.104	-0.244, 0.037	-.273	.021	.145
PPVT raw	0.124	0.031, 0.216	.601	.108	.010*
GFTA raw	0.338	-0.432, 1.107	.122	.012	.381
ChronAge	0.067	-0.068, 0.201	.251	.018	.323
ChronAge: $F(1, 42) = 1.00, p = .323$					
$\Delta R^2 = .018, p = .294$					
$R^2 = .322, p = .002^{**}$					
Step 3					
EVT raw	-0.167	-0.301, -0.032	-.439	.050	.016*
PPVT raw	0.138	0.055, 0.221	.670	.133	.002**
GFTA raw	7.943	4.665, 11.222	-.509	.048	.000**
ChronAge	0.033	-0.091, 0.157	.126	.004	.590
ChronAge: $F(1, 41) = 0.30, p = .590^2$					
ChronAgexGFTA ³	-0.176	-0.253, -0.099	-.693	.095	.000
ChronAge x GFTA: $F(1, 41) = 21.27, p < .001^4$					
$\Delta R^2 = .095, p = .013$					
$R^2 = .418, p = .000^{**}$					

1: These are the GLMM adjusted values.

2: This is the overall F -value for the group effect.

3: The interaction term is computed using centered GFTA raw scores.

4: This is the overall F -value for the Chronological Age x GFTA interaction effect.

Predicting the Phonotactic Frequency Effect

Stage 1 GLMM hierarchical regression model showed that the PPVT raw score was a significant predictor the phonotactic frequency effect, $t(45) = 2.34, p = .022, sr^2 = .436$. The EVT raw score was not a significant predictor $t(45) = 1.71, p = .095, sr^2 = .272$. This confirms the correlations in Table 9 in the context of GLMM.

The regression analysis undertaken for the SVRT and the CVRT as potential predictors of the phonotactic frequency effect were not significant. The results of these regressions are reported in Table A6 and A7 in Appendix A.

Primary predictor: GFTA raw score. At step 1 of the GLMM hierarchical regression the GFTA raw score was not a significant predictor of the phonotactic frequency effect after controlling for the PPVT raw score ($p = .987$). The PPVT raw score was a significant covariate for step 1 of the analysis ($p = .010$), as the PPVT raw score increased, the frequency effect decreased. The PPVT was not a significant unique predictor at step 2 or step 3 of the regression. Refer to Table 10

The GFTA raw score interacted with age ($p = .031$) at step 3 of the regression indicating that this relationship varied as a function of age. To explore this relationship separate regressions for subgroups based on age were examined, however, the GFTA raw score was not a significant predictor for the phonotactic frequency effect for younger children ($n = 23$), $B = -0.067$, $F(1, 20) = 0.04$, $p = .840$, or the older children ($n = 24$), $B = 1.461$, $F(1, 21) = 1.07$, $p = .313$, leaving the interaction unexplained.

Regressions were then conducted separately for each subgroup based on articulation errors (GFTA raw score = 0, $n = 30$ and GFTA raw score > 0, $n = 17$), showed age was not a significant predictor for either group, $B = -0.044$, $F(1, 27) = 1.16$, $p = .292$, for children with no articulation errors, $B = 0.048$, $F(1, 14) = 0.242$, $p = .630$, for children with one or more articulation errors. Although these results do not clearly explain why the interaction term between GFTA raw score and age was significant in predicting phonotactic frequency, the Pearson correlations between age and phonotactic frequency was found to be significant for the children with no articulation errors, $r = -.515$, but not for the children with one or more articulation errors, $r = -.015$. Similar to the interaction between age and nonword repetition accuracy for low frequency sequences, this suggests that the interaction is more related to a restriction in age range for the children with articulation errors. However, it should be noted that since the correlation between PPVT and age was high ($r = .78$), it is possible that the interaction between age and the phonotactic frequency effect may be the result of a suppression effect within the regression analysis. Consequently, as in the previous regression analysis, this interaction cannot be interpreted.

Table 10

Unstandardized (B) and Standardized (β) Regression Coefficients, and Squared Semi-partial Correlations (sr^2) for a Hierarchical Multiple Regression Model Predicting the Phonotactic Frequency Effect from GFTA Raw Scores, Chronological Age (ChronAge), and the Chronological Age x GFTA Interaction (N = 47)

Predictors (IVs)	B	95% CI ¹	B	sr^2	p-value ¹
DV: PhonFreq					
Step 1					
PPVT raw	-0.074	-0.247, -0.001	-.435	.162	.010
GFTA raw	-0.005	-0.599, 0.609	-.002	.000	.987
$R^2 = .190, p = .010^{**}$					
Step 2					
PPVT raw	-0.060	-0.163, 0.043	-.353	.046	.247
GFTA raw	-0.001	-0.594, 0.593	.000	.000	.999
ChronAge	-0.023	-0.116, 0.069	-.106	.004	.613
ChronAge: $F(1, 43) = 0.26, p = .613$					
$\Delta R^2 = .004, p = .630$					
$R^2 = .194, p = .025^*$					
Step 3					
PPVT raw	-0.060	-0.162, 0.043	-.352	.046	.246
GFTA raw	-3.664	-6.913, -0.414	.379	.027	.028
ChronAge	0.004	-0.090, 0.097	.017	.000	.938
ChronAge: $F(1, 42) = 0.01, p = .938^2$					
ChronAge x GFTA ³	0.085	0.008, 0.162	.406	.036	.031
ChronAge x GFTA: $F(1, 42) = 4.97, p = .031^4$					
$\Delta R^2 = .036, p = .171$					
$R^2 = .230, p = .024^*$					

1: These are the GLMM adjusted values.

2: This is the overall F -value for the group effect.

3: The interaction term is computed using centered GFTA raw scores.

4: This is the overall F -value for the Chronological Age x GFTA interaction effect.

Predicting Picture-naming Reaction Time

The PPVT and EVT raw scores significantly predicted picture-naming reaction time, $t(45) = 6.13, p < .001, sr^2 = .611$ and $t(45) = 5.24, p < .001, sr^2 = .571$, respectively.

Primary predictor: GFTA raw score. Step 1 of the hierarchical regression showed that the GFTA raw score was not a significant predictor of picture-naming reaction time after controlling for the PPVT and EVT raw scores ($p = .146$). Refer to Table 11.

The PPVT raw score predicted PNrt on step 1 of the regression analysis ($p = .038$) accounting for 4.3% of unique variance in PNrt; as the PPVT raw score increased, the PNrt decreased. When age was included in the analysis the PPVT no

longer predicted unique variance of PNrt. GFTA raw score interacted with age ($p = .022$) indicating that this relationship varied as a function of age.

The interaction between GFTA and chronological age was examined using separate regressions for subgroups based on age, however, the GFTA raw score did not predict PNrt for either age group, leaving the interactions unexplained.

Regressions performed separately for subgroups split on articulatory ability instead showed chronological age was a significant predictor of PNrt for both subgroups, children with no articulation errors, $B = -.6.72$, $p = .049$, $F(1, 26) = 4.26$, $p = .049$, and children with one or more articulation errors, $B = -13.89$, $p = .013$, $F(1, 13) = 8.270$, $p = .013$. The larger negative B coefficient for the children with one or more articulation errors indicated a steeper slope and consequently a stronger relationship between articulation and PNrt for the children with poorer articulatory ability.

However, this interpretation is tentative given the possible suppression effects and the highly variable regression coefficients.

Simple verbal reaction time. Mean SVRT was a significant unique predictor of picture-naming reaction time at step 1 of the analysis after controlling PPVT and EVT raw scores ($p = .013$) and at step 2 of the analysis when age was included in the analysis ($p = .008$), accounting for 6.5% of the variance in PNrt and 5.4% respectively. The EVT raw score uniquely predicted PNrt at step 1 of the regression analysis ($p = .048$); as the EVT score increased, the picture-naming reaction time decreased. This suggests that variance in PNrt explained by expressive vocabulary is shared with age, demonstrated by the value for sr^2 decreasing from 2.1% in step 1 of the hierarchical regression to 0.0% in step 2 and 3. Refer to Table 12.

Chronological age was a significant predictor at step 2 ($p = .002$) and step 3 ($p = .003$) of the analysis, accounting for 12% of additional unique variance and 11% of additional unique variance, respectively. The interaction between SVrt and age was not significant ($p = .950$), indicating that the relationship between SVrt and PNrt did not vary as a function of age.

Table 11

Unstandardized (B) and Standardized (β) Regression Coefficients, and Squared Semi-partial Correlations (sr^2) for a Hierarchical Multiple Regression Model Predicting Picture-naming Reaction Time from GFTA Raw Scores, Chronological Age, and the Chronological Age x GFTA Interaction (N = 47)

Predictors (IVs)	B	95% CI ¹	β	sr^2	p-value ¹
DV: PNrt					
Step 1					
EVT raw	-5.306	-11.530, 0.917	-.264	.026	.093
PPVT raw	-3.869	-7.506, -0.232	-.335	.043	.038*
GFTA raw	19.256	-6.985, 45.497	.131	.015	.146
$R^2 = .410, p = .000^{**}$					
Step 2					
EVT raw	1.600	-4.667, 7.867	.080	.002	.609
PPVT raw	-1.300	-4.728, 2.127	-.119	.004	.448
GFTA raw	13.00	-10.387, 36.387	.089	.007	.268
ChronAge	-9.383	-14.906, -3.860	-.667	.129	.001**
ChronAge: $F(1, 42) = 11.76, p = .001$					
$\Delta R^2 = .129, p = .001^{**}$					
$R^2 = .539, p = .000^{***}$					
Step 3					
EVT raw	-0.460	-7.215, 6.295	-.023	.000	.891
PPVT raw	-0.836	-4.035, 2.363	-.077	.002	.600
GFTA raw	260.852	56.051, 465.652	-.300	.017	.014*
ChronAge	-10.469	-15.770, -5.169	-.744	.153	.000**
ChronAge: $F(1, 41) = 15.91, p < .001^2$					
ChronAge x GFTA ³	-5.730	-10.589, -0.870	-.426	.036	.022*
ChronAge x GFTA: $F(1, 41) = 5.67, p = .022^4$					
$\Delta R^2 = .036, p = .069$					
$R^2 = .575, p < .001^{**}$					

1: These are the GLMM adjusted values.

2: This is the overall F -value for the group effect.

3: The interaction term is computed using centered GFTA raw scores.

4: This is the overall F -value for the Chronological Age x GFTA interaction effect.

Table 12

Unstandardized (B) and Standardized (β) Regression Coefficients, and Squared Semi-partial Correlations (sr^2) for a Hierarchical Multiple Regression Model Predicting Picture-naming Reaction Time from Simple Verbal Reaction Time, Chronological Age (ChronAge), and the Chronological Age x Simple Verbal Reaction Time Interaction (N = 47)

Predictors (IVs)	B	95% CI ¹	β	sr^2	p-value ¹
DV: PNrt					
Step 1					
EVT raw	-5.729	-11.403, -0.055	-.285	.031	.048*
PPVT raw	-2.879	-6.647, 0.890	-.264	.023	.131
SVRT	0.687	0.150, 1.223	-.284	.065	.013*
$R^2 = .460, p = .000^{**}$					
Step 2					
EVT raw	1.006	-5.163, 7.175	.050	.000	.774
PPVT raw	-0.253	-4.053, 3.547	-.023	.000	.894
SVrt	0.630	0.170, 1.089	.261	.054	.008**
ChronAge	-9.253	-14.754, -3.751	-.658	.126	.002**
ChronAge: $F(1, 42) = 11.52, p = .002$					
$\Delta R^2 = .126, p = .001^{**}$					
$R^2 = .586, p = .000^{**}$					
Step 3					
EVT raw	1.084	-5.630, 7.798	.054	.000	.746
PPVT raw	-0.251	-4.068, 3.565	-.023	.000	.895
SVrt	0.693	-1.434, 2.820	.261	.054	.514
ChronAge	-9.300	-15.319, -3.280	-.661	.107	.003**
ChronAge: $F(1, 41) = 9.74, p = .003^2$					
ChronAge x SVrt ³	-0.073	-2.386, 2.241	-.005	.000	.950
ChronAge x SVrt: $F(1, 41) = 0.00, p = .950^4$					
$\Delta R^2 = .000, p = .967$					
$R^2 = .586, p = .000^{**}$					

1: These are the GLMM adjusted values.

2: This is the overall F -value for the group effect.

3: The interaction term is computed using centered SVrt.

4: This is the overall F -value for the Chronological Age x SVrt interaction effect.

Choice verbal reaction time. Step 1 of the hierarchical regression showed that the CVRT was a significant predictor of picture-naming reaction time after controlling PPVT and EVT raw scores ($p = .004$). The PPVT raw score predicted picture-naming reaction time for step 1 of the regression analysis ($p = .033$); as the PPVT score increased, the picture-naming reaction time decreased. Refer to Table 13.

At step 2 of the regression when age was included in the analysis the CVRT sr^2 value dropped from 11% to 3% ($p = .073$) indicating that some of variance accounted for by the CVRT was shared with chronological age ($p = .019$). However, age also

predicted significant unique variance at step 2 and step 3 of the hierarchical regression accounting for 5.4% of the variance in PNrt at step 2 ($p = .019$) and 6% at step 3 ($p = .011$). The CVrt did not interact with chronological age ($p = .134$) indicating that this relationship did not vary as a function of age.

Table 13

Unstandardized (B) and Standardized (β) Regression Coefficients, and Squared Semi-partial Correlations (sr^2) for a Hierarchical Multiple Regression Model Predicting Picture-naming Reaction Time (PNrt) from Choice Verbal Reaction Time, Chronological Age (ChronAge), and the Chronological Age x Choice Verbal Reaction Time Interaction (N = 47)

Predictors (IVs)	B	95% CI ¹	β	sr^2	p-value ¹
DV: PNrt					
Step 1					
EVT raw	-2.834	-8.144, 2.477	-.141	.007	.288
PPVT raw	-3.663	-7.017, -0.309	-.336	.042	.033*
CVrt	0.575	0.195, 0.955	.378	.112	.004**
$R^2 = .507, p = .000^{**}$					
Step 2					
EVT raw	1.402	-4.536, 7.340	.070	.001	.636
PPVT raw	-1.946	-5.287, 1.394	-.179	.010	.246
CVrt	0.341	-0.033, 0.715	.224	.029	.073
ChronAge	-7.018	-12.800, -1.237	-.499	.054	.019*
ChronAge: $F(1, 42) = 6.00, p = .019^*$					
$\Delta R^2 = .054, p = .028^*$					
$R^2 = .562, p = .000^{***}$					
Step 3					
EVT raw	2.032	-3.890, 7.953	.101	.003	.492
PPVT raw	-1.873	-5.292, 1.546	-.172	.009	.275
CVrt	1.429	-0.067, 2.925	.217	.028	.061
ChronAge	-7.371	-12.984, -1.758	-.524	.060	.011*
ChronAge: $F(1, 41) = 7.03, p = .011^2$					
ChronAge x CVrt ³	-1.989	-4.616, 0.637	-.116	.013	.134
ChronAge x CVrt: $F(1, 41) = 2.34, p = .134^4$					
$\Delta R^2 = .013, p = .273$					
$R^2 = .574, p = .000^{**}$					

1: These are the GLMM adjusted values.

2: This is the overall F -value for the group effect.

3: The interaction term is computed using centered CVrt.

4: This is the overall F -value for the Chronological Age x CVrt interaction effect.

Predicting Speech Discrimination Ability

The PPVT and EVT raw scores were significant predictors of d-prime, $t(45) = 6.64, p = .008, sr^2 = .364$, and $t(45) = 2.50, p = .016, sr^2 = .320$, respectively. The GFTA raw score was not a significant predictor of d-prime after controlling for the PPVT and EVT raw scores ($p = .909$). Refer to Table A8 in Appendix A.

The PPVT and EVT raw scores when included with age were no longer significant predictors of speech discrimination ability suggesting the variance explained by vocabulary measures is shared with age. GFTA raw score did not interact with age ($p = .964$) indicating that the relationship between GFTA and d-prime did not vary as a function of age. The results of the hierarchical regression with GFTA raw score as primary predictor are reported Table A10 in the Appendix.

Simple verbal reaction time. SVRT was not a significant predictor of d-prime at step 1 or step 2 of the analysis after controlling PPVT and EVT raw scores, although the p value was close to reaching significance ($p = .077$) at step 1. It was significant at step 3 of the analysis ($p = .037$), accounting for 5% ($sr^2 = .047$) of the variance in d-prime, indicating as reaction time increased, speech discrimination sensitivity decreased.

Chronological age was also significant at step 3 of the analysis ($p = .042$) uniquely accounting for 2% ($sr^2 = .024$) of the variance in d-prime. The interaction between SVRT and age was also significant ($p = .038$), indicating that this relationship varied as a function of age. Furthermore, there was a significant increase in the variance explained in Step 3 of the hierarchical regression ($\Delta R^2 = 8\%$), thus demonstrating a robust finding. Refer to Table 14.

To examine the relationship between d-prime and age we looked at subgroups based on age, splitting the children into younger ($n = 23$) and older groups ($n = 24$). The younger group revealed a significant relationship with SVRT, $B = -0.002$, $F(1, 19) = 5.31$ $p = .033$, indicating for the younger children that as reaction time increased sensitivity decreased. For the older group the SVRT did not uniquely predict d-prime, $B < 0.000$, $F(1, 20) = 1.33$, $p = .263$.

Choice verbal reaction time. The CVRT was a significant predictor of d-prime after controlling PPVT and EVT raw scores ($p = .033$), indicating that CVRT uniquely predicts speech discrimination ability, accounting for 9.5% of unique variance, over and above that predicted by vocabulary. Refer to Table 15.

When age was included at step 2 of the regression the CVRT was no longer significant ($p = .061$), accounting for 7% of the variance, indicating that as reaction time increases speech discrimination sensitivity decreased. Chronological age was not a significant predictor of d-prime at step 2 ($p = .850$) or step 3 of the analysis (p

= .639). The CVRT did not interact with chronological age ($p = .124$), indicating that the prediction of d-prime did not vary as a function of age.

Table 14

Unstandardized (B) and standardized (β) Regression Coefficients, and Squared Part Correlations (sr^2) for a Hierarchical Multiple Regression Model Predicting d-prime from Simple Verbal Reaction Time, Chronological Age, and the Chronological Age x SVRT Interaction ($N = 47$).

Predictors (IVs)	B	95% CI ¹	β	sr^2	p-value ¹
DV: d-prime					
Step 1					
EVT raw	0.003	-0.004, 0.010	.130	.007	.349
PPVT raw	0.002	-0.002, 0.006	.151	.008	.292
SVRT	-0.001	-0.002, 0.000	-.256	.052	.077
$R^2 = .188, p = .028^*$					
Step 2					
EVT raw	-0.000	-0.008, 0.007	-.017	.000	.908
PPVT raw	0.001	-0.004, 0.005	.045	.001	.781
SVRT	-0.001	-0.002, 0.000	-.246	.048	.091
ChronAge	0.005	-0.002, 0.011	.289	.024	.135
ChronAge: $F(1, 42) = 2.33, p = .135$					
$\Delta R^2 = .024, p = .260$					
$R^2 = .213, p = .036^*$					
Step 3					
EVT raw	-0.007	-0.016, 0.002	-.284	.018	.126
PPVT raw	0.000	-0.003, 0.004	.033	.000	.818
SVRT	-0.006	-0.011, -0.000	-.243	.047	.037*
ChronAge	0.009	0.000, 0.017	.518	.066	.042*
ChronAge: $F(1, 41) = 4.39, p = .042^2$					
ChronAge x SVRT ³	0.006	0.000, 0.012	.325	.080	.038*
ChronAge x SVRT: $F(1, 41) = 4.61, p = .038^4$					
$\Delta R^2 = .080, p = .038^*$					
$R^2 = .293, p = .012^*$					

1: These are the GLMM adjusted values.

2: This is the overall F -value for the group effect.

3: The interaction term is computed using centered SVRTs.

4: This is the overall F -value for the Chronological Age x SVRT interaction effect.

Table 15

Unstandardized (B) and Standardized (β) Regression Coefficients, and Squared Semi-partial Correlations (sr^2) for a Hierarchical Multiple Regression Model Predicting d-prime from Choice Verbal Reaction Time, Chronological Age (ChronAge), and the ChronAge x Choice Verbal Reaction Time Interaction (N = 47)

Predictors (IVs)	B	95% CI ¹	β	sr^2	p-value ¹
DV: d-prime					
Step 1					
EVT raw	-0.000	-0.006, 0.006	-.002	.000	.989
PPVT raw	0.003	-0.001, 0.007	.214	.017	.167
CVRT	-0.001	-0.001, -0.00	-.348	.095	.033*
$R^2 = .231, p = .010^*$					
Step 2					
EVT raw	-0.000	-0.007, 0.007	-.017	.000	.910
PPVT raw	0.003	-0.002, 0.007	.203	.013	.214
CVRT	-0.001	-0.001, -0.00	-.337	.067	.061
ChronAge	0.001	-0.006, 0.007	.036	.000	.850
ChronAge: $F(1, 42) = 0.04, p = .850$					
$\Delta R^2 = .000, p = .903$					
$R^2 = .231, p = .023^*$					
Step 3					
EVT raw	-0.002	-0.009, 0.005	-.074	.002	.621
PPVT raw	0.002	-0.002, 0.007	.190	.012	.254
CVRT	-0.003	-0.007, 0.001	-.325	.062	.103
ChronAge	0.001	-0.004, 0.007	.082	.001	.639
ChronAge: $F(1, 41) = 0.22, p = .639^2$					
ChronAge x CVRT ³	0.004	-0.001, 0.010	.212	.043	.124
ChronAge x CVRT : $F(1, 41) = 2.46, p = .124^4$					
$\Delta R^2 = .043, p = .125$					
$R^2 = .275, p = .018^*$					

1: These are the GLMM adjusted values.

2: This is the overall F -value for the group effect.

3: The interaction term is computed using centered CVRTs.

4: This is the overall F -value for the Chronological Age x CVRT interaction effect.

Discussion

The purpose of study 1 was to pilot a set of tasks designed to target different components of phonological competency and speech motor ability in children with TD. Our goal was to investigate whether the younger and older children differed in phonological competency and speech motor ability and if the measures obtained from these tasks could be useful in examining the predictive relationship between phonological competence and speech motor ability in younger and older children. An additional goal was to determine if vocabulary predicted phonological competency over and above that predicted by age, replicating previous studies, and if

speech motor measures predicted phonological competence independent of the contribution made by vocabulary. Our key hypothesis was that age related developmental changes in phonological competency and speech motor ability would be observed and speech motor measures would predict phonological competency to a greater extent in younger children. The findings are discussed in relation to the individual tasks and in relation to the regression analysis.

Developmental Differences in Phonological Competency

In relation to measures of phonological competency, overall our findings indicate that the younger children in our study are less accurate at nonword discrimination, have poorer speech discrimination skills and are slower at picture-naming, than older children. These findings are consistent with previous studies and consistent with the understanding that younger children have less robust phonological representations than older children (Brooks & Mac Whinney, 2000; Edwards et al., 1999; Edwards et al., 2002; Munson & Babel, 2005). The younger children in the current study were less accurate than the older children on nonword repetition accuracy and high frequency nonwords were repeated more accurately than the nonwords that contained low phonotactic frequency sequences, consistent with the view that the phonological representations of younger children are less abstracted than representations of older children (Edwards et al., 2004; Munson, Edwards, et al., 2005b). Phonotactic frequency has been shown to influence lexical acquisition in preschool and school aged children with more common sound sequences being acquired more rapidly than less common sequences (Storkel & Rogers, 2000). Although the interaction between age and the phonotactic frequency effect just missed out on reaching significance ($p = .051$), planned comparisons showed the age groups differed on nonword repetition accuracy for the low phonotactic frequency sequences with the younger children being significantly less accurate for low frequency sequences, demonstrating a disadvantage for these items and indicating a developmental difference between the groups.

The reaction time data from the picture-naming task revealed clear developmental results with the younger children having significantly slower reaction times than the older children. This finding is consistent with the view that younger children have less efficient picture-naming skills, indicative of a less developed word retrieval system and potentially less efficient phonological encoding skills. In

addition, the younger children produced significantly more naming errors than the older children. The younger children revealed a larger phonological facilitation effect than the older children, although this difference was not statistically significant. We predicted that the degree of facilitation would be sensitive to phonological encoding efficiency, thereby reflecting the quality and stability of the underlying phonological representations. This was based on the understanding that as development proceeds phonological representations progress from holistic to more fine-grained representations, which has been shown to be a result of vocabulary expansion (Metsala, 1997; Walley, 1993). In keeping with this view, we assumed that the younger children would be less efficient in processes such as phonological retrieval and encoding and therefore the younger children would benefit to a greater degree than the older children when hearing related primes, demonstrating a larger phonological facilitation effect. One plausible explanation why the groups did not differ in phonological facilitation effects could be that they were too close in age to show a significant difference in phonological facilitation effects with no marked developmental changes in output phonological encoding of speech across the age range of the children in the current study. The fact that the children differed significantly for reaction time data would suggest that developmental differences were detectable. However, reaction time data incorporates other levels of processing such as object recognition, lexical-semantic activation and/or speech motor skills, in addition to encoding of the phonological form after the target word is selected, and therefore differences in efficiency at these other stages in picture-naming could explain the age group difference in reaction time. Alternatively, the task design could have comprised our finding no significant difference in phonological facilitation between the younger and older children. In the current study onset primes were used as auditory distractors, whereas Brookes and MacWhinney (2000) showed a bias for the offset prime (i.e., rime prime) in younger children, demonstrating greater phonological facilitation for the rime prime condition compared to the onset related prime condition. Furthermore, our task included only one SOA (at 116ms post picture presentation) to reduce the length of the task and make it more manageable for the younger participants. According to Brooks and MacWhinney (2000) this SOA may be too early for the younger participants to show enhanced facilitation effects. Brooks and MacWhinney (2000) found that the

maximum priming effect was delayed in their five-year-old group in comparison to the older children and adults and priming effects were at their peak at 150ms post picture presentation. Consequently, it is plausible that the SOA implemented was not ideal for the younger children and therefore may have contributed to a smaller priming effect that did not differ in magnitude to the older children. Future developmental studies should compare priming effects across different SOAs. Furthermore, phonological facilitation did not correlate with any other measures of phonological competence or age or vocabulary, suggesting that the current version of the picture-naming task may not be a reliable measure of competence at the level of output phonological representations.

The younger children revealed a significantly larger inhibition effect compared to the older children, consistent with the view that younger children are more susceptible to interference (Dempster, 1992). Although the inhibition effects observed could be associated with age related differences in phonological development and interpreted as a sensitive measure of the efficiency of phonological encoding they could also reflect others stages of the picture-naming process, as previously mentioned, or something more generic like resource allocation or attentional control, which vary as a function of age (Rueda et al., 2004). Consequently the inhibition effects need to be interpreted with caution, as they cannot be taken as a pure measure of phonological encoding efficiency or the abstractness of output phonological representations. The picture-naming task replicated a number of findings from earlier studies demonstrating its merit. Robust priming effects were demonstrated with the related priming condition having a significantly faster reaction time than the unrelated priming condition, consistent with previous studies (Brooks & Mac Whinney, 2000; Jerger et al., 2002). Furthermore, the silence condition showed a significantly faster reaction time than the unrelated condition, demonstrating a robust inhibition effect of the distractor, also consistent with previous research (Brooks & Mac Whinney, 2000; Jerger et al., 2002; Schriefers et al., 1990).

The speech discrimination task revealed findings consistent with previous research (Edwards et al., 1999; Edwards et al., 2002; Munson, Edwards, et al., 2005b). Both the younger and older children performed at a very low level for the gated conditions with no significant difference revealed on speech discrimination

ability between groups when all gates were included in the analysis, also found in Edwards et al.'s study (1999). However, consistent with Edwards et al. (1999), the younger children demonstrated significantly poorer speech discrimination skills than the older children for the whole word condition, demonstrating that this measure is a sensitive measure of developmental changes in speech discrimination ability.

Developmental Differences in Speech Motor Ability

In relation to the speech motor measures the simple verbal reaction time task revealed clear developmental differences in the processes required for picture-naming, with the younger children being slower overall than the older children. For the SVRT there was a significant word frequency effect with the low frequency words being named slower than the high frequency words. Although this is consistent with our understanding that high frequency occurring words are retrieved faster than low frequency words (Newman & German, 2002; Vitevitch & Sommers, 2003), it was an unexpected finding for the SVRT as word frequency effects are assumed to be located at the level of retrieving lexical phonology (Jescheniak & Levelt, 1994). Given the word is planned in advance since the response is known, word frequency should not impact on response latency. Consequently, the word frequency effect is more likely to be associated with extraneous variables, possibly speech production parameters. The high frequency target words began with the stop consonants /p/, /d/ and /k/ and low frequency items began with /b/, /k/, /n/ and /f/. A stop consonant is produced when air-flow is temporarily obstructed by the vocal tract, therefore, when airflow resumes this results in a surge of air, which triggers the microphone and in doing so captures the reaction time instantaneously. For fricatives and nasals, such as /f/ and /n/ there is no sudden surge of air, and therefore the microphone may not be triggered instantaneously to capture the reaction time as effectively, thus compromising reaction time data. It is therefore plausible that the word frequency effect for the SVRT is the result of factors associated with the small number of task stimuli in each condition. Regardless, the word frequency effect potentially undermines the validity of the SVRT as a measure of speech motor ability, since the word frequency effects are assumed to occur during lexical processing. Thus suggesting that some of the reaction time associated with the SVRT includes higher level processes associated with lexical retrieval. The SVRT did not reveal a word length effect, consistent with the view that the phonetic code is

planned in advance during SVRT and as a result the length of the target word should not influence response latency, consistent with previous findings (Klapp, 1995) and indicating the task has merit as a measure of the efficiency of on-line speech execution processes.

For the CVRT task the younger children were slower than the older children in pictures naming, indicating that this task is sensitive to developmental differences in planning and execution of verbal responses. In addition, the difference in reaction time between the younger and older children was greater for the CVRT compared to the SVRT, indicating that the CVRT was more difficult for the younger children compared to the SVRT, although this was not analysed statistically. The word frequency effect for the CVRT was not significant, consistent with our understanding that the CVRT requires minimal processes at higher levels of processing, such as lexical retrieval, consequently word frequency should not influence reaction time. The CVRT did not demonstrate a word length effect, as hypothesized, therefore compromising its validity as a measure of speech motor planning. Word length effects have been shown to be confounded by a number of variables, such as age of acquisition, word frequency effects (Markus et al., 2010) and given age of acquisition was not controlled in the current study this may have compromised finding a word length effect. It is also plausible that the absence of a length effect could be due to the speaker initiating their response as soon as the first syllable was encoded and ready for execution, thereby diminishing any potential length effect (Markus et al., 2010). This is consistent with Meyers, Roelofs and Levelt (2003) who found that speakers used different strategies to meet response deadlines, for longer words speakers generated the motor program for the first syllable and initiated their response prior to the second syllable being programmed and ready for execution. Given the stimuli for this task consisted of one and two syllable word pairs shared the same onset syllable (e.g. cart/cartwheel, pig/piglet) it is plausible that the retrieval and execution of the onset syllable could have diminished the effect for length, as the response could be initiated for the first syllable regardless of whether the second syllable was programmed and ready for execution. Although the word length did not achieve significance, there did appear to be a pattern for the older children to have longer response latency for the two syllable words, not evident

for the younger children. This was supported by a near significant three-way interaction between group, frequency and length.

Relationship Between Speech Motor Measures and Phonological Competence

For the regression analysis we looked at the repetition accuracy for the low phonotactic frequency sequences, rather than the combined accuracy for the high and low frequency sequences, since the low frequency sequences showed a stronger correlation with the phonotactic frequency effect and measures of vocabulary (see Table 8). We did not include measures of phonological facilitation in our regression analysis, since the difference in the phonological facilitation was not significant for the younger children compared to the older children.

Measures of vocabulary, both expressive (EVT) and receptive (PPVT), and age were tested initially, to determine if they predicted outcome measures prior to the primary predictors being tested. Primary predictors included the GFTA (i.e., measure of articulatory ability) and the speech motor measures (i.e., the SVRT and the CVRT). Overall vocabulary proved to be a significant predictor of phonological competence, consistent with the view that phonological development is driven to a large extent by vocabulary expansion beyond the effects of age or articulatory ability (Edwards et al., 2004; Munson, Edwards, et al., 2005b; Munson, Kurtz, et al., 2005). In the current study the PPVT predicted nonword repetition accuracy, consistent with Munson et al. (2005), whereas Edwards et al. (2004) found that the EVT accounted for a significant proportion of unique variance beyond that accounted for by the GFTA and age. The regression analysis confirmed that both receptive and expressive vocabulary were significant predictors of the phonotactic frequency effect, accounting for a significant portion of the variance, consistent with the view that the decline in the frequency effect is related to vocabulary growth (Edwards et al., 2004; Munson, Edwards, et al., 2005b). The EVT and PPVT were also significant predictors of picture-naming reaction time. Speech discrimination ability was also predicted by both measures of vocabulary in the regression analysis, consistent with our hypothesis and the premise that accurate phonological input representations emerge as a child accumulates lexical items in their mental lexicon (Edwards et al., 2002; Munson, Edwards, & Beckman, 2005a; Munson, Edwards, et al., 2005b).

Findings from the regression analysis with the GFTA as the main predictor were equivocal. The GFTA did not predict nonword repetition accuracy or the phonotactic frequency effect, as hypothesised. It did however interact with age for both of these outcome measures. Further investigation of this interaction revealed that the significant interaction could not be explained by age group but was explained by articulatory ability. It might also be a result of the distribution of the GFTA raw scores where only a small number of the older children showed articulation errors, consequently the age effect for the children with articulation errors was weakened compared to the group without articulation errors, which included a more balanced group of younger and older children. These interactions are difficult to interpret as they could be the result of suppression effects given GFTA was not a significant predictor at step 1 for either of these regressions. Regardless, it was a noteworthy finding that these interactions were better explained in the context of articulatory ability as opposed to age.

The GFTA also failed to predict picture-naming reaction time after accounting for measures of vocabulary, however, it did interact with age at step 3 of the regression, suggesting that the relationship between PNrt and GFTA varied as function of age. We investigated this relationship as a function of articulatory ability and found that the slope of the relationship between age and PNrt was steeper for the children with articulation errors compared to children with no articulation errors, indicating that articulatory ability has some influence over PNrt, at least for children with poorer articulatory skills. However, given the possible suppression effects this is a tentative explanation and therefore cannot be interpreted.

The GFTA was not a significant predictor at any step of the regression analysis with d-prime, nor did the GFTA interact with d-prime as a function of age, verified by the correlations between GFTA and d-prime. In Edwards et al.'s (2002) study, the GFTA raw score contributed additional unique variance to d-prime, indicating that articulatory ability impacted on speech discrimination ability, consistent with proposal that there is an ongoing interdependency between speech output (articulatory maps) and speech input (acoustic maps) during speech development (Edwards et al., 2002). However, the children in Edwards et al.'s (2002) study included children with PD and TD in contrast to the children in the current study who were all typically developing and subsequently had a higher overall percentile

ranking ($M = 57\%$) compared to the children with PD in Edward et al.'s (2002) study ($M = 36\%$). It is plausible that children with PD are more likely to be susceptible to poorer performance on tasks that require access to underlying phonological representations, such as speech discrimination, based on the assumption that their deficit is phonological in nature. Consequently, it is possible that the GFTA predicts speech discrimination ability when a phonological deficit is present, as was the case in Edwards et al.'s (2002) study. This proposition warrants further investigation, as it would be of interest to determine at what point, of articulatory skill or age, this interdependency diminishes and also to determine if the relationships are the same for children with a phonological deficit and for children with a speech motor deficit, such as CAS.

Overall the findings from the regression analysis with the GFTA as the primary predictor were tenuous. However, our supplementary investigations relating to the interactions between the GFTA and group indicate that children with more articulatory errors, indicative of less abstracted phonological representations, are more likely to be compromised in tasks that require access to underlying phonological representations to a greater extent than children with better articulatory skills. We also found that age (approaching significance, $p = .065$) predicted articulatory ability for the children with the better articulatory skills, but not for the children with poorer articulatory skills, demonstrating that age predicts phonological competency for children with no errors. However, this relationship is unclear for children with one or more errors and could be due to the small number of older children in the group with one or more errors, as the association may have been compromised by restriction in the age range of children in this group. Further research is therefore needed to tease out these propositions by including a more balanced cohort of younger and older children with articulation errors. Although this finding is tenuous it nevertheless highlights the importance of accurate articulatory maps and the impact they potentially have on developing robust phonological representations, emphasising the role of sensorimotor integration in speech and language development (Perkell et al., 1997; Westermann & Miranda, 2004). Furthermore, the relationship between articulatory ability and access to underlying phonological representations implies that age may have a greater influence on

phonological competence only when a certain level of articulatory ability has already been achieved.

The findings from the regressions with the SVrt were varied. The SVrt was not a significant predictor of nonword repetition accuracy or the phonotactic frequency effect. However, the SVrt predicted picture-naming reaction time at step 1 and 2 of the regression analysis, after accounting for vocabulary and age, respectively. It was not a significant predictor at step 3, nor did it interact with age indicating that the relationship between the PNrt and SVrt did not vary as a function of age. The SVrt was intended to target execution of the speech motor plan, which we presumed would predict output phonological encoding (as measured by the PNrt) after accounting for other covariates. However, the validity and reliability of the SVrt as a pure measure of speech motor execution was not supported. The SVrt revealed a word frequency effect and given word frequency effects are presumed to be located at the level of lexical retrieval this questioned the specificity of this task. We also hypothesized that the relationship between SVrt and PNrt would vary as a function of age, with the SVrt having a greater influence on PNrt for the younger children compared to the older children. This was not shown as demonstrated by the lack of a significant interaction between age and SVrt. It is plausible that the developmental differences in the predictive relationship between speech motor ability and phonological encoding skills was not detectable in the children in the current study. It could also mean, given both are naming tasks, that the association between SVRT and PNrt reflects individual differences due to a common or shared factor linked to the naming process, such as visual perceptual processing or task related cognitive control, independent of variance explained by age and vocabulary.

The SVrt did not predict speech discrimination ability at step 1 or 2 of regression analysis but was a significant predictor at step 3 accounting for 5% of unique variance in d-prime, over and above that predicted by vocabulary and age. It did however interact with age, indicating that the relationship between SVrt and d-prime varied as a function of age. This interaction should be treated with caution given the SVrt did not predict speech discrimination ability in steps 1 and 2 of the regression, although *p* values were close to reaching significance. Nevertheless, when we explored this interaction, split by age, a significant relationship with SVrt emerged for the younger group only, suggesting that as the SVrt increased speech

discrimination ability decreased for the younger children. This was not the case for the older children, indicating that SVrt did not impact on speech discrimination for the older children, suggesting greater autonomy of the speech input and output systems in older children. The SVrt was implemented as a measure of the efficiency of the speech motor system, with faster reaction times indicating more efficient processing in initiating preplanned verbal responses. This is in keeping with our understanding that speed of response is a reflection of efficiency at multiple levels of the speech system involved in picture-naming. The different relationship between the SVrt and age for the younger and older children suggests a tighter association between speech motor execution (SVrt) and speech discrimination ability (d-prime) younger in younger children. However, since the validity of the SVrt as a pure measure of speech motor execution is tenuous, this warrants further investigation and refinement of a measure of speech motor execution for Study 2.

The CVrt did not predict nonword repetition accuracy or the frequency effect. It did however predict picture-naming reaction time and speech discrimination ability at step 1 of the regression analysis. The CVrt accounted for 11% of unique variance in PNrt over and above that accounted for by receptive vocabulary. For speech discrimination ability it was approaching significance ($p = .06$), at step 2 of the regression when age was included in the analysis indicating that as CVrt increased speech discrimination ability decreased. The CVrt did not interact with age for either the PNrt or speech discrimination ability indicating that neither reaction time nor speech discrimination ability varied as a function of age. The lack of significance when vocabulary and age was included, and no interaction with age, and the absence of a word length effect, makes it unclear as to whether the CVrt is a suitable choice for being a measure of phonetic encoding efficiency or speech motor planning.

Limitations and Future Directions

This study had a number of limitations. First and foremost the sample size was relatively small ($N = 47$). Longitudinal studies have greater potential to reveal the causal relationships between different processing abilities without the confounding factors associated with cross sectional studies, implemented in the current study.

Additional limitations relate to the tasks design and stimuli used for some of the experimental tasks. The picture-naming task included only one SOA and used onset primes thereby potentially compromising the likelihood of the groups differing

in phonological facilitation effects. Further research is needed to determine whether phonological facilitation and inhibition effects in typically developing children can be observed by altering the task design to include two or three SOA's and rimes and onsets simultaneously, as auditory primes. The simple and choice reaction time tasks had a limited number of trials for each test item and a limited number of items contributing to the mean reaction time. This may have compromised the reliability of these measures and might have undermined the length effect and potentially introduced effects associated with intrinsic differences in the articulation onsets of those sounds. Furthermore, the children were observed to be quite bored with these tasks in general. Greater consideration to intrinsic differences in articulation onset of the different test items through a larger variety of items would be of value. Although the SVrt showed promise in demonstrating a unique developmental relationship between speech motor development and input phonological processing skills, a different methodology that includes more trials and a greater variety of items could be used. A delayed picture-naming task would fulfill these requirements.

This study implemented regression analysis and in doing so demonstrated the merit of using this type of analysis to reveal differences in constraints between levels of processing in the developing speech and language system. Regression analysis has considerable potential to reveal developmental changes in the different levels of processing of the speech system and the interaction between these different levels. The regression analysis revealed that age predicted nonword repetition accuracy for the children with no articulation errors, but not for the children with articulation errors. Likewise, age predicted picture naming reaction time for children with no errors, but not for the children with articulation errors. This suggests that there was a tighter coupling between articulatory ability and these measures of phonological competence for children with poorer articulatory skills, and indicating that age had a greater influence on phonological competence when a certain level of articulatory ability had already been achieved. Consequently, it would be of interest to determine at what stage of articulatory or phonological development this occurs. Future research could include investigating the role of articulatory ability in the development of phonological competency and how this relationship varies with age.

Conclusion

The specific aim of Study 1 was to explore the relationship between

phonological competency and speech motor ability in younger and older children with TD and to determine if this relationship changed with age.

This study replicated a number of findings from previous research. Nonword repetition accuracy and speech discrimination ability were better and the phonotactic frequency effect smaller in older children. Moreover, vocabulary was the best predictor of nonword repetition accuracy and the phonotactic frequency effect, over and above that predicted by age, replicating a number of studies (Edwards et al., 2004; Munson, Edwards, et al., 2005b; Munson, Kurtz, et al., 2005). Vocabulary also predicted both picture-naming reaction time and accuracy, over and above that predicted by age, consistent with the view vocabulary is a driving force in speech and language development resulting in phonological representations becoming more refined and autonomous with this lexical development (Metsala, 1999).

Despite the null finding from the picture-naming task in relation to the lack of an age difference in the phonological facilitation effect between the younger and older children with TD, the inclusion of a measure of output phonology was still warranted in Study 2. The task was amended to include two SOA's and used auditory rimes instead auditory onsets as primes. In relation to the speech motor measures the SVrt revealed more robust findings overall. The findings from the CVrt were ambiguous given the length effect predicted for the CVrt did not materialize, therefore questioning the suitability of the CVrt measure as a measure of speech motor programming and execution. A delayed picture-naming task was developed to assess speech motor execution skills to overcome some of the limitations of the SVrt. Prior to undertaking Study 2 with children with CAS and PD we needed to establish an accurate and transparent classification protocol. This resulted in a systematic review of classification protocols used in the research literature in CA and development of a classification protocol. The following chapter gives a detailed account of this systematic review and Chapter 4 provides details on the classification protocol.

Chapter 3

Identifying Childhood Apraxia of Speech: A Review and Investigation of Diagnostic Features

Introduction

As stated in Chapter 1, CAS is a multi-deficit speech disorder and is diagnosed based on a cluster of symptoms (ASHA, 2007). However, numerous features have been identified as consistent with a diagnosis of CAS resulting in a diverse set of features selected for classification purposes, clinically and in research (Forrest, 2003; McCabe et al., 1998). Consequently, classification protocols implemented by researchers are diverse and are typically based on clinical judgements, without operational definitions. This occurs not only in relation to the features selected but also the number of features deemed necessary to warrant classification as CAS. Despite many researchers acknowledging the ongoing debate with regard to the most prevalent features of CAS, and the precise origin of CAS, not much has been achieved with regard to developing a specific protocol that can be used for classification purposes (Bahr, 2005; Bahr et al., 1999; Forrest, 2003; Shriberg et al., 2012; Thoonen et al., 1997; Thoonen, Maassen, Wit, Gabreels, & Schreuder, 1996). A more stringent protocol with operationally defined features is needed to advance our understanding of CAS, reiterating comments made by ASHA's technical report (ASHA, 2007) and some researchers alike (Shriberg et al., 2012).

ASHA (2007) and RCSLT (2011) have released position statements on CAS that give an overview of CAS to date. ASHA's (2007) technical report lists the three core features required for a diagnosis of CAS, these include; (a) inconsistent errors on consonants and vowels in repeated productions of syllables or words; (b) difficulty transitioning between sounds and syllables, and (c) disrupted prosody. Other than these two position statements the only other detailed appraisals undertaken in this area include a systematic review that assessed the efficacy of intervention targeting children and adolescents with CAS (Morgan & Vogel, 2009) and a systematic review that looked at treatment outcomes for children with CAS (Murray et al., 2012). Both of these reviews were undertaken subsequent to the release of the ASHA position statement in 2007 and what is evident from these studies is that the classification protocols used were more stringent with regard to the methodological approach of the papers reviewed, consistent with the recommendations made by ASHA in their technical report. Consequently, none of the papers in the Morgan and Vogel's (2009) review met methodological inclusionary criteria and only six papers fulfilled inclusionary criteria in Murray et

al.'s (2012) review. This suggests that the release of the ASHA statement has impacted not only on the features considered relevant for classification of CAS (namely those identified in the review), but also with regard to researchers being more discerning when it comes to methodological approach and classification protocols. Accordingly, the ASHA statement has provided a platform for advancing research in CAS by identifying the most relevant features for diagnosis of CAS and highlighting the significant shortcomings with regard to methodology implemented in this area of research. However, despite this advancement ASHA concedes that operationalized features of CAS have yet to be agreed upon that can be reliably and validly used for diagnosis of CAS (ASHA, 2007).

Aims and Rationale

The basis for the current review was to assess classification protocols used in research in CAS. A systematic review methodology was used to examine the classification protocols used in research to classify children as having CAS. The aim was to identify the most prevalent features used to classify children as having CAS and how many features a child needed to have to warrant classification of CAS, and which of these features had been operationalized. A further aim was to examine changes in the prevalence of CAS related features over time. The primary motivation for this review was to choose an appropriate set of features to develop an operationalize classification protocol to help ensure that participants with CAS identified in Study 2 had the characteristics, that are, to some extent representative of the population of children that researchers have previously classified as such.

According to the principles of scientific research, dependent variables must be defined in unambiguous terms by describing the method of measurement, including the tools used and procedures followed to obtain that measurement (Lum, 2002; Portney & Watkins, 2011). Consistent with this perspective, an operational definition describes a variable according to its unique meaning within a specific study and should be sufficiently detailed to ensure another researcher can replicate the procedure (Portney & Watkins, 2011). In the context of CAS, the variables being measured relate to the features of CAS, therefore, in order to validate the presence or absence of a feature, a clearly defined measurement needs to be provided for each feature. For example, one well cited feature of CAS is the inconsistent articulation of familiar utterances. The presence of the inconsistency feature has

been operationalized by a cut-off score (greater than 40%) on the inconsistency subtest of the Diagnostic Evaluation of Articulation and Phonology (Dodd, Hua, Crosbie, Holm, & Ozanne, 2002). However, some features that occur in CAS rely more on subjective, expert clinical judgment, for example, the presence of groping (see Table 16 for definition). In these instances, the clinician/researcher needs to clearly describe how the feature is identified to ensure that the reader understands the researcher's conceptualization of the feature in question (Portney & Watkins, 2011). If participants are not classified according to the rules of scientific research, that is, with operationally defined features, then this diminishes the validity of the findings and calls into question the reliability of the population being investigated.

Forrest (2003) and McCabe et al. (1998) investigated the features used to classify children as having CAS, implemented by clinicians. They revealed a significant number of features used for classification purposes and also significant variability in the application of these features. Consequently, a number of gaps exist in our knowledge with regard to the features of CAS. For example, it would be useful to know the prevalence and variety of classification features. This would help researchers know the extent their sample fits with previous research in the field and in doing so would potentially give greater consistency to the children classified as CAS. Furthermore, it would be of benefit to determine if the feature-based approach has changed over the years and to ascertain if and how CAS features have been operationalized. An operationalized feature includes the specific criterion that is used to identify whether a feature is present. A review of the literature adopting a systematic review methodology can help provide this information and fill these gaps.

This review covered the period from 1993 to 2013. This review resulted in the development of the protocol that was used for subject recruitment for this PhD, which occurred in 2013. A significant amount of research was undertaken during this period in CAS and CAS was also recognized as a discrete disorder, separate from SSD in general (ASHA, 2007).

Method

The search strategy followed the guidelines as outlined by the PRISMA Group (Moher, Liberati, Tetzlaff, & Altman, 2009).

Types of Studies

This review included all research papers published in peer-reviewed journals that investigated CAS from 1993 to 2013, which was when this review was undertaken. Relevant databases were searched using key words; childhood apraxia of speech, developmental apraxia of speech (DAS) and developmental verbal dyspraxia (DVD), a term used in the UK, the two latter names are names that have been associated with this disorder in the past. Apraxia of speech was also used as search key words as some have combined both children with CAS and adults with AOS.

Data Extraction and Management

Titles and abstracts were independently searched and screened on line using the following databases: Medline (ProQuest (1972 to 2014), PsychInfo, ScienceDirect, Scopus, Web of Science, Wiley Online Library, AMED (Ovid 1985 to April 2014), CINAHL, Embase (Ovid 1974 to April 2014). Copies of articles were obtained and assessed against inclusion criteria. Inclusionary criteria included: (a) papers were peer reviewed published articles with the population being investigated, being children having CAS or suspected CAS, (b) papers contained information on how children were diagnosed/classified as having CAS and (c) papers were written in English.

Of the 104 papers reviewed 48 papers were excluded from the review, 18 (38%) related to SSD and were not specific to CAS; 12 (25%) had recruited children who were not assigned to experimental groups, that is the participants were not classified as CAS, therefore no classification criteria were reported; 18 (38%) papers did not report features consistent with CAS to classify participants in relation to features; and 3 (6%) were not investigating CAS. Details of studies excluded from the review can be seen in Table B1 in Appendix B.

Fifty-six (40%) of the papers reviewed reached criteria for inclusion. Data extracted included author(s), year, features used for diagnosis, number of features required for a positive diagnosis (when reported), operational measures used to identify or quantify the presence of features (when reported), number of participants, participants' age range, and other groups included in the study. Details of the studies included in the review can be seen in Appendix B2.

Classification Protocols

Participants in the studies reviewed were children that were diagnosed as CAS, or suspected CAS (sCAS), based on the presence of features that were deemed to be consistent with this disorder. Classification protocols varied with some papers reviewed based on a referral from a speech language pathologist (SLP) with features identified as consistent with CAS but not specified in relation to the specific participants. In other papers features of CAS were mentioned in general terms (i.e., features of CAS were listed) but the number of features required to warrant classification as CAS were not reported. Many studies, predominantly studies with five or less subjects, adopted a single subject design methodology whereby features were reported per subject. In these instances features were included in the review but when scores were provided on specific assessments they were not included as operational definitions, unless specified as such.

Some of the papers included in the review adopted an alternative approach with regard to classification of participants by implementing screening protocols specifically designed for CAS. Three of these papers (Marion et al., 1993; Marquardt et al., 2002; Sussman et al., 2000) used the Screening Test for Development Apraxia of Speech (Blakeley, 1983), which gives an overall probability score of the child having CAS. This screening tool has four subtests, one that measures the discrepancy between expressive and receptive language, and the remaining subtests assess prosody, verbal sequencing and articulation and provides the examiner with a probability of the child having CAS. In these instances, the features reported were included in the review, however, because the specific details on performance for each feature was not reported (other than an overall probability score) these could not be reported in the review as operationally defined features.

A number of researchers implemented Ozanne's (1995) Diagnostic Model for classification purposes whereby participants needed to present with processing errors at the level of phonological planning (Lewis, Freebairn, Hansen, Taylor, et al., 2004; Mc Neil et al., 2009; McLeod, 2009; Moriarty & Gillon, 2006). The Khan-Lewis Phonological Analysis (KLPA) was another tool used to assess if a child had deficits consistent with CAS. This tool assesses 10 developmental phonological processes yielding standard scores, percentiles and test ages equivalents. However, when this tool was implemented scores were not reported on which phonological processing

errors were present and for this reason these data could not be included in the review as operationally defined features. The Apraxia Profile (1997) was also used for classification purposes for one study (Froud & Khamis-Dakwar, 2012). This tool assesses oral structures, volitional verbal and nonverbal movement, articulation and prosody. In this instance scores on subtests were reported for each participant, and although features were included in the review, these were not counted as operationally defined features since they did not specifically state a cut-off score to either signify the presence or absence of the feature. A number of the papers that adopted a screening profile for CAS, identified features but did not provide cut-off score for the features listed, consequently, these were not included as operationally defined features.

Some of the features reported for classification purposes in the research papers were ambiguous in relation to the specific deficit they were referring. In these instances it was difficult to ascertain the specific speech deficit the authors were referring. For example, Groenen et al. (1996) cited high incidence of context related sound substitutions, giving the example of metathetic errors, whereas Lewis et al. (2004) referred to metathetic errors as deviant errors. Shriberg et al. (2003) also listed metathetic errors as a feature of CAS, however, in this instance they categorized these errors as sequencing errors. These features were discussed and analysed by the authors of this paper and a specific deficit was agreed upon. This was undertaken to ensure that features that were ambiguous were assigned, as accurately as possible, to the most appropriate deficit. A list of the assigned feature labels along with the original labels used by the authors can be seen in Table B4 Appendix B. This table is not a complete list of the features associated with CAS, it comprises only features that had various labels associated with one specific speech deficit. Authors' full details are listed in Appendix B6. A description of each of these features is explained in Table 16.

In summary, the papers included in this review needed to explain how children with CAS met the classification criteria to be classified as such. In order to do this the features (operationalized or not) consistent with a diagnosis of CAS were required to be listed in the methodology section. For a feature to be operationally defined a specific demarcation point or score to warrant presence of the feature was required. The number of features necessary to warrant classification as CAS was

included in the review when reported. Full details of the studies included in the review, including author(s), year of publication, details on participants and features (with and without operational definitions) can be seen in Appendix B2.

Results

Features of CAS

A list of features used for classification of participants as having, or being consistent with having CAS, was compiled from the papers reviewed. The list was extensive and varied, and included many different labels describing speech deficits observed in children with CAS. An initial list, respecting the labels used by authors, comprised 44 features. In an attempt to avoid duplication and inflation of number of features associated with CAS, those that could be seen to refer to the same deficit or speech characteristic were assigned to a common label. For example, limited consonant/vowel repertoire, low PCC and many phonemic errors were interpreted as one deficit, which we called limited phonetic inventory. This may or may not been the intention of the respective authors, however, since no details were reported regarding specific criteria, these definitions were combined to reflect a general feature associated with limited phonetic inventory. The intention was to identify discrete features associated with CAS as used in the broad published research literature for classification purposes. A list of the merged features, comprising 18 features, can be seen in Table 16. A list including the papers that cited these features can be seen in Appendix B3.

Prevalence of CAS Classification Features

The 18 features considered relevant for classification as having CAS are listed in order of prevalence in Table 17. This list is ordered according to frequency of use in the papers included in the review depicting the most frequently used features for classification of children as having CAS reported in published research literature during the period 1993 to 2013 (See Appendix B2 for full list of publications). As can be seen in Table 17 there is a large range in prevalence values for these features. The most prevalent feature, sequencing deficit, was reported in over 80% of the papers reviewed during the period, compared to the least prevalent feature, poor phonemic awareness, reported in less than 2% of the papers reviewed. Table 17 shows that only five features were reported in almost half of the papers reviewed,

indicating only a small number of features were consistently used.

Table 16

Description of Final 18 CAS Features Based on Systematic Review of CAS Research Publications from 1993-2013

Feature	Definition
Sequencing deficit	Difficulty producing sequences of sounds/syllables
Inconsistent speech	Differences in multiple productions of the same target word or syllable (ASHA, 2007)
Limited phonetic inventory	Reduced ability to produce sounds appropriate for age
Prosody-Stress Errors	Inappropriate assignment of stress (including intonation, pitch, inappropriate prosodic or metrical patterns of strong versus weak syllables)
Vowel errors	Errors producing vowels
Omissions/simplifications	The omission of a sound or reduction of a cluster including syllable deletions
Groping	Difficulty initiating a word resulting in oral search
Reduced DDK rate	Reduced ability to produce rapid repetition of same or alternating syllable
Deviant errors	Errors that do not follow typical development trajectory (e.g., final devoicing – “dog” produced as “dok”, initial consonant deletion – “dog” produced as “_og”)
Difficulty imitating sounds/words	Inability to reproduce sounds or syllables when model provided
Unintelligible speech	Difficult to understand (including words and sentences)
Gap between receptive and expressive language	Self explanatory
Context related errors	Switching sounds within syllables or words
Voicing errors	Producing a voiceless sound in place of a voiced sound or vice versa
Slow response to treatment	Poor progress with therapy
Slow speaking rate	Slow speaking rate (this relates to normal speaking condition, different from test conditions such as in DDK task, where nonsense syllables are produced as fast as possible)
Inappropriate loudness	Variability in loudness and inappropriate loudness for speaking environment
Poor Phonemic Awareness	Diminished knowledge of sound and sound sequences and ability to manipulate sounds and syllables (different from a sequencing deficit)

Number of Features

Of the 56 papers included in the review 23 (40%) did not report the minimum number of features required to warrant classification as CAS. For the 34 papers that did include a number of features needed to meet classification criteria the number of features ranged from three to eight. These statistics reflect the condensed features, namely the 18 features as opposed to the original 44 features identified. Consequently, there may be discrepancies in the data with regard to the number of features cited in the original articles and those reported in this thesis.

Table 17

Features Used to Classify Children as CAS and Number and Percentage of Papers Citing Each Feature

Feature	# cited	%
1. Sequencing deficit	47	82.5
2. Inconsistent speech	45	78.9
3. Limited phonetic inventory	34	59.6
4. Prosody-stress errors	28	49.1
5. Vowel errors	27	47.4
6. Omissions/simplifications	23	40.4
7. Groping	23	40.4
8. Reduced DDK rate	14	24.6
9. Deviant errors	13	22.8
10. Difficulty imitating sounds/ words	12	21.1
11. Unintelligible speech	10	17.5
12. Gap between receptive and expressive language	10	17.5
13. Context related errors	7	12.3
14. Voicing errors (not including vowels)	5	8.8
15. Slow response to treatment	5	8.8
16. Slow speaking rate	4	7.0
17. Inappropriate loudness	3	5.3
18. Poor phonemic awareness	1	1.8

Of the 34 papers that reported how many features were required to be classified as having CAS, four features were reported in 14 of these papers; six features were reported in six papers; three features were reported in seven papers; eight features were reported in five papers, five features were reported in the one paper and seven features were reported in just one paper. A number of the papers in this review reported using the same set of features used in previous papers, in most cases due to

the overlap of author(s) or because more recent papers used a subsample from previously undertaken studies.

Operational Definitions

Seventy five percent of the papers in this review failed to use any operational definitions, as defined in this thesis, for the features required for classification as CAS. There are a number of standardized tools that can be used to operationally define some of the most common features of CAS, including articulation ability, inconsistency of speech, sequencing ability, DDK rate, prosody-stress errors and oromotor skill. A list of the tools used to measure these features and the number of papers that cited used these tools are set out Table 18.

Despite the availability of a number of tools that measure phonetic inventory, only nine papers implemented a tool to operationalize this feature. Of the 45 papers that cited inconsistent speech only five papers used an operational definition for this feature. The most frequently used tool to measure inconsistency in the papers reviewed was the Diagnostic Evaluation Articulation and Phonology (DEAP) Inconsistency Subtest with a score of 40% or more indicating inconsistent speech, as recommended in the manual. An additional six papers that used the DEAP inconsistency subtest failed to report a specific cut-off to identify the presence of this feature. There are a number of tools available that have subtests that measure sequencing ability. These include the DEAP (Dodd et al., 2002), Verbal Motor Production Assessment (Hayden & Square, 1999) and the Oral and Speech Motor Control Protocol (Robbins & Klee, 1987). A sequencing deficit was cited 47 times in the papers reviewed, however, only five of these papers reported an operational definition for the feature.

Prosody/stress errors is a feature that has been predominantly judged perceptually in research in CAS, consequently, only one of the thirty papers that included prosody-stress errors as a feature of CAS used a tool to define this feature. Velleman and Shriberg (1999) used the Prosody Voice Speech Profile to identify inappropriately placed or excessive stress. Similarly, only one paper reported the use of a tool with an operational definition that measured oromotor skill. Oromotor deficits have been also predominantly observed by trained clinicians in CAS, however, the oromotor subtest of the DEAP provides an operational definition for a deficit in oromotor control, demonstrated by a score of less than eight indicating a

deficit. DDK rate was measured using two tools, the OSMCP (Robbins & Klee, 1987) and the Apraxia Profile Prosody Subtest (Hickman, 1997). Both of these tools were only used on one occasion, consequently of the 14 papers that reported “reduced DDK rate” as a feature of CAS only two operationally defined this feature reporting a specific measure to confirm the presence of this feature.

Table 18

Final CAS Features that included Operational Definitions

Feature	Operational Definition	# papers
Sequencing	DEAP oromotor subtest SS < 8 VMPAC < 85%	5
Inconsistency	DEAP IS => 40% on inconsistency subtest CSIP > 25%	5
Limited phonetic inventory	< 5% on the GFTA DEAP articulation subtest >= 1.5 SD below mean BBTPOS >= 1.5 SD below mean TDSTA 2 SD below mean	9
Prosody-Stress	Prosody Voice Speech Profile (PVSP) 80% or less of Appropriate Stress scores for Prosody-Voice Code 15; Excessive/Equal/Misplaced Stress.	1
Groping	DEAP oromotor subtest SS < 8	1
DDK rate	OSMCP >= 2 TFS SD below mean Apraxia Profile Subtest (Hickman, 1997)	2

Note. BBTOPAS = Bankson-Bernthal Test of Phonology Articulation Subtest (Bankson & Bernthal, 1990); CSIP = Consonant Substitute Inconsistency Percentage (Iuzzini & Forrest, 2010); DDK = diadochokinetic; DEAP IS = Diagnostic Evaluation of Articulation and Phonology Inconsistency Subtest (Dodd et al., 2002); GFTA = Goldman Fristoe Test of Articulation (Goldman & Fristoe, 2000); OSMCP = Oral and Speech Motor Control Protocol (Robbins & Klee, 1987); SD = Standard Deviation; SS = Standard Score; TDSTA = Templin Darley Screening Test of Articulation (Templin & Darley, 1969); TFS = Total Function Score on Apraxia Profile Subtest (Hickman, 1997); VMPACSS = Verbal Motor Production Assessment Sequencing Subtest (Hayden & Square, 1999).

Single Subject Methodology

Of the total 56 papers included in this review 29 of these (51% of total 56 papers) adopted a single subject design methodology; five of these papers (9% of the total 56 papers) were single case studies and 24 of these papers (42% of total 56 papers) had five or less participants. The remaining 28 papers (49% of total 56 papers reviewed) had six or more participants. Participants ranged in age from at 3.2 years to 14.4 years. For the 29 papers that adopted a single subject design methodology performances by participants on assessments administered were reported on an individual basis, providing a detailed account of individual skills and deficits, however operational definitions were not specified.

Alternative Approach For Classification

Nine papers (16%) included in this review adopted an alternative approach using screening tools to classify participants. The features reported in these papers were included in this review, however, operational definitions were not included as specific details pertaining to individual features were not reported.

Changes to the features of CAS used in research over time

Given the time frame covered in this review it was of interest to determine if there had been any changes to the type and prevalence of CAS related features used for classification of participants during this period. To ascertain potential shifts in features used for classification we divided this period into three equal periods; 1993 to 1999, 2000 to 2006 and 2007 to 2013. The prevalence of features for each period is depicted in Table 19. The number of papers included in the review differed for each of the periods; 11 papers were included for the first period (1993-1999), 25 for the second period (2000-2006) and 21 for the third period (2007-2013).

As can be seen in Table 19, sequencing deficit and inconsistent errors were the most prevalent features across each of these periods. Prosody/stress errors and reduced DDK rate increased in prevalence. A Spearman correlation between the % score for each of the features for each period was undertaken. This revealed a strong positive correlation between all periods. The correlation for the first and second period was .90, second and third period was .82 and first and third period was .72. The top eight features were relatively consistent for each of the periods covered by the review. The remaining 9 features were variable in prevalence, with a number of the features (e.g., receptive-expressive vocabulary gap and slow response to treatment) featuring less frequently over time, whereas others (e.g., voicing errors and slow speaking rate) increased in use. The number of features required to warrant classification of CAS did not vary greatly for the three different periods covered by the review. However the most common number of features required for classification of CAS was four features, which changed to three features for the final period. More importantly, the number of operational definitions reported increased considerably from the first period to the last. Only three operational definitions were reported in the first period and five in the second, which increased substantially to 15 operationally defined CAS features reported in the most recent period.

Table 19

Features Consistent with CAS Used in Research Literature During the First Period (1993 to 1999), Second Period (2000 to 2006) and Third Period (2007 to 2013).

Feature	1993-1999 (n = 11)		2000-2006 (n = 25)		2007-2013 (n= 21)	
	#	%	#	%	#	%
Sequencing deficit	11	100	18	72	18	86
Inconsistent errors	10	91	17	68	18	86
Limited Phonetic Inventory	7	64	16	64	11	52*
Vowel Errors	6	55	11	44*	10	48
Groping	5	46	8	32*	10	48
Omissions/Simplifications	4	36	12	48*	7	33*
Deviant errors	4	36	3	12*	6	29*
Prosody-Stress Errors	4	36	12	48*	12	57*
Difficulty imitating	4	36	6	24	2	10*
Gap expressive/receptive	4	36	5	20	1	5*
Unintelligible speech	3	27	5	20	2	10*
Context related errors	3	27	2	8*	2	10*
Reduced DDK rate	2	18	6	24*	6	29
Inappropriate loudness	2	18	1	4*	0	0
Slow response to treatment	2	18	2	8*	1	5*
Poor phonotactics	1	9	0	0	0	0
Voicing errors	0	0	0	0	5	24*
Slow speaking rate	0	0	0	0	4	19*

Note. Features are ordered in terms of frequency for the first period. An asterisk * denotes ranking of feature in second and third periods not consistent with ranking order for first period.

Discussion

CAS is by nature a complex disorder that is fraught with controversy with regard to etiology, diagnosis and treatment (ASHA, 2007). Research in this field has been attempting to find diagnostic markers of CAS for over 30 years. The lack of defining features was identified as a significant problem and the challenge set by Guyette and Diedrich (1981) to develop a set of features, specific to CAS, has yet to be realized. Numerous studies have attempted to do this and in doing so have identified an extensive list of potential markers, however none have achieved a “gold standard” for CAS (Shriberg et al., 2012). This gold standard pertains to a set of features that should be unique to CAS and not present in other children with SSD (e.g., PD), if in fact CAS is distinct from idiopathic SSD. Research is therefore needed in both domains, understanding the disorder and establishing valid and reliable methods of diagnosis.

The purpose of this review was to ascertain the most prevalent features used in peer reviewed research literature on CAS for classifying children as having CAS. Our principal objective was to develop a classification protocol based on good practice, using a representative set of features used in the research literature, which could then be used in Study 2 to classify participants as CAS. As a result of the detailed review we not only found that there was significant variability between researchers regarding the features they deemed most pertinent to a diagnosis of CAS, but also the number of features researchers deemed necessary to validate a diagnosis of CAS. Even more surprising was the degree of ambiguity with regard to how features were identified as being present, mainly due to the lack of operationally defined features.

Of the 104 papers reviewed 56 reached criteria for inclusion and identified 44 features used for classification of CAS, similar to the number of features identified by Forrest (2003). However, some of the features identified in the current review were ambiguous with regard to the particular deficit they were referring. For example, increase in number of errors in longer utterances and predominant use of simple syllable shapes, unless clearly defined, could be interpreted as the same underlying deficit, namely a sequencing deficit. However, in the context of a theoretical model such as the WEAVER model (Roelofs, 1997) these features could reflect a problem with the underlying representations, a problem with the mental

syllabary, or a problem at the level of phonetic planning, indicating very different deficits. Ambiguities such as these highlight the need for operationally defined features to remove any confusion how a feature is deemed present. We merged features to a common label that appeared to refer to the same deficit to minimize over-inflation of features present. We also wanted to ensure that the features reflected the common characteristics observed in children that researchers have previously classified as having CAS and thereby generate a list of features that could be (if not already) operationally defined.

There were also some incongruities between authors with regard to the level of processing implicated when referring to the same feature. For example, Lewis et al. (2004) interpreted metathetic errors as deviant errors, whereas Shriberg et al. (2003) interpreted the metathetic errors as a sequencing deficit. However, in these instances, the relevant features were categorized according to each individual author's theoretical perspective when specified, which also contributed to ambiguity and subjectivity regarding classification of participants.

None of the merged 18 features were cited in 100% of the papers when looking at all of the papers reviewed for the 20-year period. When we looked at each of the different periods, the sequencing deficit feature was reported in 100% of the papers reviewed for the first period. However, the term sequencing deficit was given to a range of features identified in the 11 papers included for this period. These features included difficulty sequencing, inability to produce complex phonemic sequences and increased errors with increased complexity. However, as previously mentioned, the inability to produce a complex sequence could be due to a motor deficit or underspecified phonological representations.

Of the 18 features that emerged from the review the first eight features were consistent for the three periods covered, indicating that these features reflect the primary population of children classified with CAS during the period investigated. The subtle shifts in the order of prevalence of these features related mainly to the prosody-stress errors, indicating that this feature gained more credibility as a key feature of CAS as research progressed. The remaining 10 features varied in their frequency of use, indicating they are not consistently used as markers of CAS. The number of features considered necessary to warrant classification as CAS also varied, ranging from three to eight features, with the most common number of

features being 4, reported in a quarter of the papers reviewed. Moreover, almost half (40%) of the papers reviewed did not report how many features were required to warrant classification as CAS, thus questioning the validity these participants being representative cases of CAS given it is not known how the children met classification criteria.

One of the major findings from the review was the lack of operationally defined features. The absence of operationally defined features is a serious shortcoming given a number of tools are available that can be used to operationally define the features of CAS. As previously stated, it is the author's responsibility to clearly define the variables and justify the operational definition in terms of the purpose of the research (Portney & Watkins, 2011). Consequently, features of CAS should be operationally defined to remove ambiguity and in doing so enable replication. If all features identified in the review were operationally defined then there would be less ambiguity regarding the presence of a feature. Furthermore, by quantifying the features of CAS provides opportunities for ongoing research in CAS, whereby quantitative methodologies can be used to explore the nature of speech deficits associated with CAS in more detail.

A major improvement in research in CAS over the period covered by the review was the considerable increase in the use of operationally defined features for the last period reviewed. The ASHA position statement released in 2007 highlighted the limitations in research with regard to "methodological constraints" hampering our advancement in understanding CAS. This appears to have had a positive impact with some researchers being more stringent in their classification of participants thereby using operational definitions for classification purposes.

Limitations and Future Directions

Features that were defined in ambiguous terms were assigned a label consistent with the authors' viewpoint. Consequently, the interpretation of these features may not be in congruence with other professional opinions. Although age of participants was reported in the published papers (details of which are included in Appendix B2) no data was reported on potential features that may be more prevalent with specific stages of development. Given some of the characteristics associated with CAS have been shown to vary with age (Forrest, 2003) it would be of interest to explore this topic given the broad age range of children included by the papers in this review (3

to 15 years). Furthermore, if a highly productive researcher used features that were not widely accepted, this could potentially inflate the prevalence of said features.

Despite this review identifying the most prevalent features of CAS for the period 1993-2013, it does not validate these features as the best features for differentiating children with CAS from SSD. The features identified in this review reflect the most frequently used features to classify children as CAS and therefore the implementation of these features for classification purposes would be consistent with researchers in this field. By using the most prevalent features that have emerged from this review, researchers have some assurance knowing they have recruited participants who share characteristics with the broader population of children with CAS, recruited by researchers over the past two decades. However, whether these features differentiate between children who are a distinct taxonomic group with a unique core deficit in speech motor control, consistent with CAS, and are distinguishable from other children with SSD needs to be ascertained.

Further research, therefore, is needed to identify which features are likely to be most effective to differentiate between CAS and PD. Some of the features identified in this review are useful in distinguishing speech disorder from typical development, but may be questionable for differential diagnosis. This review has highlighted the most prevalent features of CAS used in the research literature, however, not all these features have clear operational criteria, highlighting this as a potential focus for future research. Future research needs to be more consistent and transparent with regard to classification protocols to ensure that participants are accurately classified as CAS, thereby validating research outcomes. Assessment protocols need to be developed that can be used in the clinical setting, using readily available tools, to ensure children are diagnosed correctly as CAS and treated accordingly. Expert opinion is highly valued, both in research and clinically, but protocols need to have clear operationally defined criteria for classification thereby reducing subjective judgements.

Conclusion

An ongoing problem in research in CAS is the high probability of misdiagnosis, predominantly due to the absence of diagnostic markers for this disorder and confirmed by many inconsistencies with regard to classification protocols used to classify children as CAS revealed by this review. Many of the

papers in this review failed to include the number of features required to warrant classification as such, but more importantly, failed to operationally define these features, highlighting significant shortcomings in classification of children as CAS. For a population, such as CAS that is by definition highly heterogeneous, it is paramount to minimize variability as a result of ambiguous classification protocols. A failing to apply a more rigorous and consistent methodology with regard to classification of participants is a clear shortcoming of research in this field, reiterating ASHA's remark that the lack of clear and stringent diagnostic protocol is the largest impediment to advancement in our knowledge of CAS (ASHA, 2007).

Of the 18 merged features, only eight of these features were consistently used throughout the periods covered by the review. However, some of these features are also evident in children with PD and therefore we need to be more discerning with regard to the features that have the potential to differentiate between children with CAS and SSD. Furthermore, given the shared view of many researchers is that the underlying deficit in CAS is a speech motor deficit, these features need to be explored in this context and given CAS is assumed by many to be a developmental speech disorder, features need to be explored from this developmental context.

This review did not permit validation of the features at a theoretical level. The following chapter focuses on this idea by using exploratory factor analysis to investigate the underlying construct associated with the features of CAS. This type of exploratory analysis has not yet been undertaken, most likely due to the lack of progress in the use of operationalized definitions for features of CAS. Consequently, prior to undertaking this analysis, we needed to firstly, identify the features unique to CAS and then, to operationally define these features. Exploratory factor analysis was then used to evaluate the inter-relationships between the features unique to CAS and to ascertain the underlying construct associated with these features. In the context of CAS a single underlying construct could justifiably be interpreted to reflect an underlying deficit within the speech motor control system.

Chapter 4

Development and validation of a classification protocol for CAS using operationalized features.

Introduction

CAS is a symptom complex with diagnosis based on the presence of characteristic features (ASHA, 2007; Dewey, 1995; Le Normand, Vaivre-Douret, Payan, & Cohen, 2000; Lewis, Freebairn, Hansen, Iyengar, et al., 2004; McCabe et al., 1998; RCSLT, 2011; Shriberg, Green, et al., 2003; Shriberg et al., 2012). The general consensus is that symptoms observed in CAS are consistent with a deficit at the level of speech motor control, however, what this entails is not always clearly defined in the context of CAS. A variety of studies have presented deficits consistent with a speech motor deficit. Maassen (2002) proposed children with CAS had an under-developed mental syllabary, Peter and Stoel-Gammon (2005) demonstrated timing deficits, and a number of studies have demonstrated deviant/distorted transitions between sounds and syllables (ASHA, 2007; Nijland, Maassen, van der Meulen, Gabreels, Kraaimaat, & Schreuder, 2002; Nijland, Maassen, van der Meulen, et al., 2003). Consequently, how the speech system of a child with CAS is specifically impaired and whether a variety of sub-types exist, is still unknown.

Computational models, such as the DIVA model (introduced in Chapter 1) have explored different possible underlying deficits that could result in the broad range of speech deficits observed in children with CAS. The DIVA model consists of feedforward and feedback control loops that interact to optimise learning auditory targets, which in turn refines output speech production (Tourville & Guenther, 2011). This model has been utilized to explore the underlying deficit in CAS by manipulating the feedforward and feedback mechanisms (H. Terband et al., 2009). Overall, the simulation results revealed an increase in coarticulation, speech sound distortions, and searching articulatory behaviour (determined by comparing the formant frequencies at beginning, middle and end of each speech sound) as the reliance on feedback control increased. These findings indicated that the symptoms of CAS could be due to over reliance on feedback mechanisms due to deficient feedforward commands. Terband and Maassen (2010b) expanded on this research by exploring the source of the potentially degraded feedforward commands. They wanted to determine if the underlying deficit was due to impaired somatosensory feedback, consistent with the belief that children with CAS had reduced or degraded oral sensitivity, or if the deficits observed in CAS could be explained by the presence

of neural noise. The simulation results did not differentiate between the two different hypothesized deficits, with both simulations leading to similar characteristics. Although ambiguous, this finding suggests that the derived deficits in the speech motor control system could have emerged due to an underlying interference with the typical development trajectory. Consequently, it is plausible that the disparate features observed in children with CAS could be driven by a shared set of interactions within the emergent mapping processes of the speech motor control system, regardless of the specific origin of the deficit. This suggestion is consistent with the view that the nature of the core speech motor control deficit in CAS is unidimensional (ASHA, 2007; Maassen, Nijland, & Terband, 2010; Shriberg, 2010). The identification of the core features of CAS that reflect this underlying dimension, if it exists, is fundamental in ensuring accurate diagnosis of children with CAS, both clinically and in research, and critical to advance our understanding of this anomalous disorder.

Aims and Rationale

The overall aim of this study was to develop and validate a classification protocol that could be used to differentiate between children with CAS from children with SSD and to utilize this protocol to substantiate classification of the children with CAS and PD in Study 2. The systematic review undertaken in Chapter 3 confirmed that there is some consensus across researchers, over a lengthy period, with regard to certain features deemed consistent with CAS. The most common and consistently used features, for all three periods covered in the review, are listed in order of prevalence in Table 11 in Chapter 3. These features reflect the most common characteristics used for classification purposes but their prevalence does not necessarily validate that these are the best features for differentiating between children with CAS and SSD. Consequently, the final selection of features chosen for the exploratory analysis was based not only on the most prevalent features identified from the review, but also on features that were more likely to have higher levels of specificity for CAS. Furthermore, because the CAS feature variables were quantified (due to being operationalized), the underlying construct of these features could be examined using exploratory factor analysis. This analysis was undertaken to explore the underlying dimensionality of these CAS related features to determine if they reflected a singular latent structure, consistent with CAS being a

unidimensional disorder. If this transpired, then these features could justifiably represent an underlying shared deficit in speech motor control and the factors scores could be used to confirm classification of CAS participants in Study 2, using discriminate function analysis (DFA). DFA is a multivariate regression technique that can be used to classify two or more groups on the basis of a set of continuous or binary predictor variables, permitting the statistical significance and accuracy of the classification to be assessed.

Features of CAS

The final list of features selected for exploratory analysis included inconsistency, prosody/stress errors, vowel errors, simplification/ omission errors (represented by syllable deletions), reduced DDK rate and deviant errors. These features were turned into feature variables using an operationalized procedure and then the level of each operationalized feature variable needed for a child to have that feature was defined. The DDK feature represented the sequencing deficit. The DDK task included both ability to repeat an alternating syllable and speed of repetition, consequently if a child was unable to repeat an alternating syllable then they were assumed to have the sequencing deficit. Limited phonetic inventory was excluded from the analysis as this feature also occurs in children with PD and is not specific to CAS. The groping feature was also excluded from the final list because it is based on expert judgement and did not easily meet the requirement of being operationalized. We included an additional feature, deviant errors, which was the next most prevalent CAS related feature (see Chapter 3, Table 11). Initial consonant deletion (ICD) was used to represent this feature as ICD does not occur in typical speech development and has the potential to be indicative of a deviant speech motor system, consistent with CAS (Bowen, 2009).

Participants for this study were the same children recruited for Study 2. The children with SSD were identified as having CAS or PD by the practicing SLP at the Language Development Centre. Children with TD, also recruited for Study 2, were used as the comparison sample. To determine if a child with CAS and PD had each of the feature variables they were compared to the TD sample. The Crawford and Howell (1998) single case *t*-test was used to determine if a child's score differed significantly from the TD children for each of the feature variables. A *t*-test with an overall *p* value of less than .05 was considered to demonstrate a significant

difference compared to the TD sample and therefore determined the presence of the feature. Each child with CAS and PD were compared individually to the TD sample for each of the feature variables that had a normal distribution.

A number of the features did not have a normal distribution, therefore, a single *t*-test could not be used to determine if the score deviated from the normal sample. For these features, box plots were used as a non-parametric approach to define a score as deviating from the TD sample. Consequently, each feature was defined based on a continuous scale of measurement and the presence of the feature was defined as a score that exceeded a cut-off along each feature dimension relative to the TD sample (using either the single case *t*-test approach or box plot criteria for an extreme score). For each of these features the speech disordered participants were individually combined with the TD sample to determine if their score was an extreme score in relation to the TD sample, if it was then the individual was deemed to have that feature. This method of comparison not only permits a comparison to TD children but with children who have other types of speech sound disorders, such as children with PD. Furthermore, this approach to operationally define the presence of a feature along a continuum allowed quantitative analysis to investigate the relationship between these feature variables.

Method

Participants

Participants were 18 children (four girls and 14 boys) with suspected CAS, 18 children with PD (four girls and 14 boys) and 18 TD children (eight girls and 10 boys). The youngest participant was 4 years 2 months and the oldest 8 years 4 months, the age range was 4 years 2 months to 7 years 11 months for the children with CAS, 4 years 5 months to 8 years 4 months for the children with PD, and 4 years 2 months to 7 years 10 months for the children with TD. The children were recruited after ethics approval was obtained from Curtin University Human Research Ethics Committee (HREC) and parents provided informed consent. The children with CAS and PD were recruited from four Language Development Centers, two in the Perth metropolitan area and two in South West Region of Perth. Children were classified as CAS and PD based on expert opinion from a senior speech pathologist

at each of the centers (classification criteria not reported). The TD children were recruited from a primary school located in the Perth metropolitan area.

All children completed the EVT (Williams, 1997) and the PPVT-IV (Dunn & Dunn, 1997), used to measure vocabulary development; the GFTA 2, used to assess articulation ability and the Diagnostic Evaluation of Articulation and Phonology (DEAP) Inconsistency Subtest (Dodd et al., 2002) was used to measure inconsistency of speech (described below), but also percent consonants correct (PCC). The children with TD showed normal vocabulary development with standard scores of more than 85 (or a percentile rank more than 17) on the EVT and PPVT-IV (Munson, Swenson, et al., 2005). None of the children with TD had a history of speech, language or hearing impairment and all were progressing normally through school, as per parent/ school report. None of the children with CAS or PD had attention deficit disorder or autism, as per parent report. All children with CAS and PD had normal Oral Mechanism Examination (OME), as per speech pathologist report. All participants had English as a first language. Demographic data are presented in Table 20.

One participant, a male 8.6 years of age, was excluded from the CAS group and further inclusion in the study, as he did not meet the inclusionary criteria. He performed more than two standard deviations below the mean on the EVT (standard score of 75) and almost two standard deviations below the mean on the PPVT-III (standard score of 78), indicative of severe language impairment and presented with no features of SSD.

Statistical analysis using one way independent samples ANOVA revealed that groups did not differ on chronological age, $F(2,50) = 2.05$, $p = .139$. They did differ significantly on the EVT standard score, $F(2,50) = 19.22$, $p < .001$. Independent sample t -tests revealed that the children with CAS had a significantly higher mean standard score on the EVT than the children with PD, $t(33) = 2.78$, $p = .009$, $d = 0.97$, and a significantly lower mean standard score than the children with TD, $t(33) = 3.44$, $p = .002$, $d = 1.20$. The children with PD had a significantly lower standard score from the children with TD on the EVT, $t(34) = 6.01$, $p < .001$, $d = 2.06$. The groups also differed significantly on the PPVT, $F(2,50) = 13.65$, $p < .001$. Independent samples t -tests revealed that the children with CAS had a higher mean standard score on the PPVT than the children with PD, but this difference was not

significant, $t(33) = 1.01$, $p = .320$, $d = 0.35$, and a significantly lower mean standard score than the children with TD, $t(33) = 4.05$, $p < .001$, $d = 1.41$. The children with PD also had a significantly lower standard score from the children with TD on the PPVT, $t(34) = 4.63$, $p < .001$, $d = 1.55$. Therefore, the children with CAS and PD, although within normal limits for expressive and receptive vocabulary, both had significantly lower standard scores than the children with TD.

Table 20

Demographic Data and Test Scores for Children with CAS, PD and TD with Standard Deviations (SD) (N = 53)

	CAS		PD		TD	
	(n = 17)		(n = 18)		(n = 18)	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Age in months	68	14	76	13	70	11
EVT SS	97	7	90	8	105	7
PPVT-III SS	97	7	94	9	108	9
GFTA-2 SS	63	17	85	12	105	7
GFTA-2 %-ile	5	5	14	10	51	23
PCC	46	17	81	9	94	6

Note. EVT SS = Expressive Vocabulary Test Standard score (Williams, 1997); PPVT III = Peabody Picture Vocabulary Test III Standard score (Dunn & Dunn, 1997); GFTA-2 SS = Goldman-Fristoe Test of Articulation 2 Standard Score (Goldman & Fristoe, 1986); GFTA-2 %-ile = Goldman-Fristoe Test of Articulation 2 Percentile Rank (Goldman & Fristoe, 1986); PCC = percentage consonants correct.

The groups differed significantly on the GFTA, $F(2,50) = 50.03$, $p < .001$. The children with CAS had a significantly lower GFTA standard score than the children with PD, $t(33) = 4.58$, $p < .001$, $d = 1.59$, and a significantly lower standard score than the children with TD, $t(33) = 9.70$, $p < .001$, $d = 3.38$. The children with PD also had a significantly lower standard score on the GFTA than the children with TD, $t(34) = 6.14$, $p < .001$, $d = 2.10$. The groups differed significantly on PCC, $F(2,50) = 81.12$, $p < .001$. The children with CAS had a significantly lower PCC than the children with PD, $t(33) = 7.67$, $p < .001$, $d = 2.67$, and children with TD,

$t(33) = 11.21, p < .001, d = 3.90$. The children with PD also had a significantly lower PCC than the children with TD, $t(34) = 5.27, p < .001, d = 1.81$.

Operationalized Features

Diagnostic Evaluation of Articulation and Phonology (DEAP). The Inconsistency Subtest of the DEAP was used to assess speech inconsistency. The DEAP was standardized in the UK and Australia in 2001-2002. This assessment has strong test-retest, inter-rater reliability and concurrent validity. The inconsistency subtest requires the child to name 25 pictures, increasing in complexity from single syllable (e.g. shark) to multisyllabic words (e.g. helicopter). The pictures have to be named as three separate trials with each trial being separated by another activity. The DEAP inconsistency subtest was used to operationalize the inconsistency feature, consistent with the current research literature, a score of 40% or greater was indicative of inconsistent speech and validated the presence of the inconsistency feature (Dodd, 2002).

Prosody/ Stress Errors. Prosody/stress errors were rated perceptually by the first author. Ten items were selected from the DEAP inconsistency subtest, which included two single syllable words, five multisyllabic words and three two-word phrases. A list of these words and the rating scale used can be seen in Appendix C Table C1. The 30 items (10 items x 3 trials) were extracted from the continuous recording and listened to via headphones. Judgment was based on whether the word was spoken with appropriate or inappropriate stress. Due to the repetitive nature and simplicity of the DEAP inconsistency subtest some of the children viewed this task as a game and subsequently named the pictures with playful intonation patterns. Consequently, a rating scale was used consisting of three levels; 0 = normal prosody, 1 = unusual intonation (including playful intonation pattern), 2 = severe distortion. A rating of 1 and 2 were assigned to capture the severity of how prosody was altered. A score of 2 was reserved for severe distortions, indicating that the child has difficulty producing a stress pattern. A score of 1 indicated a more moderate deviation, including playful intonation, which does not suggest a limited capacity to produce a stress pattern. The two ratings were included in the analysis to minimize over inflation of this feature, this was based on the view that by including two ratings the listener could be more discerning to whether or not the child had difficulty assigning appropriate stress or whether the child was being playful. Only words that

were rated as having a severe deviation from a normal stress pattern, that is, words given a rating of 2, validated the presence of this feature. The children with TD did not produce any words with a rating of 2 resulting in no variability for these children. Neither the single case *t*-test or box plots could be used to determine if the children with CAS and PD were significantly different from their TD peers for this feature. Consequently, a cut-off score was selected based on the number of words that had received a score of 2 indicating severely distorted prosody for that word. If a child produced three words or more with severely distorted prosody, resulting in a raw score of 6 (or more), then it was assumed that the child had a difficulty in assigning appropriate lexical stress.

Vowel Errors. We included two features that represented vowel errors; these included a measure for vowel inaccuracy and an additional measure for vowel inconsistency. For the vowel inaccuracy feature the total number of errors, that is the total vowel errors (either substitutions or distortions) across all words produced during the DEAP inconsistency subtest, was divided by the total number of vowels produced, taking into consideration syllables deleted and words not attempted. (Some of the children did not name all the pictures presented for all of the trials of the DEAP and therefore these items were not counted in the total number of vowels to ensure percent scores were an accurate reflection of the number of errors.) The proportion of vowel errors was converted to a percentage score for each participant. The vowel error data did not have a normal distribution therefore we could not use a single case *t*-test to compare the children with CAS and PD to the children with TD. Consequently, box plots were used to compare each child's score, one at a time, to the children with TD. Each child's score was compared individually to the TD sample. If the child's score was an extreme score within the box plot when combined with the TD sample then it was deemed that the score deviated significantly from the comparison sample, thereby meeting the criterion for that child having the vowel error feature. In SPSS boxplots an extreme score is three times the interquartile range above or below the top and bottom of each box.

The vowel inconsistency feature was also obtained using the responses from the DEAP inconsistency subtest. Each word was awarded either a 1 or 0 depending on whether the vowels differed or were produced the same (when comparing the three trials of each of the 25 words), regardless of whether they were correct or

incorrect. This raw score was converted to a percent score for each participant and compared to the TD sample using the same box plot criterion. If the child's percentage score was extreme in the box plot when combined with the TD children, then the participant was assigned vowel inconsistency feature.

Omissions/simplifications. Due to low intelligibility of the children with CAS the omission/simplification feature was represented by syllables deleted, as these were reliably detected and indicative of delayed development of syllable and word structure, consistent with observations of CAS (Bowen, 2009). The single word repetitions from the DEAP were used as the speech sample. The number of syllables for all words attempted was counted and number of syllables deleted resulting in a percent score for each participant. The percent score for each participant was compared individually to the TD sample using box plots. If the participant's score was an extreme score in relation to the TD children then the participant was awarded this feature.

DDK. Each child completed a DDK task. For this task each child was asked to produce a rapid repetition of an alternating syllable, as many times as they could in a single breath (i.e., /pataka/), replicating Thoonen (1996). If the participant was unable to imitate this on first trial then they were trained on the single syllables (i.e., /pa/, /ta/ and /ka/), starting with the individual syllables and building to disyllables (i.e., /pata/ and /taka/), and finally attempting the trisyllable. An example was provided for both the single syllables and the alternating syllables as needed. Each child's responses for the standardized assessments was recorded for transcription purposes using an I-phone 5 and later uploaded to an Apple Macintosh for further analysis. The presence of the feature was substantiated in two ways; if the child was unable to produce a trisyllable (i.e., /pataka/) then the participant was deemed to have this feature. Repetition rate was analysed using PRAAT (Boersma & Weenink, 1995) acoustic analysis software, to maximise measurement accuracy by being able to locate the start and finish of the syllable train in the acoustic waveform (with spectrographic display) and from that measure obtain the total train duration and identify the total number of syllables produced. A reliability analysis for the DDK task was not undertaken given the measurement was undertaken using acoustic analysis based on objective visual features. Each child's repetition rate was then compared to the repetition rate of children with TD using Crawford and Howell

analysis. Consequently, children who were (a) unable to produce a trisyllable or (b) had a rate significantly slower than the TD sample then the participant was awarded this feature.

Deviant errors. For this feature we counted the number of ICDs, as these are considered deviant errors that do not occur in TD (Bowen, 2009), and the only deviant error observed in the current sample. The number of initial consonants for all words attempted were counted and number of initial consonants deleted resulting in a percentage score for each participant. The children with PD and TD did not produce any ICD errors, consequently, all the children that did produce an ICD error were deemed to have this feature. The number of ICD's ranged from 3 to 28 for the eight children who presented with this feature (all were diagnosed with CAS).

Procedure

Each child was tested individually in a quiet area of the language development centers or school, respectively, to minimize distractions. The standardized assessments were administered in the following fixed sequence to vary the cognitive demands across the successive tasks: first trial of the DEAP inconsistency subtest; GFTA 2; second trial at the DEAP inconsistency subtest; EVT; third and final trial of the DEAP inconsistency subtest; PPVT-IV; and the DDK task. The experimental tasks required for Study 2 (presented in Chapter 5) were administered in a second session, at least one week later. Participants were given adequate breaks during both sessions to maximize performance and minimize fatigue. Verbal encouragement was given in between task presentation. Each child was awarded with participation stickers and a lucky dip prize on completion the session. Each session lasted approximately one hour for each child.

Inter-rater Reliability

Inter-rater reliability was undertaken by a qualified speech pathologist. All the continuous CAS related feature variables that were obtained from the speech samples recorded from the DEAP inconsistency subtest were assessed for inter-rater reliability. Nine children (three from each group, equating to 17% of the total sample) were randomly selected and rescored on each feature by a different rater. Good levels of inter-rater reliability were found when assessed using Pearson correlation between the original and repeated measurement for each feature variable:

prosody/stress errors ($r = .94$), vowel inaccuracy ($r = .96$), vowel inconsistency ($r = .89$), deviant errors ($r = .97$) and simplification/omission errors ($r = .88$).

Data Analysis

Exploratory factor analysis was undertaken, using principal axis factoring as the method of extraction, to evaluate the relationship among the final selection of CAS related feature variables. In particular, this analysis was used to summarize the communality relationship (i.e., interdependency) between these variables to determine their underlying structure (Gorsuch, 1983). If these features had a strong interrelationship and load onto one factor then the features may be interpreted as relating to a single underlying construct, or representing a “singular concept” (Portney & Watkins, 2011, p. 707). Such a result between these quantified CAS related features would be consistent with CAS being a unidimensional disorder, indicating a single trait like deficit in speech motor control. Furthermore, the strength of the association between the variables within the factor is also of importance. The measure of the degree of generalizability found between variables is referred to as factor loading (Gorsuch, 1983). The size of the factor loading for a particular variable reflects the quantitative relationship between that variable and the underlying construct (Gorsuch, 1983). In relation to CAS, a single underlying construct can be justifiably interpreted to reflect an underlying deficit within the speech motor control system, consistent with Terband and Colleagues (Maassen et al., 2001; H. Terband et al., 2009). In addition, the factor loading for each of the CAS related features variables included in the analysis is indicative of how tightly coupled that feature variable is to this proposed underlying construct.

A multivariate regression techniques using DFA, was then used to predict group membership from the set of predictors (Tabachnick & Fidell, 2007). The combined predictors, number of features each participant displayed combined with each participants factor score, were used to classify children into CAS and non-CAS speech disordered groups providing a more robust and rigorous method of classification of participants for Study 2.

Results

The raw scores for each of the feature variables were converted to percentage scores. Summary statistics for each of these features are presented in Table 21 for

each group, classified as such based on clinical judgment. The sequencing deficit as measured by DDK combined score (i.e., ability and rate) was coded as present or absent, therefore a binary feature and was not included in this table.

As can be seen from Table 21 the children with CAS had a higher percent score for each of the features when compared to the other groups. Statistical analysis using one-way ANOVA was used to determine if groups differed statistically on inconsistent speech, $F(2,50) = 55.81, p < .001$, partial $\eta^2 = .77$. Independent sample t-tests revealed that the children with CAS had a significantly higher mean score for inconsistency than the children with PD, $t(33) = 7.22, p < .001, d = 0.97$, and than the children with TD, $t(33) = 11.92, p < .001, d = 1.20$. The children with PD also had a significantly higher mean score for inconsistent speech than the children with TD, $t(34) = 5.86, p < .001, d = 0.97$.

Table 21

Related Features of CAS for Children with CAS, PD and TD with Standard Deviations (SD) (N = 53)

	CAS		PD		TD	
	(n = 17)		(n = 18)		(n = 18)	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Inconsistent speech (%)	59	17	25	11	7	8
Prosody/Stress Errors (%)	18	23	0	0	0	0
Vowel Errors (%)	13	13	1	1	0	0
Vowel Inconsistency (%)	23	17	1	1	0	0
Syllable Deletion (%)	9	9	3	3	1	2
ICD (%)	4	7	0	0	0	0

Due to the data being severely skewed for the other feature variables, a nonparametric Mann-Whitney U test was used to compare groups. A Mann-Whitney U tested indicated that prosody/stress errors for the children with CAS was significantly higher than the prosody/stress errors for the children with PD and TD, $U = 35.00, z = -3.97$ (corrected for ties), $p < .001$, two-tailed, $r = .67$. The children with PD and TD did not differ on prosody/stress errors as neither produce severely distorted prosody errors. The children with CAS produced significantly more ICD

errors than the children with PD and TD, $U = 81.00$, $z = -3.23$ (corrected for ties), $p = .017$, two-tailed, $r = .55$. The children with PD and TD did not differ on ICD errors, since neither produced any ICDs. The children with CAS had higher vowel inaccuracy than the children with PD, $U = 54.00$, $z = -4.0$ (corrected for ties), $p < .001$, two-tailed, $r = .68$, and the children with TD, $U = 25.00$, $z = -4.44$ (corrected for ties), $p < .001$, two-tailed, $r = .75$. The children with PD did not differ significantly for vowel inaccuracy from the children with TD, $U = 124.5$, $z = -1.46$ (corrected for ties), $p = .239$, two-tailed, $r = .24$. The vowel inconsistency scores for the children with CAS were significantly higher vowel inconsistency than the children with PD, $U = 27.50$, $z = -4.36$ (corrected for ties), $p < .001$, two-tailed, $r = .74$, and TD, $U = 24.00$, $z = -4.53$ (corrected for ties), $p < .001$, two-tailed, $r = .77$. The children with PD and TD did not differ significantly in terms of vowel inconsistency, $U = 151.5$, $z = -0.48$ (corrected for ties), $p = .743$, two-tailed, $r = .08$. Syllable deletion errors for the children with CAS were significantly higher than the syllable deletion errors for the children with PD, $U = 92.50$, $z = -2.02$ (corrected for ties), $p = .045$, two-tailed, $r = .34$, and children with TD, $U = 60.00$, $z = -3.20$ (corrected for ties), $p = .002$, two-tailed, $r = .54$. The children with PD and TD did not differ significantly in relation to syllable deletion errors, $U = 107.0$, $z = -1.89$ (corrected for ties), $p = .085$, two-tailed, $r = .32$.

Prevalence of Features Within Sample

The CAS related features are presented in Table 22 in order of prevalence from left to right that was based just on the children with a clinical diagnosis of CAS. Features that are most common among children with CAS are listed first and those that are least common are listed last. The CAS children are also ordered from top to bottom in terms of their sum of CAS related features. This way of structuring a matrix consisting of a set of items with a binary outcome (columns) for a set of cases (rows) is consistent with Guttman scaling method (Price, 2016) as discussed below. The children in the PD group are also ordered top to bottom in terms of their sum of CAS related features. As can be seen from the table the children with clinical diagnosis of CAS displayed more features than the PD group, although some of the children with PD displayed some of the CAS related features. Most of the children with PD presented with the DDK feature as a result of their production rate being slower than the children with TD, whereas for the children with CAS most of these

children had this feature due to their inability to produce a trisyllable. For the children with CAS, 12 children displayed this feature with 11 of these children unable to produce a trisyllable (depicted by an asterisk in Table 22). For the children with PD, 11 children displayed this feature, however only three of these children were unable to produce a trisyllable, whereas the remaining eight had a slower repetition rate when compared to the children with TD.

The distribution of the features displayed in Table 22 for the CAS group is consistent, although not perfectly, with a unidimensional structure according to the Guttman scaling method. Guttman scaling, commonly used in attitudinal research, is a method of analyzing whether responses to a set of items (e.g., a positive outcome such as agree vs. disagree, or a feature is present or not present) for a set of cases conform to an underlying unidimensional psychological continuum (Price, 2016). It is assumed that individual cases will fall at different points along the continuum (e.g., children with CAS will have different levels of severity of the disorder, or people will differ in their strength of attitude towards something) and that the items themselves will vary in sensitivity to different levels of that attribute continuum as well. For example, some items will be highly sensitive to the underlying attribute and a positive response will be seen for individuals ranging from low to high levels of that attribute. Other items will be less sensitive and only positive for cases that fall at the high end of the continuum (e.g., more severe cases will show a particular feature, or only those with a stronger attitude will agree to a certain item). That being the case, then those cases that show a positive outcome to items that are responsive to the high end of the continuum should also show a positive outcome to all items responsive to the lower ends of the continuum. Therefore, a perfect outcome expected for a unidimensional continuum, according to Guttman scaling method, would be seen where the positive outcomes (e.g., crosses in the matrix) all cluster together, more or less, in the top left diagonal of the matrix without any gaps with the clustering. The length of the columns should decrease systematically from left to right.

From looking at the matrix for the CAS group, it is evident that while there are some gaps, the distribution of features is broadly consistent with the expected pattern where the CAS related features target a unidimensional attribute. For example, the children with a more severe deficit (a higher sum of CAS features) present with a

combination of less common features (e.g., prosody/stress errors) and more common features (e.g., vowel inconsistency), compared to children with a less severe CAS speech deficit, who mostly presented with more common features. Some features, for example, the ICD feature, showed a more marked exception to the expected pattern (resulting in more gaps), indicating that it may not be as closely related to this underlying construct as the other features, such as the inconsistency feature and vowel inaccuracy. On the other hand, the children with PD do not appear to fit the same unidimensional model and did not present with the less common features, which tended to only occur in children with more severe CAS.

Some of the children classified as CAS presented with only one or two features, or in one case, no CAS related features. Based on the conventional approach to classifying children with CAS where a threshold number of CAS features is required for classification, this suggests that some of the children classified as CAS many have been misdiagnosed. Given these children meet criteria for a speech sound disorder and there is no evidence of motoric involvement, it would suggest that they would be more appropriately classified as having PD.

A Pearson correlation between PCC and sum of features (reflecting severity of speech deficit) for the children with CAS revealed a significant negative correlation ($r = -.72, p = .001$) indicating that the higher the PCC score the lower the sum of features. This highlights the link between the number of CAS features and the severity of speech sounds deficit with children with more severe CAS displaying more articulatory errors. The correlation between PCC and sum of features was not significant for the children with PD ($r = -.47, p = .05$), although it was very close to significance. The children with TD did not display any of the features of CAS.

Exploratory Factor Analysis

A more sophisticated test of the unidimensional latent structure was undertaken with exploratory factor analysis. This was feasible because we had operationalized the CAS related features and in doing so, created continuous CAS related feature variables. All three groups were included in the analysis, however, prior to running the analysis we needed to examine the data to ensure that all the variables were suitable. This examination revealed multi-collinearity between the inconsistency measure and the vowel inconsistency measure (Tabachnick & Fidell, 2007). Communality was 0.799 for inconsistency feature and 0.930 for vowel inconsistency.

On this basis the vowel inconsistency feature variable was removed as redundant from any further analysis (Allen & Bennett, 2010). The remaining variables were not normally distributed revealed by Shapiro-Wilk tests of normality ($ps < .05$). However, from examination of the histograms and given the robust nature of factor analysis, these deviations were not considered problematic (Tabachnick & Fidell, 2007), except for the syllable deletion feature, which was severely skewed ($p < .001$). For this reason the syllable deletion scores were transformed using a base 10 log transformation (Tabachnick & Fidell, 2007).

Principal axis factor analysis was conducted on the following features: inconsistency percentage score, vowel inaccuracy percentage score, DDK combined feature (a binary variable reflecting either normal DDK production or inability to produce a trisyllable, or repeating tri-syllables at a slow rate), distorted prosody percentage score, ICD percentage score and log of percentage of syllable deletion errors. Principal factor analysis seeks the least number of factors that can account for a common variance among a set of variables (Tabachnick & Fidell, 2007). The correlation matrix generated by the factor analysis revealed that the feature variables were linearly correlated with one another and were therefore suitable for factor analysis, indicated by a correlation rating (Pearson's r) above .3 (Allen & Bennett, 2010). The correlation matrix is presented in Table 23. All statistically significant correlations are depicted with an asterisk (*).

The exploratory factor analysis resulted in a one-factor solution (based on an Eigenvalue greater than 1) accounting for 56.70% of the variance in scores with an Eigenvalue of 3.402. Eigenvalues tell us how much of the total variance is explained by a factor (Portney & Watkins, 2011). Consequently, a loading of .5 or greater indicates a high loading and reflects the strength of the relationship with the underlying construct and also tells us the covariance with other features included in the analysis (Allen & Bennett, 2010). As can be seen in Table 24 all features, excluding ICD, achieved a factor score greater than .5. The factor loading for ICD was less than .5, indicating that this feature does not have a strong relationship with the latent construct, or a strong inter-relationship with the other CAS related feature variables. It was therefore excluded from the further analysis in order to keep the number of variables within recommended limits given the small sample size (Allen & Bennett, 2010).

Table 22

Related Features of CAS in Children with CAS and PD (N = 35)

Part	Incon	Vowel Incon	Vowel Inacc	DDK	Pros/Stress	ICD	SyllDel	Sum
CAS								
2	X	X	X	X*	X	X	X	7
4	X	X	X	X*	X	X	X	7
5	X	X	X	X*	X	X	X	7
8	X	X	X	X*	X	X	-	6
13	X	X	X	X*	X	-	X	6
14	X	X	X	X*	X	X	-	6
6	X	X	X	X*	X	X	-	6
11	X	X	X	X*	X	-	X	6
7	X	X	X	X*	X	-	X	6
9	X	X	X	X*	X	-	-	5
3	X	X	X	-	-	X	-	4
15	X	X	X	X	-	-	-	4
17	X	X	X	-	-	X	-	4
18	X	X	X	-	-	-	-	3
12	X	-	-	X*	-	-	-	2
16	-	X	-	-	-	-	-	1
1	-	-	-	-	-	-	-	0
PD								
12	X	X	X	X	-	-	-	4
17	X	-	X	X	-	-	-	3
4	-	X	X	-	-	-	-	2
2	-	X	X	-	-	-	-	2
6	-	-	X	X	-	-	-	2
9	-	X	-	X	-	-	-	2
8	-	-	-	-	-	-	-	1
18	-	-	-	X*	-	-	-	1
3	-	-	X	-	-	-	-	1
14	-	-	-	X	-	-	-	1
1	-	-	-	X*	-	-	-	1
5	-	-	-	X	-	-	-	1
15	-	-	-	X	-	-	-	1
16	-	-	-	X*	-	-	-	1
7	-	-	-	-	-	-	-	0
10	-	-	-	-	-	-	-	0
11	-	-	-	-	-	-	-	0
13	-	-	-	-	-	-	-	0

Note. Part = Participant; Incon = inconsistency; VowIncon = vowel inconsistency; Vowel Inacc = vowel inaccuracy; DDK = diadochokinetic syllable repetition task; Pros/Stress = Prosody/Stress Errors; ICD = Initial Consonant Deletion (representing deviant errors); SyllDel = Syllable Deletions (representing simplification/omission errors). * Indicates feature is present due to inability to produce trisyllable (as opposed to repetition rate being significantly slower than the children with TD).

The final exploratory analysis, without the ICD feature, yielded a one-factor solution accounting for 63.37% of the variance in scores with an eigenvalue of 3.169. As can be seen from Table 25 all the features had a loading greater than .5 indicating a strong inter-relationship. The inconsistency measure had a very high load of .89 indicating that this feature is very strongly associated with the singular underlying construct. Vowel inaccuracy and syllable deletion errors also had a high loading indicating a strong association with the singular latent construct. Although DDK combined feature had sufficient loading to be included in the one factor solution, this feature was not as strongly associated with the one factor solution or as closely connected to the other features. This may be attributed to this feature reflecting both, inability to produce a trisyllable and repetition rate being significantly slower than the children with TD. Regardless, the resultant factor scores from the exploratory factor analysis can justifiably be interpreted as the degree to which each child exhibits these combined set of CAS related features. The resultant factor score for each participant can be interpreted as a type of CAS trait score reflecting the degree of severity of CAS.

Table 23

Correlation Matrix Between Feature Variables of CAS (N = 53).

	Incon	VowInacc	DDK	ProsDis	ICD	LogSyllDel
Incon	-	-	-	-	-	-
VowInacc	.702*	-	-	-	-	-
DDK	.551*	.402*	-	-	-	-
ProsDis	.588*	.543*	.424*	-	-	-
ICD	.449*	.362*	.266	.423*	-	-
LogSyllDel	.645*	.587*	.420*	.512*	.154	-

Note. Incon = Feature using the inconsistency subtest of the DEAP (Dodd et al., 2002); Vowel Inacc = vowel inaccuracy percent score; DDK = diadochokinetic syllable repetition task binary feature; ProsDis = prosody distorted percent score; ICD = Initial Consonant Deletion percent score (representing deviant errors); LogSyllDel = log of syllable deletion errors percent score (representing simplification/omission errors).

Statistical analysis using a one-way independent groups ANOVA revealed that the groups were significantly different in the mean factor score, $F(2, 50) = 50.15$, $p < .001$. The children with CAS had a significantly higher factor score ($M = 1.06$, $SD =$

.89) than the children with PD ($M = -0.22$, $SD = .33$), $t(33) = 5.7$, $p < .001$ and the children with TD ($M = -0.78$, $SD = .22$), $t(33) = 8.5$, $p < .001$. The children with PD had a significantly higher factor score than the children with TD, $t(34) = 6.88$, $p < .001$.

Table 24

Factor Loading for Features of CAS onto a One-Factor Solution (N = 53).

Item	Factor 1
Inconsistency	.905
Vowel Inaccuracy	.776
DDK Combined Feature	.582
Prosody Distorted	.714
ICD	.455
Log Syllable Deletion	.694

Note. Extraction method: Principal Axis Factoring.

Final Classification Protocol

The features of CAS revealed a one-factor solution resulting in a factor score for each participant, which can reliably be interpreted as reflecting the severity of the speech deficit associated with these features. Based on DFA, which used both the number of CAS features and the factor score (a measure of CAS trait) as predictor variables, the cases that were classified back into their original diagnostic category were taken as clear cases for each group (i.e., 14 clear cases of CAS, 12 clear cases of PD and 18 TD children), but with the proviso that children in the CAS group also needed a minimum number of 3 CAS features. The children who were put in an alternative group (suggesting misclassification), were then assessed according to the following to determine whether they should stay in the new group based on DFA. For children in the original CAS group who were classified as PD, if their number of CAS features was less than 3 and their factor score was more than 1 SD (.71) below the mean of the CAS group (1.33), then this was taken as evidence against motoric involvement consistent with CAS and they remained in the PD group. For all three candidates who changed from CAS to PD the number of CAS related features was less than 3, consequently, they were reassigned to the PD group. The children with

PD needed to meet the same criteria in order to be re-classified as having CAS, that is, they needed to have more than 3 CAS related features and have a factor score above 1 SD below the mean of the CAS group.

Table 25

Factor Loading for Features of CAS Loading Onto a One-Factor Solution Excluding ICD (N = 53).

Variables	Factor 1
Inconsistency	.891
Vowel Inaccuracy	.774
Log Syllable Deletion	.737
Prosody Distorted	.689
DDK Combined Feature	.582

Note. Extraction method: Principal Axis Factoring.

The DFA resulted in 14 clear cases of CAS (see Table 26). The three children, originally in the CAS group (participants CAS1, CAS12, and CAS16) displayed zero, two and one CAS related feature, respectively, and factor scores more than one SD below the mean of the CAS group. These children were therefore reassigned to the PD group given their profile was more similar to this group. This is consistent with commonly used diagnostic criteria for PD in research, where children meet criteria for a speech sound disorder (e.g., below the 15th percentile on a test of articulation ability) but do not meet any further criteria for motor involvement (e.g., CAS) (Munson et al., 2005). Two children with PD were identified from the DFA as possible CAS candidates (participants PD12 and PD17). However, these participants remained in the PD group since they did not meet both of those criteria. The DFA analysis also identified 4 children with PD (participants PD7, PD10, PD11 and PD13) as possible candidates for the TD group, however, since these children displayed features consistent with SSD using standardized testing they remained in the PD group.

Table 26

Discriminate Function Analysis Resulting in Reassigned Group Membership (N = 53)

	Predicted Group Membership			Total
	CAS	PD	TD	
CAS	14	3	0	17
PD	2	12	4	18
TD	0	0	18	18

Note. 83.0% of original grouped cases correctly classified.

Sensitivity and Specificity

A sensitivity and specificity analysis (Portney & Watkins, 2011) was undertaken to determine the effectiveness of the individual CAS related features used in the classification protocol. Sensitivity refers to the diagnostic tool's ability to reliably identify the presence of a CAS related feature when it is in fact present (i.e., sensitivity = true positive frequency as a proportion of the sum of the true positive and false negative frequency) and specificity refers to how well the tool can capture a negative response when the CAS related feature is not present (i.e., specificity = true negative frequency as a proportion of the sum of false positive and true negative frequencies) (Portney & Watkins, 2011). The sensitivity and specificity of these features can be seen in Table 27. This analysis included only the children with CAS ($n = 14$) and PD ($n = 21$).

In the current study, the DDK combined feature failed to differentiate between children with CAS and children with PD, as there was almost an equal number of participants in both groups that displayed this feature. However, as previously mentioned seven of the ten children with PD who presented with this feature had it due to a slower rate than the TD children, it was therefore of interest to determine if ability alone was a more discerning feature to differentiate between children with CAS and PD. As can be seen from the analysis in Table 19 DDK ability compared to DDK combined feature (i.e., ability and rate) has a lower sensitivity (71 compared 79) but much higher specificity (81 compared to 43), indicating that the combined DDK feature is somewhat better for identifying the presence of the CAS feature, but DDK ability is better at identifying when the CAS feature is not present. In the

current sample, ICD occurred in children with CAS and not in children with PD, consequently it had a high specificity but low sensitivity.

Table 27

Sensitivity and Specificity Values for Operationalized Features of CAS (N= 35)

Measure	Sensitivity	Specificity
Specific CAS Features		
Inconsistency >40% *	100	86
Vowel Inconsistency *	100	76
Vowel Inaccuracy *	100	71
Prosody *	79	100
DDK Combined *	79	43
DDK Ability	71	81
Syllable Deletion Errors*	43	100
Nonspecific CAS Features		
PCC Feature severe <55%	86	100
ICD	57	100
GFTA <5%	64	81

Note. * Denotes features used for factor analysis. DDK = diadochokinetic task; PCC = percentage consonants correct; ICD = initial consonant deletion; GFTA = Goldman Fristoe Test of Articulation

Additional measures used for classification of children with CAS (see Chapter 3) were included in this analysis. PCC with a severity rating of less than 55% consistent with a severe speech deficit demonstrated the highest sensitivity of all features analysed (i.e., 86%). However, it is a feature of SSD in general given children with severe PD can have a very low PCC. Likewise children with severe PD can have a very low percentile ranking on the GFTA (i.e., less than 5%), consequently this feature has poor sensitivity, similar to ICD, for differentiating between children with CAS and PD.

Discussion

An ongoing problem, highlighted in the CAS literature, is the concern regarding diagnostic uncertainty (ASHA, 2007), consequently, the purpose of this

study was to develop a robust and transparent classification protocol for CAS using operationally defined CAS related features. Prior to doing this, we identified the most prevalent features used in the systematic review in Chapter 3. We then operationalized these features and in doing so created continuous variables. These continuous variables could then be used for exploratory factor analysis to test the hypothesis that the underlying speech related deficits in CAS, as quantified by the CAS related feature variables, was a unidimensional disorder with a singular underlying construct, consistent with a speech motor deficit. In addition, this method of analysis allowed us to investigate the communality between the CAS related feature variables and enabled us to investigate the validity of the CAS related features and their capacity to differentiate between children with CAS and PD. We also were able to develop a protocol to test the accuracy of group classification using the associated factor scores and the sum of CAS related features displayed by each participant.

Prior to undertaking the exploratory analysis we organized the CAS related features in order of prevalence for the CAS participants, consistent with Guttman scaling method. The ideal Guttman scale, which aims to measure individual differences along a unidimensional construct, is a set of items (in this case, CAS related features) that are ranked in a specific order of prevalence (Trochim & Donnelly, 2007), consistent with a unidimensional latent construct. The Guttman table resulted in a relatively close fit to the ideal Guttman scale, consistent with a unidimensional model, for the children in the CAS group. Overall, the children with CAS displayed more features than the children with PD, although some of the CAS related features were also observed in the PD group. In addition, there was a trend in the data for the children with CAS who presented with the less common features (e.g. syllable deletion and ICD) to also present with more common features (e.g., inconsistency), and more features in total. This suggests that some of the features, namely those are more prevalent, such as inconsistency, occurred in most of the children with CAS, whereas, the CAS related features that were less common, such as prosody and syllable deletion, only occurred in the more severe cases of CAS, not the case for the children with PD. Interestingly, the correlation between PCC and the sum of features was significant for the children with CAS indicating that there is a strong relationship between a low PCC score (i.e., less than 55%) and severity of

motor speech deficit (indicated by sum of features). For the children with PD, the correlation between the PCC and the sum of CAS related features just missed out on reaching significance suggesting this relationship was not as strong for the children with PD compared to the children with CAS. This is possibly related to a restriction in range of severity of deficit for the PD group and warrants further investigation. For the children with CAS the strong correlation between these measures indicates there is degree of interdependence between PCC and severity of motor speech deficit, represented by the number of CAS related features for these children. In keeping with the general view that the underlying deficit in CAS is a motor speech deficit, the link between sum of CAS related features and PCC supports the proposal that higher level phonological /linguistic deficits (such as low PCC) observed in children with CAS occur as a result of the interdependency between the different levels of processing during development. This proposition is consistent with theories on the dynamic nature of speech and language development in relation to children with CAS (Maassen, 2002) and will be further explored in the general discussion in Chapter 6.

Some features, such as ICD, did not fit with the Guttman scaling expectation, which were later supported by the results of the exploratory factor analysis demonstrated by the ICD not loading onto the single factor solution. This suggests that the ICD feature is not closely coupled with the underlying singular construct, or with the other CAS related features included in the analysis. It is therefore questionable whether this feature is part of the same underlying deficit in speech motor control or if it is associated with a higher-level phonological deficit. If this is the case, then why do children with CAS present with deviant errors, such as ICD. One plausible explanation could be that the lower level speech motor deficit constrains development of higher-level phonological representations, consistent with the dynamic nature of developmental disorders. The general discussion elaborates on this proposal.

The exploratory factor analysis provided a statistically robust method of investigating the underlying deficit of CAS. Consistent with our hypothesis, the selected CAS related features were consistent with CAS being a unidimensional disorder indicating the independence of a speech motor deficit. Furthermore, in the context of the DIVA model (Guenther & Perkell, 2004) and its adaptation by

Terband and Maassen (2010) to children with CAS this underlying speech motor deficit, the one factor solution found provides empirical evidence to support the proposal that the multiple deficits in CAS can arise as a result of a singular underlying construct in speech motor control.

We originally looked at six CAS related features but given the ICD feature resulted in a low factor loading (less than .5), we reran the analysis without this feature. The final analysis resulted in a one-factor solution for the five remaining CAS related features indicating that these features were closely coupled with the underlying construct and also one another. The inconsistency feature had the highest loading and DDK combined feature had the lowest loading. It was not surprising that the DDK combined feature had the lowest factor loading and although sufficient to be included in the one factor solution, the low loading suggests that this feature was not as tightly coupled with the other features, or the underlying construct. This feature was present in almost equal numbers of children with PD in our original sample (based on rate as opposed to ability) and therefore DDK ability alone may have been a more appropriate and CAS specific feature. All other variables had a loading of .7 or greater, suggesting a strong relationship between the feature variables and the underlying construct.

Sensitivity and specificity values calculated for each of the CAS related features did not achieve the ideal “gold standard” as suggested by Shriberg and colleagues (2012), which is 90% for both sensitivity and specificity. The features that showed the greatest potential for identifying children with CAS were the inconsistency measure with high sensitivity (100%) and specificity (86%). However, the goal was not to identify and implement one specific feature to classify children with CAS, but to operationalize and define a number of features that were specific to CAS and could be used collectively to classify children as such. The other features analysed, not specific to CAS, did not achieve sensitivity and specificity values sufficient to justify their inclusion as CAS specific features. The PCC feature, although it achieved high sensitivity and 100% specificity is not unique to children with CAS since this feature is present in children with PD with a severe phonological deficit.

The DFA used to reclassify children with CAS and validate their CAS status resulted in three of the children, previously diagnosed as CAS (based on expertise

clinical judgement), being reassigned to the PD group. These three children failed to reach classification criteria, based on a combination of factor score (from herein referred to as the CAS trait score) and sum of features present. This newly developed method of classification was shown to be a reliable and transparent method to classify participants with CAS. When research participants are selected based solely on clinical judgment it is difficult to determine which diagnostic criteria were considered for this evaluation. Furthermore, as previously highlighted, Davis et al. (1998) and Forrest (2003) have revealed significant disagreement between clinicians regarding the criteria for diagnosing CAS potentially resulting in an over inflation of participants classified as such and highlighting the need for a more rigorous diagnostic protocol.

In summary, this study has generated a number of novel findings. Exploratory factor analysis has not been used as a method of exploring the underlying construct associated with CAS. It has been used in SSD (Lewis et al., 2006) to classify children with SSD by investigating the factor structure of a number of early speech and language measures (e.g., GFTA, PCC) and reading and spelling measures (e.g., spelling, word attack) to determine the underlying deficit that these measures represent. It has also been used in CAS, but as a method of data reduction (E. Teverovsky, Bickel, & Feldman, 2009a). Teverovsky et al. (2009a) used factor analysis to detect structural relations among classification codes used to describe and categorize functional problems of children with CAS. Their primary goal being to describe the functional abilities in children with CAS using codes provided by the International Classification of Functioning Disability and Health (ICF) and included codes to describe daily activities, body functions, social participation and environmental factors that may influence these abilities. The current study used exploratory factor analysis to explore the underlying dimensionality of CAS related features that have been conventionally used for their classification. This has not been done in the CAS literature, to date, and is therefore a unique and novel contribution, not only with regard to its application, but in relation to the findings. The fact that the features loaded onto a single factor solution is consistent with the proposal that CAS is a unidimensional disorder. The factor score, along with the number of CAS related features present, can reliably be interpreted as an indicator of severity of speech motor deficit and can therefore be used for diagnostic purposes.

Consequently, the factor score in conjunction with sum of features was validated as an accurate method of classification for the participants in Study 2, thereby restricting the CAS group to clear cases of CAS.

Limitations and Future Directions

The sample size for this study is a key limitation. A larger sample size is a better representation of the population being investigated, consequently the findings would be more reliable if the cohort was larger. One aspect overlooked in this study was the developmental shift in the presentation of features, particularly in children with CAS, which has been highlighted in the literature (Pennington & Bishop, 2009). The age range of the participants varied from 4 to 8 years of age and the developmental changes that can occur in typical development during this period is extensive. Therefore, a larger sample size would have permitted groups to be split by age to determine if age range had an impact on findings. Furthermore, the children with CAS and PD were recruited from community clinics where the children are likely candidates for having a language deficit. However, all the participants, excluding one child with CAS (who was omitted from the study for this reason) had vocabulary measures within the normal limits, indicating that their expressive and receptive language, at least in terms of their lexical-semantic knowledge, was within normal limits.

Although the children in the PD group showed fewer CAS features than the children with CAS, some displayed CAS related features. The correlation between PCC was not significant for the PD group, however, given how close this was to reaching significance ($p = .05$), it would be of interest to determine if a significant relationship between sum of features and PCC would emerge for the children with more severe PD compared to children with a milder speech deficit. It would be useful to explore the prevalence of these CAS related features in children with mild PD and severe PD. It is plausible that there would be little overlap in the features shared between children with mild PD and children meeting criteria for CAS, whereas there may be a greater overlap between severe PD and CAS. The findings from the factor analysis are consistent with CAS having a singular underlying deficit, indicating a discrete and distinct speech motor deficit in children with CAS. The children with PD did not fit this profile, demonstrated by this group not fitting with the Guttman scaling method and the lack of correlation between PCC and sum of features for the

children with PD. Future research could explore underlying dimensionality in PD using exploratory factor analysis, which could provide a more equitable platform to compare children with CAS and other SSD with PD. It could also help determine if these two disorders CAS and PD are distinctly different or if they in fact have a similar causal origin but differ in terms of severity.

Conclusion

Children with CAS have historically been classified based on the presence of features, and classified as such based on expert opinion. However, as revealed in the systematic review a number of ambiguities and inconsistencies emerged with regard to the classification protocols currently used in the research literature. This applies not only to the ambiguous nature of how features are identified, but also in relation to the number of features considered sufficient to warrant classification as CAS. The protocol developed in the current study, using the factor scores derived from a set of operationalized CAS related feature variables and the number of CAS related features, provides an alternative and transparent method of classification. It has the advantage over other methods of classification by identifying the most relevant features of CAS used in the research literature and which presumably are an accurate representation of the CAS population. Not only have the features been operationalized using a readily available tool (i.e., DEAP), permitting a replicable and reliable approach to validate the presence of features, but the factor analysis, resulting in a one factor solution and corresponding factor score, indicating severity of speech motor deficit, further enhances this method of classification.

Over diagnosis of CAS has been identified as an ongoing problem and a widely discussed professional issue (ASHA, 2007). ASHA contends that over diagnosis is a result of the lack of information available to speech language pathologists regarding the key diagnostic characteristics of this disorder (Davis et al., 1998; Forrest, 2003). Furthermore, many of the features used for classification purposes also occur in children with SSD (ASHA, 2007). It was therefore paramount to establish a more robust and transparent protocol for Study 2 to ensure that the children were classified as accurately as possible, which was interpreted in the present study as targeting children who are representative of the key characteristics that previous research has used to classify children with CAS.

Chapter 5

Exploring Predictors of Phonological Competence in Children with CAS and PD

Introduction

The evidence to date suggesting a motor disorder underlying the deficits observed in CAS and a phonological deficit underlying the deficits observed in PD is limited, especially given the significant overlap of deficits observed in these children. Children with CAS have been shown to demonstrate deficits at the phonological encoding level (Marion et al., 1993; Marquardt et al., 2004; Marquardt et al., 2002) and likewise, children with PD have demonstrated deficits at the level of speech motor control (Dodd et al., 2005). However, determining the source of the deficit is difficult in developmental disorders due to the interaction of the different levels of the speech and language system during development (Maassen, 2002). The primary aim of this study was to determine the degree of shared deficits within the phonological and speech motor systems of children with CAS and PD and then to compare both groups in terms of the relationships between levels of speech motor ability and phonological competence, with the principal goal to reveal different patterns of relationship that could suggest difference in the source or causal origin of these shared deficits.

This proposition was considered in the context of Meehl's (1992) theory of taxonomy. Meehl's (1992) theory proposes that taxons that have different aetiologies can have similar pathologies resulting in overlapping symptoms, consistent with the shared speech impairments observed in children with CAS and PD. The proposal that CAS is a unidimensional disorder (as demonstrated by the single factor solution from the factor analysis), distinct from PD, is consistent with this theory. This thesis aimed to test this notion of independent taxons to CAS and PD, despite CAS and PD presenting with similar deficits at multiple levels of the speech processing system. We used regression analysis, extending the paradigm used by Munson et al. (2005b) to compare children with CAS and PD in terms of predictors of phonological competence as a strategy to test for differences in the nature of the constraints on the developing system of children with CAS and PD. If the pattern of relationship between predictors differed for children with CAS and PD then, other factors being equal, this would support the proposal that CAS and PD have had different developmental trajectories and consistent with the proposal that CAS and PD have different etiologies, and providing empirical evidence that CAS and PD are discrete disorders. The main hypothesis under investigation is that

speech motor measures will uniquely predict outcome measures of phonological competence, after controlling for covariates of phonological competence such as receptive and/or expressive vocabulary in children with CAS but not (or to a less extent) in children with PD and TD.

In particular, it is assumed that children with CAS begin speech and language development with a core deficit at the motor level of speech production. Because these levels of processing mutually influence each other during development this core motor deficit can constrain the emergence of phonological representations of speech, and this was predicted to give rise to a tight coupling between measures of speech motor ability and phonological competence, independent of how higher level vocabulary knowledge might influence phonological development. Therefore, we predicted for children with CAS that measures of speech motor ability will uniquely predict phonological competence after controlling for vocabulary knowledge.

In contrast, children with PD are assumed to have a deficit originating at a higher linguistic-phonological level. Although this deficit is not well understood, it could be associated with an intrinsic problem in phonological processing and/or difficulty learning the phonological rules of the child's language (Gierut 1998; 1999). It may also interact directly with the development of vocabulary knowledge. By hypothesis, however, the speech motor control system has normal capacity to learn and will get off to a good start in development through the prelinguistic babbling stage, which is a stage of early speech motor development that is believed not to be contingent on establishing phonological representations for producing speech (Levelt et al., 1999). During the prelinguistic stage the sensorimotor mappings for basic articulatory routines (e.g., canonical and advanced canonical babbling) that will be the basis of early intentional speech production are formed (MacNeilage, 1997; 1998; 2000). There is, therefore, greater potential for early independence in the development of the phonological system and the speech motor control system in children with PD compared to children with CAS. While a phonological deficit can also potentially impact speech motor control because of dynamic interactions, we predicted less coupling between measures of speech motor ability and phonological competence in the case PD provided their deficit originates at a higher linguistic level. For example, children with PD are characterised by producing a variety of speech errors such as omissions and substitutions because of

poorly formed phonological representations. But, in view of the above argument, they should have relatively normal capacity to articulate the speech sounds for the phonological representations that are correctly formed. This is also reflected in the fact that children with PD by definition do not meet clinical criteria for having an impaired speech motor system (although some deficits in processing at the speech motor level are apparent, as highlighted in Chapter 1). While speech motor ability may be associated with phonological competence in children with PD, we predicted that speech motor measures were unlikely to uniquely predict phonological competence after controlling for vocabulary knowledge (this is based on the assumption that the variance linking phonological competence with the vocabulary and speech motor levels of processing is more likely to be shared rather than unique).

The approach taken in this thesis was to evaluate different aspects or levels of phonological competence and to test the main hypothesis in each of these domains. In addition, more than one measure of the speech motor system was used. This strategy of using multiple measures and multiple tests of the research hypothesis was therefore adopted in this thesis to cast a fairly wide net to capture the hypothesized relationships, an approach was deemed appropriate given the preliminary and exploratory nature of this research.

Phonological Competence

We evaluated phonological competence using the tasks implemented with typically developing children in Study 1 (chapter 2). The NWR task and speech discrimination task used in Study 2 were unchanged from the tasks administered in Study 1. However, the picture-naming task was modified to include two SOAs and the auditory primes were changed from onset primes to rime primes. SOA refers to the transition between the picture presentation and the presentation of the auditory prime and the sensitivity of lexical generation to phonological interference and facilitation effects have been shown to differ with age (Brooks & Mac Whinney, 2000; Jerger et al., 2002). The facilitation effect of a related auditory prime has been shown to be greater when the prime lags the picture presentation (Jerger et al., 2002). If the prime is presented too soon, before the presentation of the picture, then it is hypothesized that the phonological activation will have decayed prior to the onset of phonological encoding (Schriefers et al., 1990). We therefore included two SOA's, both presented after picture presentation at 50ms and 150ms, to maximize priming

effects for the age of the participants in the current study, based on the assumption that children with CAS and PD are more likely to have slower processing skills. The auditory onset primes were replaced with rimes based on the premise that younger children are more likely to be influenced by rimes than onsets, in the context of picture-naming (Brooks & MacWhinney, 2000).

Speech Motor Ability

Measures of speech motor ability included the GFTA raw score, a CAS trait score (i.e., factor score obtained from the factor analysis) and an additional measure of speech motor control obtained from a delayed picture-naming task, developed for this study.

The delayed picture-naming task reaction time (DPNrt) is a picture-naming task with a delayed response. In the present version of the DPNrt, when the picture is presented the child must wait for a signal to appear on the computer screen prior to initiating the response and naming the picture. The general assumption and rationale for using DPNrt is that by delaying the onset of the verbal response this allows sufficient time for the pre-execution processes, such as picture identification, lexical retrieval and phonological encoding, to be completed (Kawamoto, Qiang Liu, & Sanchez, 2008). According to theories of speech motor control and execution, in delayed picture-naming a fully specified motor program is delivered and held in the motor program buffer (consistent with the articulatory buffer in the WEAVER model) and remains there until the signal to respond is detected (Kawamoto et al., 2008). Consequently, the time between the presentation of the signal to respond and speech output will be sensitive to response execution processes and, in particular, the time it takes to retrieve, unpack and initiate the execution of the speech motor plan (Sternberg et al., 1988). Word frequency was manipulated to include high and low frequency target words to check on this assumption. This was based on the perspective that high frequency words have a faster reaction time compared to low frequency words (Newman & German, 2002), but for delayed picture-naming, where the response is planned in advance, word frequency should not have an effect on reaction time, given there is sufficient time for pre-execution processes to be completed, regardless of the target word frequency (Lunganaro & Xavier Alario, 2006). The motivation for this task comes from the SVrt in Study 1, which was shown to be a valid measure of speech execution but was limited in that it only had a

small number of items and a small number of trials per item contributing to the overall mean, therefore potentially compromising the findings. The SVRT also demonstrated a frequency effect, which might be related to a lexical effect, or a stimulus design confound, thus undermining the validity of the task. The DPNrt is similar to the SVrt except that the task involves a greater sample of verbal responses. Study 1 provided evidence that the SVrt was related to the development of speech discrimination skills in TD children, therefore suggesting that a similar measure of speech motor ability, such as the DPNrt is a good candidate to examine different degrees of constraint between speech motor and phonological levels of processing in children with CAS and PD.

Hypotheses

There were a number of hypotheses for the experimental tasks used to assess phonological competence and speech motor ability in children with CAS, PD and TD. For the nonword repetition task we hypothesized that the children with CAS and PD would be less accurate than the children with TD in overall nonword repetition accuracy. This hypothesis was based on the deficits expected in children with CAS and PD in speech motor ability and phonological knowledge respectively, Consistent with previous findings, we hypothesized that the nonwords that contained high frequency sequences would be repeated more accurately than the nonwords that contained the low frequency sequences (Munson, Edwards, et al., 2005b). The hypothesis relating to the phonotactic frequency effect (i.e., difference in repetition accuracy between the high and low frequency sequences) in children with CAS and PD was speculative. We hypothesized that there would be an interaction between group and the phonotactic frequency effect indicating that the phonotactic frequency effect would differ between groups, but the direction or magnitude of this interaction was difficult to predict. Either the children with CAS and PD would perform like younger children with TD and demonstrate a larger frequency effect indicative of more holistic phonological representations (Edwards et al., 2004; Metsala, 1999; Munson, 2001a; Storkel, 2001, 2002) or they would perform similar to the children with TD with no observed difference in phonotactic frequency effect consistent with Munson et al.'s findings (2005b).

For the picture-naming task with auditory primes we hypothesized that the children with CAS and PD would be slower overall at picture-naming than the

children with TD. This hypothesis was based on the assumption that children with CAS and PD have weaker output phonological encoding, although the difference in reaction time may be due to different stages involved in picture-naming for the children with CAS and PD (refer to the WEAVER model in Chapter 1). Consistent with previous findings we hypothesized that the phonologically related priming condition would reveal a faster reaction time compared to the unrelated priming condition, demonstrating a phonological facilitation effect (Brooks & Mac Whinney, 2000). We also hypothesized that there would be a group by priming interaction based on the premise that children with CAS and PD have less efficient output phonological representations, consistent with previous research with children with dyslexia, who are also assumed to have a phonological processing deficit (Truman & Hennessey, 2006). However, the extent that the children with CAS would differ from the children with PD, if they differed at all, was speculative, given this paradigm has not been implemented in children with CAS and PD simultaneously. Either the children with CAS and PD will have a larger facilitation effect, consistent with previous research in children with dyslexia (Truman & Hennessey, 2006), or they will not differ at all, consistent with Munson and Krause (2017). Munson and Krause (2017) found that children with SSD did not differ in the magnitude of phonological facilitation or inhibition effects when compared to children with TD.

For the speech discrimination task we hypothesized that the children with CAS and PD would reveal poorer speech discrimination sensitivity (i.e., d' -prime) than the children with TD. This is in keeping with previous research that has shown that children with PD require more acoustic information than children with TD to accurately discriminate between words when the acoustic signal is degraded (Edwards et al., 2002). Children with CAS have also been shown to have a speech perception deficit, albeit for a different reason than children with PD (Nijland, 2009). Children with CAS have been shown to have auditory processing deficits (Groenen et al., 1996; Nijland, 2009) and poorer auditory discrimination skills of consonants (Groenen et al., 1996) and vowels (Maassen et al., 2003). Consequently, we hypothesized that the children with CAS would need more acoustic information to discriminate between words than the children with TD and show diminished sensitivity to acoustically degraded stimuli. We also proposed that the children with

CAS and PD might differ in d-prime depending on their severity of deficit at the level of input phonological representations.

For the delayed picture-naming task we hypothesized that the children with CAS would have a slower reaction time than the children with PD and TD indicative of less efficient speech motor execution system, consistent with a diagnosis of CAS. We also hypothesized that the children with PD would demonstrate a slower reaction time than the children with TD, since there is children in the PD sample that also have deficient speech motor control.

The main research hypothesis was that speech motor measures, after controlling for vocabulary knowledge, would uniquely predict measures of phonological competence to a greater extent in children with CAS but not in children with PD and TD. This hypothesis was based on the premise that CAS is associated with an underlying deficit in speech motor control and this speech motor deficit, through dynamic interactions over the course of development, has an effect of constraining the emergence of the linguistic-phonological system in children with CAS. In PD, however, the speech motor system has potential to develop normal capacity to articulate the speech sounds that are within the child's phonemic repertoire independently of the higher level phonological deficit. Furthermore, if there are limitations that the phonological system places on speech motor development, the variance between those two systems will be shared with vocabulary knowledge, which is also potentially constrained by the same underlying phonological processing deficit.

Three measures capturing different aspects of speech motor competence were used as the primary predictors (GFTA, DPNrt and CAS trait score). We hypothesized that articulation ability, as measured by the Goldman Fristoe Test of Articulation (Goldman & Fristoe, 2000), would predict nonword repetition accuracy, speech discrimination ability and the phonotactic frequency effect in children with CAS and PD. Edwards et al. (2004) showed that the GFTA did not predict the phonotactic frequency effect for the children with TD and, Munson et al. (2005) also showed that the GFTA did not predict the phonotactic frequency effect in children with PD and TD. However, when the children with PD were analysed in relation to severity of deficit (i.e., mild versus severe) the GFTA did predict the phonotactic frequency effect for the children with a severe speech deficit (Munson et al., 2005), indicating

that articulatory ability when severely compromised is a contributing factor to whether or not articulatory ability predicts the phonotactic frequency effect. Although speculative, we hypothesized that the GFTA would predict picture-naming reaction time and the phonological facilitation effect for the children with CAS and PD. In relation to group differences we predicted that the GFTA would uniquely predict measures of phonological competence in the children with CAS but not (or to a lesser extent) in children with PD and TD.

In summary, our overall goal was to implement tasks that more selectively targeted phonological and speech motor ability in children with CAS, PD and TD with the overall goal to explore predictors of phonological competence using hierarchical mixed moderator regression analysis. By examining the levels of linguistic (phonological) representation and lower level speech motor control, at the same time in both groups, addresses the question: do children with PD and CAS share deficits at both levels.

Method

Participants

The children were assigned to the CAS and PD groups based on discriminate function analysis undertaken in Chapter 4. Final classification of participants resulted in 14 children with CAS, 21 children with PD and 18 children with TD. For further details see the section on Final Classification Protocol in Chapter 4. The youngest participant was 4 years 2 months and the oldest 7 years 11 months, the age range was 4.2 to 7.9 for the children with CAS, 4.5 to 8.4 for the children with PD, and 4.2 to 7.10 for the children with TD. Refer to Table 28.

Statistical analysis using ANOVA revealed that groups differed significantly on the EVT standard score, $F(2,50) = 16.77, p < .001$, partial $\eta^2 = .40$. The children with CAS had a significantly higher EVT standard score than the children with PD, $F(1, 33) = 4.42, p = .043$, partial $\eta^2 = .12$; and a significantly lower standard score from the children with TD, $F(1, 30) = 12.14, p = .002$, partial $\eta^2 = .28$ and the children with PD had a significantly lower EVT standard score than the children with TD; $F(1, 37) = 29.88, p < .001$, partial $\eta^2 = .45$.

The groups also differed significantly on the PPVT standard score, $F(2,50) = 13.72, p < .001$, partial $\eta^2 = .35$. The children with CAS did not differ from the

children with PD, $F(1, 33) = 1.19, p = .29$, partial $\eta^2 = .03$, but had a significantly lower standard score than the children with TD; $F(1, 30) = 13.72, p = .001$, partial $\eta^2 = .31$. The children with PD also had a significantly lower standard score than the children with TD; $F(1, 37) = 23.10, p < .001$, partial $\eta^2 = .38$.

Table 28

Demographic Data and Test Scores for Children with CAS, PD and TD with Standard Deviations (SD) (N = 53)

		CAS		PD		TD	
		(n = 14)		(n = 21)		(n = 18)	
		M	SD	M	SD	M	SD
Age in months		65	13	77	13	70	11
EVT	Standard Score	96	6	91	9	105	7
	Raw Score	51	9	55	11	63	12
PPVT	Standard Score	97	7	94	9	108	9
	Raw Score	85	19	96	22	107	21
GFTA	Percentile Rank	5	5	13	10	51	23
	Standard Score	63	17	82	15	104	7
	Raw Score	39	10	19	9	5	5

Note. GFTA = Goldman-Fristoe Test of Articulation (Goldman & Fristoe, 1986); EVT = Expressive Vocabulary Test (Williams, 1997); PPVT-III = Peabody Picture Vocabulary Test – III (Dunn & Dunn, 1997).

The groups differed significantly on chronological age, $F(2,50) = 3.73, p = .031$, partial $\eta^2 = .14$. The children with CAS were significantly younger than the children with PD, $F(1, 33) = 6.57, p = .015$, partial $\eta^2 = .17$. This was an unexpected outcome arising from the final classification process used to assign children to the CAS and PD groups, even though the original samples were age-matched at the group level. The CAS children did not differ significantly from the children with TD for chronological age, $F(1,30) = 1.45, p = .239$, partial $\eta^2 = .05$, and the children with PD did not differ significantly from the children with TD, $F(1, 37) = 2.64, p = .113$, partial $\eta^2 = .07$. Age was therefore included as a covariate in the analysis to remove any confounding factors associated with the children with CAS and PD being significantly different in relation to age.

The groups differed significantly on GFTA raw score, $F(2,50) = 70.54$, $p < .001$, partial $\eta^2 = .74$. The children with CAS had a significantly lower GFTA raw score than the children with PD, $F(1, 33) = 44.76$, $p < .001$, partial $\eta^2 = .58$, and the children with TD, $F(1, 30) = 157.60$, $p < .001$, partial $\eta^2 = .84$, and the children with PD did differ significantly from the children with TD, $F(1, 37) = 29.6$, $p < .001$, partial $\eta^2 = .44$.

Stimulus Materials

The stimulus materials for the nonword repetition task and the speech discrimination task were the same stimuli used in Study 1. Refer to Stimulus Materials in Chapter 2 for further details.

Picture-naming task. Stimuli for this task consisted of 15 digitized black and white pictures of everyday objects. We reduced the number of items from 18 items (originally used in Study 1) since we were including an additional SOA and therefore doubling the length of the task. These items were taken from the same items used in Study 1. All test items had a low frequency rating of less than nine occurrences per million (Kučera & Francis, 1967). Furthermore, the phonological primes used in Study 1 were changed from onsets to rimes/offsets. Phonological primes consisted of two priming nonsense syllables formatted for each picture name target word, one related and one unrelated. Related auditory priming nonsense syllables shared the rime, vowel and consonant with the target (e.g., target = *frog*, related prime = *og*) whereas unrelated auditory primes did not share rime vowel and consonant with the target (e.g., target = *sock*, unrelated prime = *um*). Nine practice items (five 1 syllable words and four 2 syllable words) were included to permit participants to practice naming pictures under each of the conditions prior to the test trials. All priming nonsense syllables, including practice primes, were digitally recorded at a sampling frequency of 44100 Hz in 16 bit by the female Australian adult speaker from Study 1 using PRAAT. Each sound file was edited to ensure a lead-time of 15 ms prior to onset and finishing at offset. Primes were randomly allocated to different targets, therefore the same primes were used for both the related and unrelated conditions, controlling for any prime specific differences. For a complete list of the stimuli refer to Appendix D1.

Delayed picture-naming task. The delayed picture-naming task required the participants to name pictures presented on a computer screen after a stimulus was

presented. The name of all pictured objects were concrete nouns that were chosen to be familiar to children within the range of age of the children in this study based on age of acquisition (AOA) data from MRC Psycholinguistic Database (Coltheart, 1981). Stimuli were all 1 syllable words and were different from the items used in the picture-naming task with auditory primes. The pictures varied in frequency of occurrence to include 10 high frequency items, with a mean frequency rating of 64 occurrences per million and 10 low frequency items with a mean frequency rating of six occurrences per million (Kucera & Francis, 1967). Ten practice items, which were different from the test items, were included for each ISI. For a complete list of the stimuli refer to Appendix D2.

Procedure

Each child was tested individually in a quiet area of the language development centers or school, respectively, to minimize distractions. Participation took place over two sessions, at least a week apart. The standardized assessments were administered in the first week, described in Chapter 4, as part of the development of the classification protocol. The experimental tasks were administered in the second session in the same order for each participant; speech discrimination task, picture-naming task, NWR task and delayed picture-naming task. Participants were given adequate breaks during both sessions to maximize performance and minimize fatigue. Verbal encouragement was given in between task presentation. Each child was awarded with participation stickers and a lucky dip prize on completion each session. Each session lasted approximately one hour for each child. See the method section in Chapter 2 for procedural details of the NWR task and the speech discrimination task.

Picture-naming task. The pictures to be named were presented in the middle of the computer screen on a white background and were approximately 8cm by 6cm in size, replicating Study 1. The 15 test items were randomly split into three sets of five, with each set allocated to a different priming condition: phonological related prime, unrelated prime and no prime (i.e., silence). The items were presented in three cycles with the set rotated across conditions so that each item was presented once in each condition and each item appeared only once in a cycle. The items within each cycle were randomly presented each time it was run, therefore controlling for any order effect of those cycles across children. Each test item was

presented in each condition, once with phonologically related prime, an unrelated prime and with no prime (i.e., silence). Auditory primes were presented at two SOAs, +50ms and +150ms, after picture presentation so that each child was presented three cycles of 15 trials for each SOA. Presentation of SOA was manually counterbalanced so that half of the participants received SOA at +50 ms first and the other half received SOA at +150 ms.

Each participant named all test items prior to the commencement of the test trials. A practice phase with auditory primes was undertaken to familiarize the participants with the task and further instruction was provided as needed, prior to the commencement of the test trials. The practice phase included phonologically related primes, phonologically unrelated primes and the silence condition to ensure the participant understood the task. Participants were instructed to name the picture on the computer screen as quickly as possible and ignore what they heard via the headphones. They were also told that sometimes they would not hear anything at all.

During the test trials the picture disappeared from the screen when triggered by voice input, in the absence of a voice input it disappeared from the screen at a time out set at 3500 ms. The audio input used to trigger the voice key was automatically recorded on each trial, and each verbal response was later checked for accuracy. Reaction time was measured from onset of the verbal response. The wait time after each picture disappeared from the screen was set at 4000 ms to ensure the response was recorded in its entirety.

Delayed picture-naming task. The 20 items selected for the DPNT were presented at two ISIs, resulting in each participant naming 40 items in total. The 40 items were presented in random order, with ISI also presented in random order. Each participant was instructed to wait until a prompt (i.e., a star) appeared prior to naming the picture. Each picture was presented in the middle of the computer screen on an off-white background and was approximately 8cm x 6cm in size. Each participant named all the items prior to test trials. Practice items comprised of 10 pictures to be named. These were presented prior to the test trial to familiarize the participants with the task and further instruction was provided, as needed, prior to the commencement of the test trials. The picture disappeared from the screen when triggered by voice input. Alternatively in the absence of a voice input the picture

disappeared from the screen at a time out, set at 3500 ms. Responses were recorded on-line for later checking.

Scoring of Dependent Measures

Accuracy of NWR was scored replicating Munson et al. (2005) as detailed in Study 1 in Chapter 2. The outcome measures included; the mean NWR accuracy (percent correct) and the difference in accuracy between high and low frequency sequences, namely the PhonFreq effect. Outcome measures from the picture-naming task included; mean reaction time for the silence condition only (measured in milliseconds), the difference in reaction time between related and unrelated auditory primes (i.e., the PhonFac effect), and mean picture-naming accuracy (percent correct). For the speech discrimination task d-prime values were used as the outcome measure for all the statistical analysis undertaken, replicating Edwards et al. (2002) and consistent with Study 1. Refer to Table 29.

Data Analysis

Data analysis was undertaken in two stages. The first stage used GLMM to explore group differences on experimental tasks, consistent with the analysis undertaken in Study 1. An alpha level of .05 was used for all analyses. See Chapter 2 for further detail on the analysis approach. The second of the data analysis used a series of GLMMs implemented through SPSS (Version 22) to test predictors of phonological competency. The primary predictors were measures of speech motor ability and included the GFTA raw score, CAS trait score and the DPNrt. Potential control variables were age and the measures of expressive (EVT) and receptive (PPVT) vocabulary raw scores. Each control variable was assessed to determine if they correlated with the outcome measure first. If they were found to correlate with the outcome measures then they were included as fixed effects in the GLMM to ensure that they did not confound the relationship between the predictor and the outcome variable. The effect sizes reported from Stage 1 are partial eta squared obtained from the GLM ANOVA. Parameters omitted from the maximum likelihood output (namely, the standardized regression coefficient, the part-correlation, and the multiple correlation coefficient) were therefore obtained from the linear multiple regressions. Further explanation and rationale for this approach can be found in Chapter 2.

Table 29

Summary of Outcome Measures for Tasks Evaluating Phonological competency and Speech Motor Ability

Measure	Description
Phonological Competency	
NWR Accuracy	The mean segmental accuracy score for the low frequency sequences.
Phonotactic Frequency Effect (PhonFreq)	The difference in mean segmental accuracy scores between high and low frequency sequences in the NWR task.
Picture-naming RT (PNRT)	Mean picture-naming reaction time for the silence condition for pictures named correctly, measured in milliseconds.
Phonological Facilitation (PhonFac)	The difference in reaction time between related and unrelated auditory prime rimes.
D-prime	Measure of sensitivity for correct identification.
Speech Motor Ability	
Delayed Picture-naming RT (DPNrt)	Mean reaction time for pictures named correctly.

Results

Nonword Repetition

For the NWR task the GLMM examined group differences in the mean accuracy of high and low frequency sequences. A three-way mixed design was used for this analysis with group as a between groups IV with 3 levels (CAS, PD and TD), frequency of diphone sequence with 2 levels (high and low phonotactic frequency) and length of nonword with 2 levels (2 and 3 syllables). Prior to undertaking the GLMM we needed to test for significant covariates. Chronological age and the PPVT standard score were both significant: $F(1, 204) = 5.40, p = .021$, partial $\eta^2 =$

.14, and $F(1, 204) = 9.76, p = .002$, partial $\eta^2 = .33$, respectively. We therefore needed to include these covariates in the GLMM. Refer to Table 30.

The GLMM revealed a significant main effect for group after accounting for significant covariates, $F(2, 194) = 27.35, p < .001$, partial $\eta^2 = .53$. The main effect for frequency was significant, $F(1, 194) = 17.31, p < .001$, partial $\eta^2 = .069$, with the high frequency items ($M = 71.18\%$, $SEM = 2.83$) repeated more accurately than the low frequency items ($M = 66.52\%$, $SEM = 2.86$). The main effect of length was also significant, $F(1, 194) = 7.41, p = .007$, partial $\eta^2 = .024$, with the two syllable words ($M = 70.59\%$, $SEM = 2.66$) repeated more accurately than the three syllable words ($M = 67.11\%$, $SEM = 3.03$). The interaction between group and frequency was not significant, $F(2, 194) = 0.620, p = .539$, partial $\eta^2 = .050$. The interaction between length and group was not significant, $F(2, 194) = 0.14, p = .872$, partial $\eta^2 = .017$. The three-way interaction between group, frequency and length was not significant, $F(2, 194) = 0.52, p = .596$, partial $\eta^2 = .033$.

The interaction between frequency and length was significant, $F(1, 194) = 4.14, p = .043$, partial $\eta^2 = .070$. Follow up analysis revealed that there was a length effect for the high frequency sequences, $F(1, 194) = 9.96, p = .002$, partial $\eta^2 = .062$ but not for the low frequency sequences, $F(1, 194) = 0.61, p = .44$, partial $\eta^2 = .003$.

Further inspection of the main effect for group revealed that the children with CAS had a significantly lower mean accuracy ($M = 55.13\%$, $SEM = 3.44\%$) than the children with PD ($M = 74.42\%$, $SEM = 2.27\%$), $t(194) = 5.78, p < .001$, partial $\eta^2 = .538$, and the children with TD ($M = 77.01\%$, $SEM = 2.83\%$), $t(194) = 6.49, p < .001$, partial $\eta^2 = .533$. The children with PD were not significantly different on NWR accuracy from the children with TD, $t(194) = 0.70, p = .487$, partial $\eta^2 = .055$.

Although the interaction between group and frequency was not significant we examined if there were any trends in the difference in accuracy between the low and high frequency sequences (i.e., the phonotactic frequency effect) for each of the three groups. Inspection of the means revealed that the children with CAS had the lowest frequency effect ($M = 2.84\%$, $SEM = 1.42\%$), followed by the children with TD ($M = 5.28\%$, $SEM = 1.4\%$) and the children with PD had the highest frequency effect ($M = 6.11\%$, $SEM = 0.23\%$). Follow up analysis revealed that there was a significant effect of frequency for the children with PD, $F(1, 194) = 17.04, p < .001$, partial $\eta^2 = .103$ and the children with TD, $F(1, 194) = 9.89, p = .002$, partial $\eta^2 = .069$, but not

for the children with CAS, $F(1, 194) = 1.23, p = .27$, partial $\eta^2 = .016$. While the null result for CAS might be due to reduced power, it is notable the effect size is small for this group.

Table 30

NWR Mean Accuracy Scores (%) and Standard Error of the Mean for Children with CAS, PD and TD for Low and High Frequency Sequences for Two Syllable and Three Syllable Nonwords (N = 53)

		Low Frequency		High Frequency	
		2 Syll	3 Syll	2 Syll	3 Syll
CAS	<i>M</i>	53.43	53.98	59.82	53.27
	<i>SEM</i>	2.94	3.22	3.45	4.13
PD	<i>M</i>	73.12	69.61	79.98	74.96
	<i>SEM</i>	2.17	2.48	2.39	2.03
TD	<i>M</i>	74.85	74.14	82.36	76.70
	<i>SEM</i>	2.69	3.67	2.34	2.62

Note. Continuous predictors are fixed at the following values: Chronological Age = 72 and PPVT Standard Score = 100

Picture-Naming

The analysis for the picture-naming task examined group differences in mean reaction time for prime and SOA. Mean reaction time for each participant for each condition was calculated after excluding errors (38.8%). Reaction time outliers were defined as values more than 2 standard deviations above or below the mean reaction time for each participant within each priming condition. A total of 3.6% of data were excluded as outliers after errors were excluded. Chronological age was a significant covariate, $F(1, 312) = 31.78, p < .001$, partial $\eta^2 = .416$. Refer to Table 31.

The main effect for group was not significant, $F(2, 297) = 0.20, p = .820$, partial $\eta^2 = .026$, although numerically, the mean was lower for the children with CAS ($M = 1188$ ms, $SD = 42$ ms), compared to the children with PD ($M = 1234$ ms, $SD = 52$ ms) and TD ($M = 1202$ ms, $SD = 36$ ms).

Table 31

Picture-Naming Mean Reaction Time for Children with CAS, PD and TD for Unrelated, Related and Silence Prime Conditions for Short and Long SOAs (N = 53)

		Prime					
		Related		Unrelated		Silence	
		+50	+150	+50	+150	+50	+150
		SOA	SOA	SOA	SOA	SOA	SOA
CAS	<i>M</i>	1146	1064	1306	1216	1212	1182
	<i>SEM</i>	43	55	64	65	63	62
PD	<i>M</i>	1108	1198	1279	1309	1221	1287
	<i>SEM</i>	64	64	60	69	66	58
TD	<i>M</i>	1130	1129	1268	1285	1229	1173
	<i>SEM</i>	40	61	50	58	47	41

Note. SOA = Stimulus Onset Asynchrony. Continuous predictors are fixed at the following values: Chronological Age = 71 months.

The GLMM revealed a significant main effect for prime, $F(2, 297) = 24.41, p < .001$, partial $\eta^2 = .468$. Further inspection of the means revealed that the related prime ($M = 1129$ ms, $SD = 25$ ms) had a significantly faster mean reaction time than the unrelated prime ($M = 1277$ ms, $SD = 28$ ms), $t(297) = 6.81, p < .001$, partial $\eta^2 = .462$, and the silence condition ($M = 1217$ ms, $SD = 26$ ms), $t(297) = 4.02, p < .001$, partial $\eta^2 = .241$. The unrelated prime and silence condition were also significantly different, $t(297) = 2.47, p = .014$, partial $\eta^2 = .121$, with the silence condition having a faster reaction time than the unrelated condition.

The main effect for SOA was not significant, $F(1, 297) = 0.07, p = .790$, partial $\eta^2 = .001$. There were no significant two-way interaction effects group by prime, $F(4, 297) = 0.24, p = .915$, partial $\eta^2 = .006$, prime by SOA, $F(2, 297) = 0.09, p = .912$, partial $\eta^2 = .003$, and group by SOA, $F(2, 297) = 2.36, p = .096$, partial $\eta^2 = .080$. The three-way interaction was also not significant, $F(4, 297) = 0.44, p = .783$, partial $\eta^2 = .019$.

Although the interaction between group and prime was not significant, planned comparisons were undertaken to examine if there were trends in the difference in the phonological facilitation for the three groups. The numerical differences in the facilitation effect were small between groups. Children with CAS had the largest

facilitation effect ($M = 156$ ms, $SD = 156$ ms) and the children with PD had the smallest smaller facilitation effect ($M = 141$ ms, $SD = 166$ ms), when compared to the children with TD ($M = 148$ ms, $SD = 164$ ms). Follow up analysis comparing the related versus unrelated priming condition for each group separately confirmed significant phonological priming for each group, $p < .001$, for children with CAS, PD and TD.

Picture-naming accuracy was analysed separately to determine if the groups differed significantly on accuracy. Chronological age was a significant covariate for picture-naming accuracy, $F(1, 312) = 16.75$, $p < .001$, partial $\eta^2 = .267$. Refer to Table 32.

Table 32

Picture-naming Mean Accuracy (%) for Children with CAS, PD and TD for Related, Unrelated and Silence Priming Conditions for Short and Long SOAs (N = 53)

		Prime					
		Related		Unrelated		Silence	
		+ 50	+150	+ 50	+150	+ 50	+150
		SOA	SOA	SOA	SOA	SOA	SOA
CAS	<i>M</i>	55	58	50	50	52	46
	<i>SEM</i>	4	5	6	5	6	5
PD	<i>M</i>	64	63	56	53	57	59
	<i>SEM</i>	4	4	3	3	3	3
TD	<i>M</i>	75	74	69	65	76	68
	<i>SEM</i>	4	4	3	4	4	34

Note. SOA = Stimulus Onset Asynchrony. Continuous predictors are fixed at the following values: Chronological Age = 71.

The GLMM revealed a significant main effect for group, $F(2, 297) = 8.76$, $p < .001$, partial $\eta^2 = .265$. Further inspection of the group marginal means revealed that the children with CAS had numerically lower accuracy ($M = 52\%$, $SEM = 5\%$) than the children with PD ($M = 59\%$, $SEM = 3\%$) but this difference was not significant, $t(297) = 1.44$, $p = .152$, partial $\eta^2 = .170$. However, both the children with CAS and PD differed significantly from the children with TD ($M = 71\%$, $SEM = 4\%$), $t(297) =$

3.93, $p < .001$, partial $\eta^2 = .411$, $t(297) = 3.14$, $p = .002$, partial $\eta^2 = .075$, respectively.

The main effect for prime was also significant, $F(2, 297) = 13.48$, $p < .001$, partial $\eta^2 = .334$. Further inspection of the marginal means revealed that the related prime ($M = 65\%$, $SEM = 4\%$) had a significantly higher mean accuracy than the unrelated prime ($M = 57\%$, $SEM = 4\%$), $t(297) = 5.08$, $p < .001$, partial $\eta^2 = .327$, and the silence condition ($M = 60\%$, $SEM = 4\%$), $t(297) = 3.37$, $p < .001$, partial $\eta^2 = .181$. The unrelated prime and silent prime condition were not significantly different, $t(297) = 1.60$, $p = .111$, partial $\eta^2 = .028$.

The interaction between group and prime was not significant, $F(4, 297) = 0.69$, $p = .59$, partial $\eta^2 = .033$. The interaction between group and SOA was not significant, $F(2, 297) = 0.81$, $p = .45$, partial $\eta^2 = .019$. The interaction between group and prime and SOA was not significant, $F(4, 297) = 1.05$, $p = .38$, partial $\eta^2 = .034$.

Speech Discrimination

The analysis for the speech discrimination task used d-prime as the dependent variable to capture sensitivity to input phonetic features of the contrasting target consonants (see Chapter 2 for further explanation of d-prime). Chronological age was a significant covariate, $F(1, 314) = 60.26$, $p < .001$, partial $\eta^2 = .317$. d-prime values were analysed with word pair (i.e., *cap/cat* and *tap/tack*) and gating condition (i.e., whole word, short gate and long gate) as repeated measures independent variables, and group (i.e., CAS, PD and TD) as a between subjects independent variable. Refer to Table 33.

Table 33

Speech Discrimination Ability as Measured by d-Prime for Children with CAS, PD and TD for Whole Word, Short Gate and Long Gate and for Word Pairs, Cat/Cap and Tap/Tack (N = 53)

		Gate					
		Whole		Short		Long	
		Cat/Cap	Tap/Tac	Cat/Ca	Tap/Tac	Cat/Cap	Tap/Tack
			k	p	k		
CAS	<i>M</i>	1.27	0.59	0.79	0.74	0.79	0.20
	<i>SEM</i>	0.25	0.31	0.22	0.26	0.22	0.14
PD	<i>M</i>	1.61	1.05	0.77	0.34	0.77	0.03
	<i>SEM</i>	0.16	0.20	0.13	0.17	0.13	0.15
TD	<i>M</i>	1.91	1.49	0.94	0.97	0.94	0.44
	<i>SEM</i>	0.12	0.23	0.18	0.26	0.18	0.25

Note. Whole = whole word, Short = final stop-gap and release burst removed, Long = formant transition, stop gap and release burst removed.

The GLMM revealed that the groups were significantly different on speech discrimination ability as measured by d-prime, $F(2, 299) = 3.16$, $p = .044$, partial $\eta^2 = .123$. Further inspection of the main group effect revealed that the children with CAS ($M = 0.73$, $SEM = 0.233$) had a lower mean d-prime than the children with PD, ($M = 0.76$, $SEM = 0.157$), $t(299) = 0.13$, $p = .898$, partial $\eta^2 = .084$, and TD ($M = 1.11$, $SEM = 0.204$), $t(299) = 1.77$, $p = .077$, partial $\eta^2 = .132$, but neither of these differences were significant, although the CAS and TD contrast was close to reaching significance. However, the children with PD were significantly less accurate than the children with TD, $t(299) = 2.28$, $p = .024$, partial $\eta^2 = .139$.

The GLMM revealed that the main effect for gate was significant, $F(2, 299) = 29.88$, $p < .001$, partial $\eta^2 = .490$. Further inspection of the marginal means revealed that the whole word condition ($M = 1.32$, $SEM = 0.213$) had a significantly higher d-prime than the short gate ($M = 0.76$, $SEM = 0.202$), $t(299) = 6.58$, $p < .001$, partial $\eta^2 = .455$, and long gate ($M = 0.53$, $SEM = 0.178$), $t(299) = 7.48$, $p < .001$, partial $\eta^2 = .490$, and the short gate had a significantly higher d-prime than the long gate, $t(299) = 2.95$, $p = .003$, partial $\eta^2 = .128$. The main effect for word pair was significant, $F(1, 299) = 22.42$, $p < .001$, partial $\eta^2 = .244$, with the cat/cap word pair, ($M = 1.086$,

SEM = 0.177) showing stronger sensitivity, than the tap/tack word pair ($M = 0.650$, SEM = 0.219).

The interaction between gate and group was significant $F(4, 299) = 2.64$, $p = .034$, partial $\eta^2 = .075$. Further inspection of this interaction revealed that the groups did not differ significantly in accuracy for the short gate condition, $F(2, 99) = 1.83$, $p = .16$, partial $\eta^2 = .040$, or for the long gate condition, $F(2, 99) = 1.53$, $p = .22$, partial $\eta^2 = .020$. The interaction between gate and group was significant for the whole word condition, $F(2, 99) = 3.98$, $p = .022$, partial $\eta^2 = .164$. The children with CAS did not differ significantly from the children with PD, $t(99) = 1.11$, $p = .269$, partial $\eta^2 = .034$, although numerically, the children with CAS ($M = 0.96$, SEM = 0.28) had a lower d-prime than the children with PD ($M = 1.21$, SEM = 0.18). The children with CAS had a significantly lower d-prime than the children with TD ($M = 1.70$, SEM = 0.17), $t(99) = 2.56$, $p = .012$, partial $\eta^2 = .191$. The children with PD did not differ significantly from the children with TD, although this difference was close to significance, $t(99) = 1.83$, $p = .071$, partial $\eta^2 = .012$.

The interaction between gate and pair was significant when all gates were included in the analysis, $F(2, 299) = 6.84$, $p = .001$, partial $\eta^2 = .010$. The interaction for the whole word condition with the word pair was significant, $t(302) = 4.10$, $p < .001$, partial $\eta^2 = .073$, with the d-prime being higher for the cat/cap pair ($M = 1.59$, SEM = 0.18) than the tap/tack pair ($M = 1.4$, SEM = 0.25). The gate pair interaction was significant for the long gate condition, $t(302) = 3.95$, $p < .001$, partial $\eta^2 = .118$, with the d-prime being higher for the cat/cap ($M = 2.5$, SEM = 0.53) than the tap/tack pair ($M = 0.67$, SEM = 0.54). The interaction between gate and pair for the short condition was not significant, $t(302) = 1.22$, $p = .22$, partial $\eta^2 = .007$, with a nominal difference in d-prime for the cat/cap ($M = 2.5$, SEM = 0.53) and the tap/tack pair ($M = 2.05$, SEM = 0.69). The interaction between group and pair was not significant, $F(2, 303) = 0.84$, $p = .432$, partial $\eta^2 = .019$. The three-way interaction between group, gate and pair was not significant, $F(4, 299) = 0.47$, $p = .756$, partial $\eta^2 = .002$.

Delayed Picture-naming

The analysis for the delayed picture-naming task examined group differences in mean reaction time. Mean reaction time for each participant for each condition was calculated after excluding errors (33.8%). Errors included wrong names, false starts, coughing and touching microphone. Reaction time outliers were defined as

values more than 2 standard deviations above or below the mean reaction time for each participant. A total of 7% of data were excluded as outliers after errors were excluded. Chronological age was a significant covariate, $F(1, 208) = 15.78, p < .001$, partial $\eta^2 = .272$. Refer to Table 34.

Table 34

Delayed Picture-naming Mean Reaction Time for Children with CAS, PD and TD for Low and High Frequency Words and Short and Long ISI's (N = 53)

		Low Frequency		High Frequency	
		Short ISI	Long ISI	Short ISI	Long ISI
CAS	<i>M</i>	797	728	767	619
	<i>SEM</i>	68	72	77	39
PD	<i>M</i>	683	625	688	608
	<i>SEM</i>	55	39	34	34
TD	<i>M</i>	587	503	535	502
	<i>SEM</i>	50	24	33	23

Note. ISI = inter-stimulus interval.

The GLMM revealed a significant difference between groups for reaction time, $F(2, 199) = 7.03, p = .001$, partial $\eta^2 = .196$. Further inspection the marginal means revealed that children with CAS had a slower reaction time ($M = 727$ ms, $SEM = 64$ ms) than the children with PD ($M = 651$ ms, $SEM = 41$ ms) but this difference was not significant, $t(199) = 1.18, p = .239$, partial $\eta^2 = .024$. The CAS reaction time was significantly slower than the children with TD ($M = 532$ ms, $SEM = 33$ ms), $t(199) = 3.14, p = .002$, partial $\eta^2 = .239$, the children with PD also had a significantly slower reaction time than children with TD, $t(199) = 2.85, p = .005$, partial $\eta^2 = .159$. The main effect for ISI was significant, $F(1, 199) = 17.97, p < .001$, partial $\eta^2 = .268$, with the shorter ISI having a slower reaction time ($M = 676$ ms, $SD = 28$ ms) than the longer ISI ($M = 597$ ms, $SD = 21$ ms). The main effect for frequency was not significant, $F(1, 199) = 2.86, p = .097$, partial $\eta^2 = .054$. There were no significant interactions, group and word frequency was not significant, $F(2, 199) = 0.75, p = .47$, partial $\eta^2 = .013$, group and ISI was not significant, $F(2, 199) = 0.54, p = .58$, partial $\eta^2 = .015$, word frequency and ISI was not significant, $F(1, 199) = 0.23, p = .63$,

partial $\eta^2 = .000$, and the three-way interaction between group, word frequency and ISI was not significant, $F(2, 199) = 1.13, p = .33$, partial $\eta^2 = .042$

Chronological age was a significant covariate for delayed picture-naming accuracy, $F(1, 208) = 21.84, p < .001$, partial $\eta^2 = .406$. The GLMM revealed a significant main effect of group, $F(2, 199) = 8.73, p < .001$, partial $\eta^2 = .226$. The children with CAS did not differ significantly from the children with PD $t(199) = 0.39, p = .696$, partial $\eta^2 = .006$, although numerically the children with CAS had a higher mean accuracy score ($M = 63\%$, $SEM = 3.05$) than the children with PD ($M = 61\%$, $SEM = 2.41$). The children with CAS and PD both had a significantly lower mean accuracy score than the children with TD ($M = 74\%$, $SEM = 2.31$), $t(199) = 2.95, p = .004$, partial $\eta^2 = .239$, $t(199) = 3.91, p < .001$, partial $\eta^2 = .265$, respectively. The main effect for ISI was significant, $F(1, 199) = 17.97, p < .001$, partial $\eta^2 = .268$, with the shorter ISI having a slower reaction time ($M = 676$, $SEM = 39$) than the longer ISI ($M = 596$, $SEM = 39$). The main effect for frequency was not significant, $F(1, 199) = 2.86, p = .097$, partial $\eta^2 = .054$.

Regression Analysis

The second stage of analysis used a series of hierarchical regressions in the context of GLMM to examine predictors of outcome measures of phonological competence and to determine if speech motor measures predicted phonological competence and if this relationship varied across groups. Tables 35 and 36 contain bivariate correlations (Pearson's product-moment correlations) between the different outcome measures and Tables 37, 38, 39 and 40 show the correlations between the predictor measures and outcome measures from the different tasks for the whole group and for each group individually.

Inspection of these tables indicated that, as a whole group, there were a number of significant correlations between the measures of phonological competence (Table 35). The correlations between the phonological measures for the whole sample revealed significant correlations between the NWR phonotactic frequency effect and NWR accuracy for the low frequency sequences. d-prime was correlated with NWR accuracy for the high and low frequency sequences. Picture-naming reaction time and phonological facilitation effects did not correlate with any of the other measures of phonological competency. When we looked at the groups individually different correlations emerged (refer to Tables 38, 39 and 40). For the children with CAS

there were no significant correlations between the phonotactic frequency effect and other measures of phonological competence, whereas for the children with PD and TD the phonotactic frequency effect was highly correlated with NWR accuracy for the low frequency sequences. In addition, the NWR accuracy for the low and high frequency sequences was positively correlated with d-prime for the children with PD and TD, but not for the children with CAS. Picture-naming reaction time was positively correlated with NWR for the low and high frequency sequences for the children with CAS and for the children with TD, but not for the children with PD. The phonological facilitation effect correlated positively with d-prime for the children with CAS but not for the children with PD or TD.

We also explored the relationship between outcome measures with age and vocabulary. Inspection of the correlations in Tables 37 indicates that outcome measures of phonological competence were well predicted by a variety of measures of vocabulary and motor measures for the entire group. However, when we looked at potential predictors for each of the individual groups they differed in relation to these associations. For the children with CAS (Table 39) the only outcome measure reliably predicted by age and vocabulary was the picture-naming reaction time. However, a number of these correlations were close reaching significance and therefore statistical power may be in issue for these correlations. Nonword repetition accuracy for the high frequency sequences was also correlated with the EVT standard score. For the children with PD (Table 40), NWR accuracy for the low frequency sequences was positively correlated with the PPVT, PNrt was negatively correlated with age and PPVT and picture-naming accuracy was positively correlated with age and PPVT and EVT, as was d-prime. For the children with TD (Table 41), NWR accuracy was positively correlated with the PPVT and EVT, but not age; and picture-naming reaction time was positively correlated with age, PPVT and EVT, as was d-prime.

We then looked at the motor measures as potential predictors of the outcome measures of phonological competence. For the whole group all three motor measures correlated with NWR for the low frequency sequences, the high frequency sequences, picture-naming accuracy and d-prime (Table 38). The DPNrt correlated with the phonotactic frequency effect and the PNrt. None of the motor measures correlated with the phonological facilitation effect. When we looked at the groups

individually, the CAS trait score correlated negatively with NWR for the low frequency sequences, the GFTA correlated negatively with the NWR for high frequency sequences and the PNrt and the DPNrt correlated positively with the phonotactic frequency effect, for the children with CAS (Table 39). For the children with PD (Table 40), the CAS trait score negatively correlated with the NWR for low frequency sequences, the GFTA negatively correlated with the NWR high frequency sequences and the DPNrt positively correlated with the PNrt. For the children with TD (Table 41), the only motor measure to correlate with any of the outcome measures was the DPNrt, which positively correlated with the PNrt.

Although not included in the correlation tables, we looked at the correlations between the motor measures and measures of vocabulary and age. The GFTA was highly correlated with both the EVT standard score ($-.488, p < .01$) and the PPVT standard score ($-.471, p < .01$) but not with chronological age ($-.172, p > .05$). The DPNRT was highly correlated with chronological age ($-.522, p < .01$) and the PPVT standard score ($-.349, p < .05$) but not with the EVT standard score ($-.248, p > .05$). The CAS trait score was highly correlated with both the EVT standard score ($-.434, p < .05$) and the PPVT standard score ($-.534, p < .01$) and chronological age ($-.426, p < .05$). The motor measures were also correlated with one another. The GFTA was positively correlated with the DPNrt ($.442, p < .05$) and the CAS trait score ($.765, p < .01$). The DPNrt and the CAS trait score were also positively correlated ($.863, p < .01$).

The hierarchical regression analysis proceeded in two stages. Stage 1 tested whether chronological age, the PPVT raw score, and the EVT raw score were significant predictors of the various outcome measures of phonological competence and were included as covariates in the regression analysis if they were significant. Primary predictors assessing speech motor ability included the GFTA raw score, the CAS trait score (CAS_{ts}) and DPNrt. Stage 2 tested whether the primary predictor, (Step 1), group (Step 2), and the predictor x group interaction (Step 3) accounted for significant proportions of variance in the outcome measures after controlling for the covariates identified at Stage 1. The regressions that produced non-significant results for all steps at Stage 2 of the analysis are reported in the appendix.

Table 35

Pearson Correlations Between Measures of Phonological Competency for All Participants (N = 53).

	FreqEff	NWRLow	NWRHigh	PhonFac	PNrt	D-prime
FreqEff	-					
NWRLow	-.326*	-				
NWRHigh	.185	.866**	-			
PhonFac	.076	.034	.081	-		
PNrt	.176	-.123	-.032	-.046	-	
D-prime	-.147	.458**	.407**	.078	-.190	-

Note. FreqEff = phonotactic frequency effect; NWRLow = NWR accuracy for low frequency sequences; NWRHigh = NWR accuracy for high frequency sequences; PhonFac = Phonological facilitation; PNrt = Picture-naming mean reaction time; D-prime = speech discrimination ability.

Table 36

Pearson Correlations Between Measures of Phonological Competency for the Children with CAS (n = 14), PD (n = 21) and TD (n = 18)

	FreqEff			NWRLow			NWRHigh			PhonFac			PNrt		
	CAS	PD	TD	CAS	PD	TD	CAS	PD	TD	CAS	PD	TD	CAS	PD	TD
FreqEff	-	-	-												
NWRLow	-.418	-.713**	-.695**	-	-	-									
NWRhigh	.345	.063	-.135	.690**	.655**	.807*	-	-	-						
						*									
PhonFac	.114	.428	-.347	-.033	-.054	.270	.070	.385	.086	-	-	-			
PNrt	.165	.234	.217	.580*	-.162	-.525*	.740**	.022	-.545*	-.258	-.101	.184	-	-	-
d-prime	-.107	-.320	-.303	-.059	.447*	.481*	-.103	.292	.414	.545*	.024	-.193	-.184	-.091	-.269

Note. FreqEff = phonotactic frequency; NWRLow = NWR accuracy for low frequency sequences; NWRHigh = NWR accuracy for high frequency sequences; PhonFac = Phonological facilitation; PNrt = Picture-naming mean reaction time; D-prime = speech discrimination ability.

Table 37

Correlations Between Outcome Measures, Measures of Vocabulary and Articulation and Predictors for the Whole Group (N = 53).

	NWrep Low	NWrep High	FreqEff	PhonFac	PNPC	PNrt	D-prime
Age	.265	.147	-.224	.060	.505**	-.646**	.530**
PPVT Raw score	.572**	.401**	-.364**	.033	.549**	-.559**	.565**
Standard score	.541**	.422**	-.286*	-.049	.229	-.080	.232
EVT Raw score	.512**	.373**	-.297*	-.013	.562**	-.530**	.560**
Standard score	.368**	.296*	-.180	-.104	.167	.097	.130
Motor Measures							
GFTA Raw score	-.721**	-.818**	-.144	-.046	-.485**	.037	-.416**
CAS Trait Score	-.745**	-.783**	-.045	.029	-.496**	.070	-.342*
DPNrt	-.483**	-.327*	.316*	.065	-.475**	.374**	-.388**

Note. NWrepLow = mean accuracy for low frequency sequences; NWrepHigh = mean accuracy for high frequency sequences; FreqEff = phonotactic frequency effect; PhonFac = phonological facilitation effect; PNPC = picture-naming accuracy; PNrt = mean picture-naming reaction time for silence condition; D-prime = speech perception for whole word; CAS Trait Score = predicted value for each participant based on features of CAS; DPNrt = delayed picture-naming mean reaction time.

** . Correlation is significant at the .01 level (2-tailed).

* . Correlation is significant at the .05 level (2-tailed).

Table 38

Correlations Between Outcome Measures, Measures of Vocabulary and Articulation and Predictors for children with CAS (N = 14)

	NWrep Low	NWrep High	FreqEff	PhonFac	PNPC	PNrt	D-prime
Age	-.103	-.420	-.355	.314	.337	-.684**	.424
PPVT: Raw score	.160	-.234	-.493	.204	.308	-.603*	.153
Standard score	.477	.244	-.374	-.172	-.088	.048	-.487
EVT: Raw score	.068	-.171	-.274	.249	.474	-.625*	.268
Standard score	.326	.537*	.217	-.204	.153	.326	-.393
Motor Measures							
GFTA Raw score	-.389	-.602*	-.301	-.211	-.084	-.669**	-.018
CAS Trait Score	-.564*	-.392	.163	.083	-.312	-.121	.301
DPNrt	-.473	.033	.642*	-.028	-.339	.099	-.288

Note. NWrepLow = mean accuracy for low frequency sequences; NWrepHigh = mean accuracy for high frequency sequences; FreqEff = phonotactic frequency effect; PhonFac = phonological facilitation effect; PNPC = picture-naming accuracy; PNrt = picture-naming reaction time; D-prime = speech perception for whole word; CAS Trait Score = predicted value for each participant based on features of CAS; DPNrt = delayed picture-naming mean reaction time.

**. Correlation is significant at the .01 level (2-tailed).

*. Correlation is significant at the .05 level (2-tailed).

Table 39

Correlations Between Outcome Measures, Measures of Vocabulary and Articulation and Predictors for children with PD (N = 21)

	NWrep Low	NWrep High	FreqEff	PhonFac	PNPC	PNrt	D-prime
Age	.245	.111	-.221	.123	.786**	-.623**	.589**
PPVT: Raw score	.503*	.283	-.402	-.045	.536*	-.474*	.690**
Standard score	.453*	.275	-.343	-.319	-.103	.075	.402
EVT: Raw score	.354	.156	-.323	-.020	.568**	-.420	.655**
Standard score	.241	.106	-.220	-.214	-.112	.280	.261
Motor Measures							
GFTA Raw score	-.326	-.448*	.014	-.112	-.182	.184	-.224
CAS Trait Score	-.530*	-.394	.335	.003	-.049	.141	-.310
DPNrt	-.155	.238	.425	.032	-.346	.543*	-.295

Note. NWrepLow = mean accuracy for low frequency sequences; NWrepHigh = mean accuracy for high frequency sequences; FreqEff = phonotactic frequency effect; PhonFac = phonological facilitation effect; PN PC = picture-naming accuracy; PNrt = picture-naming reaction time; D-prime = speech perception for whole word; CAS Trait Score = predicted value for each participant based on features of CAS; DPNrt = delayed picture-naming mean reaction time.

**. Correlation is significant at the .01 level (2-tailed).

*. Correlation is significant at the .05 level (2-tailed).

Table 40

Correlations Between Outcome Measures, Measures of Vocabulary and Articulation and Predictors for children with TD (N = 18)

	NWrep Low	NWrep High	FreqEff	PhonFac	PNPC	PNrt	D-prime
Age	.434	.270	-.400	-.180	.236	-.676**	.478*
PPVT: Raw score	.653**	.502*	-.484*	.045	.420	-.736**	.601**
Standard score	.592**	.516*	-.365	.280	.370	-.479*	.440
EVT: Raw score	.586*	.383	-.517*	-.128	.343	-.744**	.549*
Standard score	.394	.118	-.518*	-.022	.157	-.249	.268
Motor Measures							
GFTA Raw score	-.301	-.198	.264	-.068	.077	.133	-.350
CAS Trait Score	.066	.211	.147	-.137	.234	-.344	-.127
DPNrt	-.081	-.085	.032	.275	-.306	.618**	-.096

Note. NWrepLow = mean accuracy for low frequency sequences; NWrepHigh = mean accuracy for high frequency sequences; FreqEff = phonotactic frequency effect; PhonFac = phonological facilitation effect; PNPC = picture-naming accuracy; PNrt = picture-naming reaction time; D-prime = speech perception for whole word; CAS Trait Score = predicted value for each participant based on features of CAS; DPNrt = delayed picture-naming mean reaction time.

** . Correlation is significant at the .01 level (2-tailed).

* . Correlation is significant at the .05 level (2-tailed).

Predicting Nonword Repetition Accuracy

The PPVT raw score significantly predicted the NWR accuracy for the low frequency sequences, $t(51) = 5.53$, $p < .001$, $sr^2 = .327$, as did EVT raw score, $t(51) = 5.02$, $p < .001$, $sr^2 = .263$, and chronological age, $t(51) = 2.34$, $p = .023$, $sr^2 = .070$. Refer to Table 41.

GFTA raw score predicting NWR accuracy. The GFTA raw score was a significant predictor of the NWR accuracy after controlling for EVT and PPVT raw scores and chronological age ($p < .001$); as the GFTA raw score increased, the NWR accuracy for the low frequency sequences decreased. GFTA did not interact with group ($p = .893$) indicating that this relationship held across the three groups. The PPVT raw score predicted NWR accuracy for the low frequency sequences on each of the three steps (Step 1: $p < .001$; Step 2: $p = .020$; Step 3: $p < .001$); as the PPVT score increased, NWR accuracy for the low frequency sequences also increased. Chronological age predicted NWR accuracy for the low frequency sequences for step 1 ($p < .001$) and step 3 of the analysis ($p < .001$). We expected a positive association between age and accuracy, with accuracy increasing with age, however, the association was in the opposite direction. This association may have been confounded by the groups differing on age, with the children with PD being the eldest ($M = 77$, $SD = 13$), compared to the children with TD ($M = 70$, $SD = 11$) and children with CAS ($M = 60$, $SD = 13$).

CAS trait score predicting NWR accuracy. The CAS trait score was a significant predictor of the nonword mean accuracy for the low frequency sequences after controlling for EVT and PPVT raw scores and chronological age ($p < .001$); as the CAS trait score increased, the accuracy for the low frequency sequences decreased. CAS trait score did not interact with group ($p = .883$) indicating that this relationship held across the three groups. Refer to Table 42.

Table 41

Unstandardized (B) and standardized (β) Regression Coefficients, and Squared Part Correlations (sr^2) for a Hierarchical Multiple Regression Model Predicting NWR Mean Accuracy for Low Frequency Sequences from GFTA Raw Scores, Group, and the Group x GFTA Interaction (N = 53)

Predictors (IVs)	B	95% CI ¹	B	sr^2	p-value ¹
DV: NWrep Low					
Step 1					
EVT raw	0.017	-0.322, 0.357	.014	.000	.918
PPVT raw	0.413	0.208, 0.617	.615	.076	.000**
Chron Age	-0.372	-0.613, -0.131	-.326	.045	.003**
GFTA raw	-0.557	-0.720, -0.395	-.581	.296	.000**
$R^2 = .691, p = .000^{**}$					
Step 2					
EVT raw	-0.012	-0.289, 0.265	.151	.005	.932
PPVT raw	-0.206	-0.377, -0.034	.713	.094	.020*
Chron Age	0.099	-0.132, 0.331	-.575	.080	.391
GFTA raw	-0.039	-0.280, 0.202	-.558	.079	.746
Group (D1) ²	-5.209	-15.749, 5.331	.031	.000	.325
Group (D2) ²	-1.511	-7.748, 4.727	.285	.024	.628
Group: $F(2, 46) = 0.58, p = .566$					
$\Delta R^2 = .045, p = .026^*$					
$R^2 = .736, p = .000^{**}$					
Step 3					
EVT raw	0.178	-0.145, 0.500	.142	.004	.274
PPVT raw	0.479	0.280, 0.678	.714	.094	.000**
Chron Age	-0.639	-0.974, -0.304	-.560	.072	.000**
GFTA raw	-0.569	-1.170, 0.032	-.593	.011	.063
Group (D1) ²	2.960	-9.158, 15.078	.089	.001	.625
Group (D2) ²	8.968	-0.394, 18.329	.298	.010	.060
Group: $F(2, 44) = 2.47, p = .096^3$					
Group (D1)xGFTA ⁴	-0.620	-11.698, 10.457	-.028	.000	.911
Group (D2)xGFTA ⁴	1.738	-10.075, 13.551	.041	.000	.768
Group x GFTA : $F(2, 44) = 0.11, p = .893^5$					
$\Delta R^2 = .001, p = .891$					
$R^2 = .738, p = .000^{**}$					

1: These are the GLMM adjusted values.

2: These are dummy variables; in combination, they reflect the group effect.

3: This is the overall F -value for the group effect.

4: The interaction term is computed using centered GFTA raw scores.

5: This is the overall F -value for the Group x GFTA interaction effect.

Table 42

Unstandardized (B) and standardized (β) Regression Coefficients, and Squared Part Correlations (sr^2) for a Hierarchical Multiple Regression Model Predicting NWR Mean Accuracy for Low Frequency Sequences from CAS Trait Score (CASts), Group, and the Group x CASts Interaction (N = 53)

Predictors (IVs)	B	95% CI ¹	B	sr^2	p-value ¹
DV: NWRepLow					
Step 1					
EVT raw	0.048	-0.295, 0.392	.039	.000	.778
PPVT raw	0.343	0.129, 0.557	.511	.051	.002**
Chron Age	-0.344	-0.539, -0.150	-.302	.038	.001**
CASts	-9.272	-12.433, -6.111	-.587	.275	.000**
$R^2 = .670, p = .000^{**}$					
Step 2					
EVT raw	0.136	-0.240, 0.512	.109	.003	.471
PPVT raw	0.392	0.150, 0.634	.584	.062	.002**
Chron Age	-0.500	-0.831, -0.169	-.438	.048	.004**
CASts	-6.329	-11.525, -1.134	-.401	.031	.018*
Group (D1) ²	-5.619	-16.619, 5.381	-.168	.005	.309
Group (D2) ²	-2.916	-5.169, 11.001	.097	.003	.471
Group: $F(2, 46) = 1.45, p = .246$					
$\Delta R^2 = .019, p = .263$					
$R^2 = .688, p = .000^{**}$					
Step 3					
EVT raw	0.128	-0.276, 0.532	.102	.002	.527
PPVT raw	0.382	0.153, 0.610	.568	.055	.002**
Chron Age	-0.485	-0.803, -0.166	-.425	.044	.004**
CASts	-2.531	-21.838, 16.777	-.160	.000	.793
Group (D1) ²	-8.769	-26.160, 8.622	-.262	.006	.315
Group (D2) ²	-0.718	-17.633, 16.198	-.024	.000	.932
Group: $F(2, 44) = 1.41, p = .256^3$					
Group (D1)xCASts ⁴	-3.845	-23.949, 16.258	-.178	.001	.702
Group (D2)xCASts ⁴	-5.603	-28.451, 17.246	-.087	.002	.624
Group x CASts : $F(2, 44) = 0.12, p = .883^5$					
$\Delta R^2 = .002, p = .896$					
$R^2 = .690, p = .000^{**}$					

1: These are the GLMM adjusted values.

2: These are dummy variables; in combination, they reflect the group effect.

3: This is the overall F -value for the group effect.

4: The CAS trait score is already centered.

5: This is the overall F -value for the Group x CAS trait score interaction effect.

The PPVT raw score predicted the NWR for low frequency sequences on each of the three steps (Step 1: $p = .002$; Step 2: $p = .002$; Step 3: $p = .002$); as the PPVT score increased, the NWR accuracy increased. Chronological age also predicted the NWR for low frequency sequences on each of the three steps (Step 1: $p = .001$; Step 2: $p = .004$; Step 3: $p = .004$), although, as previously mentioned, the groups differing in age have confounded the direction of this association.

DPNrt predicting NWR accuracy. The DPNrt was a significant predictor of the nonword mean accuracy for the low frequency sequences after controlling for EVT and PPVT raw scores and chronological age ($p = .031$); as the DPNrt increased, the NWR accuracy decreased. DPNrt did not interact with group ($p = .132$) indicating that this relationship held across the three groups. Refer to Table 43.

The PPVT raw score predicted the NWR for low frequency sequences on each of the three steps (Step 1: $p = .002$; Step 2: $p = .001$; Step 3: $p = .001$); as the PPVT score increased, the NWR accuracy increased. Chronological age also predicted the NWR for low frequency sequences on each of the three steps (Step 1: $p = .002$; Step 2: $p = .003$; Step 3: $p = .002$), although, as previously mentioned, the groups differing in age have confounded the direction of this association

Predicting the Phonotactic Frequency Effect

The phonotactic frequency effect was measured by the difference in mean segmental accuracy scores between high and low frequency sequences in the NWR task. The PPVT raw score significantly predicted the phonotactic frequency effect, $t(51) = 2.58$, $p = .013$, $sr^2 = .132$, as did EVT raw score, $t(51) = 2.22$, $p = .031$, $sr^2 = .088$. Chronological age was not a significant predictor of the phonotactic frequency effect, $t(51) = 1.47$, $p = .149$, $sr^2 = .050$.

GFTA raw score predicting the phonotactic frequency effect. The GFTA raw score was a significant predictor of the phonotactic frequency effect after controlling for EVT and PPVT raw scores ($p = .038$), as the GFTA score increased (indicating more articulation errors), the phonotactic frequency effect decreased. However, the GFTA did not interact with group ($p = .597$) after including group and interaction between group and GFTA, suggesting that the GFTA did not independently predict the phonotactic frequency effect after accounting for group differences in the prediction. The PPVT raw score predicted the phonotactic frequency effect on each of the three steps (albeit just failing to reach significance on Step 1) (Step 1: $p = .053$; Step 2: $p = .032$; Step 3: $p = .041$); as the PPVT score increased, the phonotactic frequency effect decreased. Refer to Table 44.

Table 43

Unstandardized (B) and standardized (β) Regression Coefficients, and Squared Part Correlations (sr^2) for a Hierarchical Multiple Regression Model Predicting NWR Mean Accuracy for Low Frequency Sequences from DPNrt, Group, and the Group x DPNrt Interaction (N = 53)

Predictors (IVs)	B	95% CI ¹	B	sr^2	p-value ¹
DV: NWRepLow					
Step 1					
EVT raw	0.135	-0.288, 0.557	.108	.003	.524
PPVT raw	0.446	0.175, 0.717	.664	.088	.002**
Chron Age	-0.540	-0.875, -0.204	-	.091	.002**
			.473		
DPNrt	-0.022	-0.042, -0.002	-	.062	.031*
			.305		
$R^2 = .457, p = .000^{**}$					
Step 2					
EVT raw	0.155	-0.248, 0.558	.124	.003	.444
PPVT raw	0.424	0.175, 0.674	.632	.074	.001**
Chron Age	-0.595	-0.983, -0.208	-	.064	.003**
			.522		
DPNrt	-0.009	-0.025, 0.006	-	.009	.242
			.128		
Group (D1) ²	-16.349	-24.739, -7.959	-	.119	.000**
			.489		
Group (D2) ²	1.128	-7.596, 9.852	.037	.001	.796
Group: $F(2, 46) = 15.96, p < .000$					
$\Delta R^2 = .210, p = .000^{**}$					
$R^2 = .667, p = .000^{**}$					
Step 3					
EVT raw	0.192	-0.193, 0.576	.153	.005	.320
PPVT raw	0.446	0.203, 0.689	.664	.081	.001**
Chron Age	-0.618	-0.992, -0.245	-	.068	.002**
			.542		
DPNrt	0.014	-0.020, 0.049	.201	.004	.398
Group (D1) ²	-16.287	-24.962, -7.613	-	.144	.000**
			.487		
Group (D2) ²	-0.099	-8.980, 8.782	-	.000	.982
			.003		
Group: $F(2, 44) = 14.22, p < .001^3$					
Group (D1)xDPNrt ⁴	-6.856	-14.553, 0.841	-	.017	.079
			.337		
Group (D2)xDPNrt ⁴	-2.719	-9.470, 4.033	-	.002	.421
			.093		
Group x DPNrt : $F(2, 44) = 2.13, p = .132^5$					
$\Delta R^2 = .023, p = .213$					
$R^2 = .690, p = .000^{**}$					

- 1: These are the GLMM adjusted values.
- 2: These are dummy variables; in combination, they reflect the group effect.
- 3: This is the overall F -value for the group effect.
- 4: The interaction term is computed using centered DPNrts.
- 5: This is the overall F -value for the Group x DPNrt interaction effect.

CAS trait score predicting the phonotactic frequency effect. The CAS trait score was not a significant predictor of the phonotactic frequency effect after controlling for EVT and PPVT raw scores ($p = .079$). The CAS trait score did not interact with group ($p = .264$) indicating that this relationship held across the three groups. The PPVT raw score predicted the phonotactic frequency effect on each of the three steps (albeit just failing to reach significance on Step 1) (Step 1: $p = .054$; Step 2: $p = .050$; Step 3: $p = .030$); as the PPVT score increased, the phonotactic frequency effect decreased. Refer to Table 45.

Table 44

Unstandardized (B) and standardized (β) Regression Coefficients, and Squared Part Correlations (sr^2) for a Hierarchical Multiple Regression Model Predicting the Phonotactic frequency effect from GFTA Raw Scores, Group, and the Group x GFTA Interaction ($N = 53$)

Predictors (IVs)	B	95% CI ¹	B	sr^2	p-value ¹
DV: FreqEff					
Step 1					
EVT raw	0.014	-0.265, 0.294	.023	.000	.918
PPVT raw	-0.016	-0.328, 0.002	-.480	.055	.053
GFTA raw	-0.143	-0.278, -0.008	-.294	.076	.038*
$R^2 = .210, p = .009^{**}$					
Step 2					
EVT raw	0.026	-0.246, 0.299	.042	.000	.846
PPVT raw	-0.174	-0.333, -0.016	-.513	.063	.032*
GFTA raw	-0.054	-0.281, 0.174	-.111	.007	.637
Group (D1) ²	-4.062	-13.418, 5.294	-.240	.011	.387
Group (D2) ²	-0.032	-4.738, 4.673	-.002	.000	.989
Group: $F(2, 47) = 0.63, p = .538$					
$\Delta R^2 = .022, p = .522$					
$R^2 = .232, p = .022^*$					
Step 3					
EVT raw	0.040	-0.219, 0.299	.063	.001	.757
PPVT raw	-0.169	-0.332, -0.007	-.513	.059	.041*
GFTA raw	0.191	-0.411, 0.793	-.394	.005	.525
Group (D1) ²	-5.228	-17.093, 6.637	-.309	.010	.380
Group (D2) ²	-3.295	-12.650, 6.061	-.166	.007	.482
Group: $F(2, 45) = 0.40, p = .673^3$					
Group (D1)xGFTA ⁴	-5.354	-15.949, 5.241	-.470	.012	.314
Group (D2)xGFTA ⁴	-3.545	-14.054, 6.963	-.166	.006	.500
Group x GFTA : $F(2, 45) = 0.52, p = .597^5$					
$\Delta R^2 = .012, p = .692$					
$R^2 = .244, p = .066$					

1: These are the GLMM adjusted values.

2: These are dummy variables; in combination, they reflect the group effect.

3: This is the overall F -value for the group effect.

4: The interaction term is computed using centered GFTA raw scores.

5: This is the overall F -value for the Group x GFTA interaction effect.

Table 45

Unstandardized (B) and standardized (β) Regression Coefficients, and Squared Part Correlations (sr^2) for a Hierarchical Multiple Regression Model Predicting the Phonotactic frequency effect from CAS Trait Scores, Group, and the Group x CAS Trait Score Interaction (N = 53).

Predictors (IVs)	B	95% CI ¹	B	sr^2	p-value ¹
DV: FreqEff					
Step 1					
EVT raw	0.028	-0.256, 0.313	.045	.000	.842
PPVT raw	-0.173	-0.349, 0.003	-.510	.061	.054
CASSts	-1.973	-4.187, 0.240	-.247	.049	.079
$R^2 = .183, p = .018^*$					
Step 2					
EVT raw	0.028	-0.239, 0.295	.044	.000	.833
PPVT raw	-0.167	-0.333, -0.000	-.490	.056	.050*
CASSts	1.596	-1.881, 5.073	.200	.010	.361
Group (D1) ²	-9.073	-17.266, -0.879	-.536	.058	.031*
Group (D2) ²	-1.509	-5.745, 2.727	-.099	.005	.477
Group: $F(2, 47) = 2.62, p = .084$					
$\Delta R^2 = .053, p = .206$					
$R^2 = .236, p = .023^*$					
Step 3					
EVT raw	0.069	-0.192, 0.329	.108	.003	.599
PPVT raw	-0.188	-0.357, -0.020	-.555	.068	.030*
CASSts	7.353	-2.278, 16.985	.921	.015	.131
Group (D1) ²	-	-21.251, -2.005	-.687	.049	.019*
	11.628				
Group (D2) ²	-5.506	-14.793, 3.782	-.361	.117	.239
Group: $F(2, 45) = 3.52, p = .038^3$					
Group (D1)xCASSts ⁴	-7.203	-17.647, 3.241	-.660	.013	.172
Group (D2)xCASSts ⁴	-3.826	-16.932, 9.280	-.117	.065	.559
Group x CASSts : $F(2, 45) = 1.37, p = .264^5$					
$\Delta R^2 = .016, p = .623$					
$R^2 = .502, p = .055$					

1: These are the GLMM adjusted values.

2: These are dummy variables; in combination, they reflect the group effect.

3: This is the overall F -value for the group effect.

4: CAS trait score is already centered

5: This is the overall F -value for the Group x CAS trait score interaction effect.

DPNrt predicting the phonotactic frequency effect. The DPNrt was not a significant predictor of the phonotactic frequency effect after controlling for EVT and PPVT raw scores ($p = .257$). DPNrt did not interact with group ($p = .156$) indicating that this relationship did not vary across the three groups. The EVT and PPVT did not predict the phonotactic frequency effect when group was included in

the analysis ($p = .661$ and $p = .161$, respectively). The results of the hierarchical regression are reported in Appendix D3.

Predicting Picture-Naming Reaction Time

The PPVT raw score significantly predicted the PNrt, $t(51) = 6.45$, $p < .001$, $sr^2 = .313$, as did EVT raw score, $t(51) = 4.97$, $p < .001$, $sr^2 = .281$, and chronological age, $t(51) = 7.43$, $p < .001$, $sr^2 = .417$.

GFTA raw score predicting picture-naming reaction time. The GFTA raw score was not a significant predictor of the PNrt after controlling for EVT and PPVT raw scores and chronological age ($p = .093$). GFTA did not interact with group ($p = .242$) indicating that this relationship did not vary across the three groups. Chronological age was the only covariate that predicted the PNrt time on each of the three steps (Step 1: $p = .001$; Step 2: $p = .021$; Step 3: $p = .029$). The results of the hierarchical regression are reported in Appendix D4.

CAS trait score predicting picture-naming reaction time. The CAS trait score was not a significant predictor of the picture-naming reaction time after controlling for EVT and PPVT raw scores and chronological age ($p = .110$). However, the regression analysis revealed that it was a significant predictor at step 2 and 3 of the analysis. As can be seen from the correlation table the CAS trait score did not correlate with picture-naming reaction and therefore these significant results could be due to suppression effects. Suppression effects occur in regression when the relationship between the IVs is stronger than the relationship between the IV and the DV, therefore enhancing the likelihood of a significant result. The CAS trait score did not interact with group ($p = .092$) indicating that this relationship did not vary across the three groups. Chronological age was the only covariate that predicted the picture-naming reaction time on each of the three steps (Step 1: $p = .001$; Step 2: $p = .036$; Step 3: $p = .021$). Refer to Table 46.

DPNrt predicting picture-naming reaction time. The DPNrt was not a significant predictor of picture-naming reaction time after controlling for EVT and PPVT raw scores and chronological age ($p = .969$), although it did interact with group ($p = .020$), indicating that this relationship differed across the three groups.

Table 46

Unstandardized (B) and Standardized (β) Regression Coefficients, and Squared Part Correlations (sr^2) for a Hierarchical Multiple Regression Model Predicting Picture-Naming Reaction Time from CAS Trait Score (CASts), Group, and the Group x CASts Interaction (N = 53)

Predictors (IVs)	B	95% CI ¹	B	sr^2	p-value ¹
DV: PNrt					
Step 1					
EVT raw	-2.357	-11.281, 6.568	-.109	.003	.598
PPVT raw	-2.063	-6.848, 2.722	-.178	.006	.390
Chron Age	-9.451	-14.943, -3.960	-.481	.096	.001**
CASts	-48.472	-108.346, 11.403	-.178	.025	.110
$R^2 = .456, p = .000^{**}$					
Step 2					
EVT raw	-1.672	-9.828, 6.483	-.078	.001	.682
PPVT raw	-2.017	-7.417, 3.383	-.174	.005	.456
Chron Age	-10.215	-19.706, -0.724	-.520	.067	.036*
CASts	-109.410	-207.175, -11.644	-.402	.031	.029*
Group (D1) ²	170.713	-49.771, 391.197	.296	.014	.126
Group (D2) ²	65.930	-138.773, 270.633	.127	.006	.520
Group: $F(2, 46) = 1.3, p = .290$					
$\Delta R^2 = .014, p = .540$					
$R^2 = .470, p = .000^{**}$					
Step 3					
EVT raw	-1.624	-10.306, 7.057	-.075	.001	.708
PPVT raw	-1.342	-6.839, 4.156	-.116	.002	.625
Chron Age	-10.959	-20.188, -1.731	-.558	.075	.021*
CASts	-322.992	-526.726, -119.258	-1.19	.025	.003*
Group (D1) ²	329.691	114.577, 544.805	.573	.027	.003*
Group (D2) ²	258.810	16.391, 501.228	.499	.019	.037*
Group: $F(2, 44) = 4.81, p = .013^3$					
Group (D1)xCASts ⁴	227.828	-8.237, 463.894	.613	.010	.058
Group (D2)xCASts ⁴	276.238	-25.079, 577.554	.249	.013	.071
Group x CASts: $F(2, 44) = 2.53, p = .092^5$					
$\Delta R^2 = .013, p = .576$					
$R^2 = .483, p = .000^{**}$					

1: These are the GLMM adjusted values.

2: These are dummy variables; in combination, they reflect the group effect.

3: This is the overall F -value for the group effect.

4: The CAS trait score is already centered.

5: This is the overall F -value for the Group x CAS trait score interaction effect.

Additional regressions were undertaken to explore the significant interaction between age group and predictor for each of the groups. The DPNrt was not a significant predictor for the children with CAS, $F(1,9) = 3.18, p = .108$, partial $\eta^2 = .29$; the children with PD, $F(1,16) = 4.01, p = .063$, partial $\eta^2 = .66$, or TD, $F(1,13) = 4.27, p = .059$, partial $\eta^2 = .26$, although the children with PD and TD almost reached significance. The Pearson correlations reflected this pattern also with the DPNrt

correlating with the PNrt for the children with PD and TD, but not for the children with CAS. Chronological age predicted the mean picture-naming reaction time on each of the three steps (Step 1: $p = .001$; Step 2: $p = .041$; Step 3: $p = .023$). Refer to Table 47.

Table 47

Unstandardized (B) and standardized (β) Regression Coefficients, and Squared Part Correlations (sr^2) for a Hierarchical Multiple Regression Model Predicting Picture-naming Reaction Time from DPNrt, Group, and the Group x DPNrt Interaction ($N = 53$)

Predictors (IVs)	B	95% CI ¹	B	sr^2	p-value ¹
DV: PNrt					
Step 1					
EVT raw	-1.670	-10.649, 7.309	-.078	.001	.710
PPVT raw	-1.207	-6.258, 3.844	-.104	.002	.633
Chron Age	-10.019	-15.934, -4.103	-.510	.106	.001**
DPNrt	0.006	-0.322, 0.335	.005	.000	.969
$R^2 = .431, p = .000^{**}$					
Step 2					
EVT raw	-1.377	-9.842, 7.088	-.064	.001	.745
PPVT raw	-1.190	-6.830, 4.450	-.103	.002	.673
Chron Age	-10.586	-20.710, -0.462	-.539	.068	.041*
DPNrt	0.048	-0.302, 0.398	.038	.001	.785
Group (D1) ²	-51.398	-200.921, 98.124	-.089	.004	.492
Group (D2) ²	15.502	-190.656, 221.660	.030	.000	.880
Group: $F(2, 46) = 0.41, p = .668$					
$\Delta R^2 = .009, p = .682$					
$R^2 = .440, p = .000^{**}$					
Step 3					
EVT raw	-0.488	-8.504, 7.527	-.023	.000	.903
PPVT raw	-0.535	-5.952, 4.883	-.046	.000	.843
Chron Age	-11.138	-20.688, -1.588	-.567	.074	.023*
DPNrt	0.665	-0.037, 1.368	.537	.026	.063
Group (D1) ²	-43.196	-202.983, 116.592	-.075	.003	.589
Group (D2) ²	-14.856	-221.689, 191.977	-.029	.000	.886
Group: $F(2, 44) = 0.17, p = .842^3$					
Group (D1)xDPNrt ⁴	-186.187	-337.948, -34.427	-.531	.041	.017*
Group (D2)xDPNrt ⁴	-55.577	-207.852, 96.698	-.111	.003	.466
Group x DPNrt : $F(2, 44) = 4.30, p = .020^5$					
$\Delta R^2 = .064, p = .069$					
$R^2 = .504, p = .000^{**}$					

1: These are the GLMM adjusted values.

2: These are dummy variables; in combination, they reflect the group effect.

3: This is the overall F -value for the group effect.

4: The interaction term is computed using centered DPNrts.

5: This is the overall F -value for the Group x DPNrt interaction effect.

Predicting Speech Discrimination Ability

Speech discrimination ability as measured by d-prime for the ungated condition was analyzed to determine significant predictors. The PPVT raw score significantly predicted speech discrimination ability, $t(51) = 5.42, p < .001, sr^2 = .328$, as did EVT raw score, $t(51) = 5.58, p < .001, sr^2 = .333$ and chronological age, $t(51) = 6.04, p < .001, sr^2 = .291$.

GFTA raw score predicting d-prime. The GFTA raw score was a significant predictor of d-prime after controlling for EVT and PPVT raw scores and chronological age ($p = .048$). GFTA did not interact with group ($p = .978$) indicating that this relationship did not vary across the three groups. Chronological age predicted d-prime at step 1 in the regression analysis (Step 1: $p = .046$), indicating that as chronological age increased so did d-prime. Refer to Table 48.

CAS trait score predicting d-prime. The CAS trait score was not a significant predictor of d-prime after controlling for EVT and PPVT raw scores and chronological age ($p = .408$). However, the CAS trait score interacted with group ($p = .038$) indicating that this relationship varied across the three groups. Further analysis revealed that the CAS trait score was not a significant predictor for the children with CAS, $F(1,9) = 0.72, p = .417$, partial $\eta^2 = .105$, or for the children PD, $F(1,16) = 0.01, p = .923$, partial $\eta^2 = .025$, but it was for the children with TD; $F(1,13) = 4.86, p = .046$, partial $\eta^2 = .043$. None of the covariates predicted speech discrimination ability, chronological age just missed out on reaching significance for step 1 of the regression analysis (Step 1: $p = .053$). Refer to Table 49

DPNrt predicting d-prime. The final regression model tested the DPNrt as a predictor of d-prime. The DPNrt was not a significant predictor of speech discrimination ability after controlling for EVT and PPVT raw scores and chronological age ($p = .664$). DPNrt did not interact with group ($p = .341$) indicating that this relationship did not vary across the three groups. None of the covariates were significant predictors of speech discrimination ability. The results of the hierarchical regression are reported in in Appendix D5.

Table 48

Unstandardized (B) and standardized (β) Regression Coefficients, and Squared Part Correlations (sr^2) for a Hierarchical Multiple Regression Model Predicting Speech Discrimination Ability (d-prime) from GFTA Raw Scores, Group, and the Group x GFTA Interaction (N = 53)

Predictors (IVs)	B	95% CI ¹	B	sr^2	p-value ¹
DV: D-prime					
Step 1					
EVT raw	0.014	-0.017, 0.046	.179	.008	.371
PPVT raw	0.006	-0.014, 0.025	.137	.004	.550
Chron Age	0.018	0.000, 0.035	.248	.026	.046*
GFTA raw	-0.016	-0.032, -0.000	-.261	.060	.048*
$R^2 = .418, p = .000^{**}$					
Step 2					
EVT raw	0.017	-0.014, 0.048	.220	.010	.263
PPVT raw	0.007	-0.014, 0.028	.167	.005	.503
Chron Age	0.012	-0.018, 0.043	.172	.007	.413
GFTA raw	-0.016	-0.046, 0.014	-.260	.015	.294
Group (D1) ²	0.036	-1.156, 1.227	.017	.000	.952
Group (D2) ²	0.167	-0.612, 0.947	.088	.002	.668
Group: $F(2, 46) = 0.16, p = .852$					
$\Delta R^2 = .004, p = .856$					
$R^2 = .422, p = .000^{**}$					
Step 3					
EVT raw	0.017	-0.015, 0.049	.215	.009	.293
PPVT raw	0.007	-0.014, 0.028	.168	.005	.502
Chron Age	0.013	-0.016, 0.043	.183	.008	.372
GFTA raw	-0.016	-0.057, 0.025	-.264	.002	.436
Group (D1) ²	0.103	-1.417, 1.623	.049	.000	.892
Group (D2) ²	0.167	-0.657, 0.991	.088	.001	.685
Group: $F(2, 44) = 0.08, p = .920^3$					
Group (D1)xGFTA ⁴	-0.047	-1.280, 1.187	-.033	.000	.940
Group (D2)xGFTA ⁴	0.051	-0.696, 0.797	.019	.000	.892
Group x GFTA: $F(2, 44) = 0.02, p = .978^5$					
$\Delta R^2 = .001, p = .979$					
$R^2 = .422, p = .001^{**}$					

1: These are the GLMM adjusted values.

2: These are dummy variables; in combination, they reflect the group effect.

3: This is the overall F -value for the group effect.

4: The interaction term is computed using centered GFTA raw scores.

5: This is the overall F -value for the Group x GFTA interaction effect.

* $p < .05$; ** $p < .01$

Table 49

Unstandardized (B) and standardized (β) Regression Coefficients, and Squared Part Correlations (sr^2) for a Hierarchical Multiple Regression Model Predicting Speech Discrimination Ability (d -prime) from CAS Trait Score (CAS_{ts}), Group, and the Group x CAS_{ts} Interaction (N = 53)

Predictors (IVs)	B	95% CI ¹	B	sr^2	p-value ¹
DV: D-prime					
Step 1					
EVT raw	0.017	-0.016, 0.050	.214	.011	.302
PPVT raw	0.006	-0.012, 0.024	.145	.004	.494
Chron Age	0.017	0.000, 0.034	.236	.023	.053
CAS _{ts}	-0.129	-0.441, 0.182	-.129	.013	.408
$R^2 = .371, p = .000^{**}$					
Step 2					
EVT raw	0.017	-0.017, 0.052	.219	.010	.314
PPVT raw	0.008	-0.013, 0.029	.193	.007	.436
Chron Age	0.014	-0.010, 0.037	.192	.009	.240
CAS _{ts}	0.336	-0.188, 0.860	.336	.022	.203
Group (D1) ²	-1.182	-2.371, 0.008	-.558	.051	.052
Group (D2) ²	-0.219	-0.795, 0.357	-.115	.005	.448
Group: $F(2, 46) = 2.07, p = .138$					
$\Delta R^2 = .055, p = .121$					
$R^2 = .426, p = .000^{**}$					
Step 3					
EVT raw	0.009	-0.028, 0.045	.109	.002	.634
PPVT raw	0.013	-0.009, 0.035	.303	.016	.248
Chron Age	0.013	-0.007, 0.032	.178	.008	.193
CAS _{ts}	-0.624	-1.486, 0.238	-.624	.007	.152
Group (D1) ²	-0.863	-1.923, 0.198	-.408	.014	.108
Group (D2) ²	0.398	-0.411, 1.207	.209	.003	.327
Group: $F(2, 44) = 2.85, p = .069^3$					
Group (D1)xCAS _{ts} ⁴	1.278	0.199, 2.356	.936	.025	.021*
Group (D2)xCAS _{ts} ⁴	0.395	-0.695, 1.485	.097	.002	.469
Group x CAS _{ts} : $F(2, 44) = 3.53, p = .038^5$					
$\Delta R^2 = .043, p = .179$					
$R^2 = .469, p = .000^{**}$					

1: These are the GLMM adjusted values.

2: These are dummy variables; in combination, they reflect the group effect.

3: This is the overall F -value for the group effect.

4: The CAS trait score is already centered.

5: This is the overall F -value for the Group x CAS trait score interaction effect.

Discussion

The main purpose of Study 2 was to investigate whether there is a predictive relationship between measures of speech motor ability and measures of phonological competence in children with CAS, PD and TD. We specifically wanted to determine if the groups differed in the nature or strength of this relationship. We hypothesized that the children with CAS would show the strongest predictive relationship between

measures of speech motor ability and phonological competence, compared to children with PD and TD, consistent with the developmental account of CAS (Maassen, 2002), whereby motor deficits constrain the emergence of higher-level linguistic processes, such as phonological representations. Prior to doing this we needed to ascertain the extent of shared deficits in children with CAS, PD and TD by implementing tasks that targeted phonological competence and speech motor ability.

Phonological Deficits in CAS and PD

As expected, children with CAS and PD shared phonological deficits, although, overall, the children with CAS presented with more severe deficits than the children with PD. The children with CAS were significantly poorer than the children with TD and PD in NWR accuracy, averaged across frequency and length. We assumed the children with CAS and PD have underspecified phonological representations and consequently, would be less efficient retrieving the required phonological units for the nonwords, resulting in poorer overall accuracy. Recent research has shown that speech motor ability predicts nonword repetition accuracy, suggesting that nonword repetition is not only a measure of phonological encoding processes but also an index of speech motor control (Krishnan et al., 2013). This would suggest that the children with CAS would be further disadvantaged repeating nonwords due to their underlying speech motor deficit. Although the children with PD were less accurate than the TD children the difference was not significant, as hypothesized. Previous research found that children with PD performed significantly poorer than the children with TD (Munson, Edwards, et al., 2005b), however, the children with PD in Munson et al.'s study had a lower mean percentile ranking ($M = 5\%$, $SD = 2$), compared to the children with PD in the current study ($M = 13\%$, $SD = 10$), indicating a more severe speech deficit than the children with PD in the current study and more comparable to the children with CAS in the current study ($M = 5\%$, $SD = 5$). Therefore, it is plausible that the children with milder speech deficits in our PD group inflated the overall mean accuracy for this group. Low statistical power may have also comprised our findings.

Consistent with previous research, the high frequency sequences were repeated more accurately than low frequency sequences and in keeping with the view that sound sequences that are high in phonotactic frequency (i.e., the sound sequences occur in more words) are acquired and accessed more easily than sound sequences

that are low in phonotactic frequency (Beckman & Edwards, 2000; Coady & Aslin, 2004). However, the groups did not differ significantly in accuracy between the low and high phonotactic frequency sequences, as hypothesized. In fact, the children with CAS were the least sensitive to the effect of frequency compared to the children with PD and TD. Munson et al. (2005b) also found that the children with PD in their study were less sensitive to frequency than the children with TD, contrary to what they had predicted. They predicted that the children with PD would perform more similar to younger children with TD, consistent with the assumption that younger children have more high frequency syllables readily available in their repertoire, compared to low frequency sequences, and therefore nonwords with high frequency sequences would be repeated with greater accuracy compared to nonwords with low frequency sequences. Given the children with PD in Munson et al.'s (2005b) study had a lower percentile ranking on the GFTA than the children with PD in our study and were more similar in relation to their percentile rank to the children with CAS in the current study, then it is not surprising that the children with CAS were least sensitive to the phonotactic frequency effect, consistent with Munson et al.'s (2005b) findings.

The children with PD had a numerically larger frequency effect from the children with CAS and TD, although the difference in the magnitude was not significantly different, it may reflect the variability in severity of deficits of the children in this group. As previously stated, the children with PD included children with mild and severe speech deficits, consequently, the larger frequency effect observed for children with PD may in fact be due to the children with the milder speech deficits demonstrating enhanced performance for the high frequency sequences, thereby potentially enhancing the overall performance of this group and resulting in an elevated phonotactic frequency effect for the children with PD as a group. Munson et al. (2005b) investigated why the children with PD in their study did not demonstrate a larger frequency effect than the children with TD, as they had hypothesized. To do this they split the children with PD into a mild and a severe group. They found that the frequency effect was smaller in children with severe PD compared to children with mild PD (Munson, Edwards, et al., 2005b). This suggests that it is plausible that children with mild PD, who have attained more abstracted phonological representations (i.e., they have developed some degree of distinct

phonemic categories) compared to children with severe PD, are more likely to demonstrate an effect of frequency, performing better for the high frequency sequences compared to the low frequency sequences. In contrast, for children with severe PD, and children with CAS, it is therefore plausible that these children did not demonstrate an advantage for the high frequency sequences (compared to the low frequency sequences), thus diminishing the likelihood of finding a phonotactic frequency effect, because they have not yet attained a sufficient degree of phonemic development, regardless of phonotactic frequency.

Munson, Kurtz and Windsor (2005) explored NWR accuracy in children with SLI and TD. They found that the children with SLI were significantly less accurate than their TD age matched peers but they did not demonstrate a particular disadvantage for the low frequency sequences, consistent with Munson et al.'s (2005b) findings. Munson et al. (2005) proposed that children with SLI have difficulty with word learning and this may in turn have a consequence of difficulties abstracting phonological representations, which is why they did not demonstrate a larger frequency effect. For a child to benefit in repeating high frequency sequences, compared to low frequency sequences, then they must have acquired a repertoire of high frequency sequences, or distinct phonemic categories that can be readily assembled in the phonological buffer, ready for execution. Consequently, for children with a severe speech deficit, such as CAS or severe PD, it's plausible these distinct phonemic categories have not been established even for high frequency sequences and therefore performance is not noticeably enhanced for the high frequency sequences compared to low frequency sequences.

The groups did not differ significantly in relation to picture-naming reaction time. We did however demonstrate a robust priming effect, with the related prime having the fastest reaction time, followed by the silence condition and the unrelated prime, consistent with previous findings (Brooks & Mac Whinney, 2000; Jerger et al., 2002). We hypothesized that the children with CAS and PD would have a slower reaction time than the children with TD. This was based on the rationale that these children have weaker output phonological encoding abilities, although this was not demonstrated. Either the children with CAS and PD do not have a deficit at this level of processing, or the nature of the deficit in output phonological does not impact on the efficiency of phonological encoding, or the task failed to pick up on

these differences as a result of design issues. The fact that the children with CAS were numerically the fastest was surprising, however, the children with CAS had also the lowest accuracy score and so it is plausible that they traded accuracy for speed, especially since all the participants named the pictures correctly prior to test trials.

The main purpose of the picture-naming task was to explore phonological priming effects as a means of assessing output phonological representations and phonological encoding efficiency. However, since no difference between the groups was demonstrated in relation to priming effects, it may be that the groups did not differ in phonological encoding, or there is a deficit that does not impact on encoding efficiency, or the task was not sensitive enough to detect differences in encoding efficiency. There is conflicting evidence in research pertaining to phonological facilitation effects in children with speech and language deficits. Phonological facilitation effects have been well documented in the literature in both adults (Schriefers et al., 1990) and children (Brooks & MacWhinney, 2000; Jerger et al., 2002), however, research in phonological facilitation effects is not as extensive in children with speech and/or language deficits. Children with dyslexia have demonstrated enhanced phonological facilitation, compared to TD peers, indicative of deficient output phonological representations (Truman & Hennessey, 2006), whereas children with SLI demonstrated facilitation effects for onsets but not offset/rime auditory primes (Seiger-Gardner & Brooks, 2008). The performance of the children with SLI is similar to the 5 to 7 years olds in Brooks and MacWhinney's (2000) study, whereby the younger children showed onset competition at the early SOA and rime/offset priming at the late SOA. However, the children with SLI did not demonstrate a rime or offset-based priming effect for the late SOA, unlike the 5 to 7 year olds in Brooks and MacWhinney's (2000) study. The absence of a rime priming effect in the children with SLI suggests that children with SLI do not perform like younger children with TD, who are assumed to have less abstracted phonemic categories and slower phonological encoding processes. More recently Munson and Krause (2017) explored phonological encoding ability in children with SSD using a cross modal priming experiment, similar to the task used in the current study. They found that the size of the phonological facilitation effect was similar for children with SSD and TD and concluded that SSD is not associated with reduced

phonological encoding ability. Given the findings in the current study with children with CAS and PD and the recent research in children with SSD, further research is needed to determine whether in fact children with speech and language deficits perform similar to younger children with TD or if they are unique in their performance. But since the findings replicate Munson and Krause's (2017) study showing similar priming for CAS and PD compared to TD children, the results provide tentative support for the conclusion that phonological encoding is not less efficient in children with both CAS and PD.

The analysis for the speech discrimination task used d-prime as the dependent variable. We replicated previous findings in that the groups did not differ on the two conditions with the least acoustic information (i.e., the short gate condition and the long gate condition) demonstrated by all groups performing at a very low level in these conditions (Edwards et al., 1999). In keeping with Edwards et al. (1999) we analysed the data for the whole word condition only and found that the children with CAS were significantly less accurate than the children with TD and less accurate, but not significantly, than the children with PD. The fact that children with CAS and PD demonstrated a diminished ability to accurately identify the words spoken for the whole word condition suggests that these children may have weaker or less well specified input phonological representations, which may be a characteristic feature of both CAS and PD, and potentially associated with a more severe speech deficit. This is indicative of a tight coupling between speech production and speech discrimination abilities for these children, and in keeping with previous research that has demonstrated speech perception deficits in CAS (Nijland, 2009) and PD (Edwards et al., 1999; Edwards et al., 2002). These findings are also consistent with developmental models that propose a strong relationship between articulatory and acoustic maps that work symbiotically to enhance speech development (Westerman & Miranda, 2004). Nijland (2009) found that children with CAS had diminished ability to discriminate between words that differed in one consonant (initial or final) or stimuli that were metatheses of each other (e.g., sut and tus). They also found that children with CAS had difficulty identifying rhymes and discriminating rhymes, in contrast to the children with PD in their study, who had difficulty identifying rhymes only (with discrimination skills intact). Nijland (2009) proposed the deficits in both discrimination and rhyming ability in the children with CAS was indicative of a

strong association between perception and production for these children. This proposition is further evaluated in the discussion relating to the regression analysis and specifically in relation to the relationship between speech discrimination ability and speech motor measures.

In relation to the speech motor measures, the delayed picture-naming reaction time (DPNrt) revealed a significant difference between groups. The children with CAS and PD were significantly slower than the children with TD, consistent with our hypothesis, but the difference in reaction time between children with CAS and PD was not significant. Given the different assumed etiologies in children with CAS and PD, we predicted that the children with CAS would be significantly slower than the children with PD. However, even though the difference in DPNrt between CAS and PD was not significant the reaction times were in the predicted direction. The difference in reaction between the children with CAS and PD was marginal (76ms), with a small effect size (.024), compared to the difference in reaction time between PD and TD, which was larger (120ms), with a large effect size (.159). The fact that the effect size was small compromised the likelihood of us finding a significant difference between the children with CAS and PD based on the lower statistical power associated with the smaller sample size for the CAS group. Furthermore, the children with PD in the current study varied in articulatory ability` and it is therefore plausible that some of the children with PD in our cohort with a more severe speech deficit also had speech motor deficits.

Relationship Between Speech Motor Measures and Phonological Competence

There were a number of correlations that emerged that differentiated between the groups. As a whole group the motor measures correlated with a number of the measures of phonological competency, however, when we looked at the correlations for the individual groups a different pattern of association emerged indicative of different relationships between the levels of processing investigated, although the bivariate correlations do not control for other confounding variables. The GFTA raw score correlated with nonword repetition accuracy for the high frequency sequences for the children with CAS and PD but not for the children with TD. This was not surprising given that both the children with CAS and PD had articulatory difficulties that potentially constrain and therefore influence higher level processing skills required for nonword repetition. The CAS trait score, on the other hand, correlated

with nonword repetition accuracy for the low frequency sequences for the children with CAS and PD. This was as expected since it is reasonable to assume that the traits associated with this score (speech motor deficits) are more likely to impact on phonotactic sequences that are low in frequency as these sequences are less likely to be attained by these children and less robust compared to high frequency sequences. What was more noteworthy was that the delayed picture-naming reaction time correlated with the phonotactic frequency effect for the children with CAS, but not for the children with PD and TD, differentiating children with CAS and PD and suggesting that speech motor ability influences the phonotactic frequency effect in children with CAS, but not PD or TD. Furthermore, DPNrt also correlated with PNrt for the children with PD and TD, but not for the children with CAS. This correlation between these measures per say is not that surprising given the number of processes these tasks share, however, it is of note that these measures did not correlate for the children with CAS indicating a point of difference that warrants further investigation, and is further examined in the regression analysis discussion. Another point of difference was the GFTA correlated with picture-naming reaction time for the children with CAS only, suggesting that articulatory ability (when severe, as is the case for the children with CAS) influences picture-naming reaction time for the children with CAS but not for the children with PD and TD.

When we looked at the correlations between the measures of phonological competence a different pattern of associations emerged for the different groups. For the children with CAS and PD there were a number of similar associations, which was not surprising given the degree of shared deficits, however, there were also a number of differences. The phonotactic frequency effect did not correlate with any of the other measures of phonological competence for the children with CAS, whereas for the children with PD and TD it correlated with NWR accuracy for the low frequency sequences. The implication being that there is a close relationship between these measures for the children with PD and TD but not for the children with CAS. This difference between the children with CAS compared to the children with PD and TD highlights the smaller and non-significant frequency effect for the children with CAS, which potentially diminished the relationship with NWR accuracy observed for the other two groups. If we assume that children with CAS have a core speech motor deficit that influences all other levels of processing then we

would expect dissociation between these measures. Whereas, for children with PD, if we assume these children have a core phonological deficit then this finding is as expected, indicative of a tight coupling between measures that tap into the same level of processing. The NWR accuracy measures (both high and low frequency sequences) were also positively correlated with the picture-naming reaction time for the children with CAS, whereas, for the children with TD they were negatively correlated and there was no correlation for the children with PD. d' -prime was positively correlated NWR accuracy for the low frequency sequences for the children with PD and TD, whereas there was no correlation for the children with CAS. Also, d' -prime correlated with the phonological facilitation effect for the children with CAS, but there was no correlation for the children with PD or TD. These points of difference between the children with CAS and PD warrant further investigation and highlight some significant differences in the constraints that can potentially emerge in speech and language development in these children.

Regression Analysis

To investigate these relationships further we used regression analysis to determine if speech motor measures predicted measures of phonological competency while controlling for the confounding effects of age and vocabulary. The primary predictors included the GFTA raw score, the CAS trait score (derived from the factor analysis with the trait score reflecting severity of CAS features) and the mean reaction time from the DPNrt. Overall the GFTA raw score was a good predictor of measures of phonological competence, however, it did not interact with group for any of these measures, indicating that this relationship did not vary with group. The GFTA raw score predicted nonword repetition accuracy after accounting for all significant covariates (i.e., EVT, PPVT and age), indicating that it is a robust predictor of NWR accuracy, accounting for an additional 29% of variance (refer to Table 41, step 1 of the regression analysis). This suggests that, for the whole group, the development of robust phonemic categories, and therefore greater accuracy at repeating nonwords with low phonotactic frequency sequences, is dependent somewhat on articulatory ability. However given the interaction between GFTA and group was not significant this indicates that the groups did not differ in this relationship, suggesting articulatory ability predicted NWR accuracy to the same extent in the children with CAS, PD and TD. Munson et al. (2005b) found that the

GFTA predicted variance in mean accuracy for the entire group and for the children with PD, but not the children with TD, when analysed as separate groups. In our study, the correlation between the GFTA and mean accuracy for both the high the low frequency sequences suggested a strong negative association for the entire group. There was a strong correlation between mean nonword repetition accuracy for the high frequency sequences and the GFTA for the children with CAS (-.602**) and PD (-.448*), but not for the children with TD (-.198) suggesting that articulatory ability is associated with repetition accuracy when articulation is compromised. NWR with high frequency sequences is sensitive to the degree of phonological development among children who have weaker phonological and articulatory skills, such as children with CAS and PD. Performance for children with TD may be close to ceiling for the words containing high frequency sequences and therefore this limits the likelihood of finding a significant correlation for these children. This finding was further supported by the larger correlation between GFTA and NWR for the low frequency sequences compared to the high frequency sequences for the children with TD. This suggests that the low frequency sequences were more sensitive to relationships with other variables, whereas the high frequency sequences were more robust and therefore accuracy was not as related to skills at other levels of processing, such as articulatory ability for the children with TD.

The GFTA was also a significant predictor of the phonotactic frequency effect, after accounting for vocabulary, indicating that articulatory ability contributes unique variance in the phonotactic frequency effect. This shows that there is a clear and unique relationship between articulatory ability and the quality of underlying phonemic categories, independent of age and beyond that accounted for by vocabulary. The negative slope in the regression model for GFTA shows that the frequency effect is weaker for those with poor articulation ability. This relationship however was not well supported by the bivariate correlations, no correlations were significant between the GFTA and the frequency effect, either for the entire group or for the individual groups, which suggests some caution that suppression effects be involved. The interaction between the GFTA and group was not significant, indicating that the relationship between the GFTA and the phonotactic frequency effect did not vary between the groups, contrary to our hypothesis and in contrast to Munson et al.'s findings (2005b). Munson et al. (2005b) found that the GFTA did

not predict the phonotactic frequency effect after controlling for vocabulary when children with TD and PD were included in their analysis, however when they analysed groups separately the GFTA was a significant predictor for the frequency effect for the children with PD but not for the children with TD, suggesting that articulatory ability only predicted the frequency effect for children with compromised articulatory ability. The negative direction of the relationship was similar to the present study, those with poorer articulation ability showed reduced frequency effects. As previously stated, the children with PD in the current sample included children with mild and severe speech deficits, whereas the children in Munson et al.'s (2005b) study included children with severe PD, indicated by the mean percentile ranking on the GFTA ($M = 5\%$). Consequently, the GFTA may only predict the frequency effect for children with severe speech deficits.

The GFTA did not predict picture-naming reaction time or interact with group indicating that there is no significant association between these outcome measures. However, the correlations showed that the GFTA was negatively correlated with picture-naming reaction time for the children with CAS ($-.669^{**}$) only, indicating that the efficiency at this level of processing was correlated with articulatory ability for these children but not for the children with PD and TD. Picture-naming involves a number of different stages one of which is articulation, consequently, it is not surprising that the GFTA correlated with this measure. However, the fact that it only correlated with GFTA for the children with CAS is of interest and warrants further investigation.

The GFTA was a significant predictor of speech discrimination ability after controlling for vocabulary and age, accounting for an additional 6% of variance in speech discrimination ability, consistent with previous findings (Edwards et al., 2002). However, the interaction between the GFTA and group was not significant and the absence of any significant correlation between the GFTA and speech discrimination ability for the groups individually reinforced this finding. This indicates that the GFTA did not vary in its ability to predict speech discrimination for the different groups.

Overall the CAS trait score predicted a number of the measures of phonological competence, but failed to interact with group, other than for speech discrimination ability. The CAS trait score was a significant predictor of mean

repetition accuracy for the low frequency sequences for the entire group after controlling for vocabulary and age. However, the interaction between the CAS trait score and group was not significant, indicating that the groups did not differ in the relationship between the CAS trait score on NWR accuracy. The correlations revealed a significant negative correlation between the CAS trait score and nonword repetition accuracy for the low frequency sequences for the children with CAS and PD, suggesting that the severity of deficits associated with CAS (i.e., the higher the CAS trait score) is related to performance on nonword repetition accuracy for children with CAS and PD. The marginal difference in the strength of the correlation ($-.564^*$ vs. $-.530^*$) between the children with CAS and PD suggests that the CAS trait score did not predict NW accuracy to a greater extent in CAS than in PD, contrary to our hypothesis. The CAS trait score was not a significant predictor of the phonotactic frequency effect and it did not interact with group. The CAS trait score did not predict picture-naming reaction time after controlling for vocabulary and age, however it predicted picture-naming reaction time at step 2 and 3 of the regression analysis. This finding was not supported by the correlations, for the whole group or the individual groups and could be the result of suppression effects, which need to be interpreted with caution. It was nevertheless of interest that the correlation between the picture-naming reaction time and the CAS trait score were negative for the children with CAS and TD but positive for the children with PD. The positive correlation is in the expected direction indicating the higher the CAS trait score (i.e., the more severe the speech deficit) then the higher the reaction time (reflecting less efficient processing).

The CAS trait score did not predict speech discrimination ability, however, it interacted with group, indicating that the relationship between the CAS trait score and d-prime varied as a function of group. Further analysis of the interaction revealed that the CAS trait score predicted d-prime for the children with TD, but not for the children with CAS and PD. This unexpected result may be due to suppression effects given the children with TD did not have a CAS trait score as they did not present with any of the features associated with CAS. The absence of any significant correlations between the CAS trait score and d-prime for the individual groups confirms there was no association between these measures for any of the groups after accounting for the covariates. An alternative explanation for this

finding is that the regression analysis may be suspect given the reduced variability in the CAS trait score for the TD group in particular.

Overall the delayed picture-naming reaction time (DPNrt) did not predict measures of phonological competence, other than for nonword repetition accuracy. On the whole the DPNrt did not interact with group, except for picture-naming reaction time. Further analysis of this interaction revealed that the DPNrt was close to being statistically significant for the children with PD and TD ($p = .063$, $p = .059$, respectively) but not for the children with CAS ($p = .108$). The method of follow up analysis is likely to have reduced statistical power, compared to the overall analysis, because of the reduced sample size that follows from conducting the regressions one group at a time. The bivariate correlations support this interpretation of the interaction as well. The DPNrt did not predict picture-naming reaction time but it did interact with group suggesting the groups varied in this relationship. The DPNrt was positively correlated with the picture-naming reaction time for the children with PD (.543*) and TD (.618**) but not for the children with CAS (.099). The correlation is not surprising given both tasks are picture-naming and therefore have a number of shared processes, such as object recognition, word retrieval etc. The reaction time for each individual may reflect a participant specific factor such as speed of processing, giving rise to an association. It suggests that the level of efficiency in processing during object recognition and word retrieval are related to the speed or efficiency of processing in a similar naming task that is sensitive to motor execution processes. However, in the context of the WEAVER framework the DPNrt does minimize the involvement of some stages of processing, such as word form encoding and phonetic encoding and in this study was used as a measure of speech motor programming and response preparation and execution (Kawamoto et al., 2008). Consequently, it is therefore noteworthy that these two measures did not correlate for the children with CAS, whereas they did for the children with PD and TD. These results suggest there is dissociation between performance on the PNrt task and the DPNrt task for the children with CAS, whereas there is a clear association for the children with PD and TD. This warrants further investigation and is further discussed in the General Discussion in Chapter 6.

In addition to determining whether speech motor measures predicted measures of phonological competence we also wanted to look at vocabulary as a predictor.

Overall, vocabulary predicted measures of phonological competence, consistent with previous research and in keeping with the understanding that children gradually develop abstract phonemic representations or more detailed or well-specified phonological representations as a consequence of learning more words. The PPVT and EVT raw scores were significant predictors for the nonword repetition accuracy for the low frequency sequences, in keeping with the previous findings (Edwards et al., 2004; Munson, Edwards, et al., 2005b). The PPVT and EVT were significant predictors of the phonotactic frequency effect, consistent with previous findings (Edwards et al., 2004; Metsala, 1999; Munson, Edwards, et al., 2005b; Munson, Kurtz, et al., 2005), and indicating that vocabulary size has greater influence over the difference in repetition accuracy between high and low frequency sequences, than age. The PPVT, EVT and chronological age all predicted a substantial portion of picture-naming reaction time, which was an expected finding given the assumption that picture-naming is a lexical task, sensitive to lexical skills and therefore relies heavily on vocabulary. The PPVT, EVT and chronological age all accounted for unique variance in speech discrimination ability. Edwards et al. (2002) found that the PPVT accounted for a substantial portion of the variance in speech discrimination ability. In the current study the PPVT and EVT accounted for 33% of variance in speech discrimination ability, and chronological age accounted for an additional 29%, consistent with the view that children with large vocabularies perform better on speech perception tasks (Edwards et al., 1999; Edwards et al., 2002; Munson, 2001b).

Limitations and Future Directions

A major limitation of this study was the sample size, especially after redefining groups, which not only reduced the number of CAS participants but also made the groups unequal in size. The cohort of children with PD included children with mild and severe speech deficits, consequently the mean performance of the children on all tasks was potentially higher than if the sample was all severe PD children, as in Munson et al.'s (2005b) study. Severity of deficit has been shown to play a role in whether potential predictors account for unique variance in an outcome measure. For example, Munson et al. (2005) found that the GFTA did not predict the phonotactic frequency effect when both children with PD and TD were included in the analysis, but when they redid the analysis with just the children with PD the

GFTA was found to be a significant predictor of the phonotactic frequency effect. The variability in severity of our cohort for the children with PD potentially confounded some of our findings and if the sample size had been larger we could have undertaken separate analysis for the children with mild PD and severe PD, to explore if differences emerged for the mild and severe groups. Furthermore, performance on the NWRep task may have been potentially confounded for the children with CAS due to their assumed speech motor deficit, thus impacting on findings.

The low accuracy scores on the picture-naming task for the children with CAS ($M = 52\%$), PD (59%) and TD (71%), compromises the reliability of the reaction time measure and the interpretation of the task results. The level of difficulty of the task may have potentially compromised the sensitivity of this task to detect group differences in priming, even though it was clearly sensitive to the robust effects of priming. The low mean accuracy scores for this task indicates that the items selected affected the success of the task. Despite the items selected for Study 2 fitting the criteria relating to age of acquisition and frequency of occurrence, greater consideration is needed regarding stimuli selection. The delayed picture naming task also had a high error rate (33.8%), especially given there was time to prepare the response prior to the stimulus presentation. However, errors included wrong names, false starts, coughing and touching the microphone, so were not solely related to incorrect naming. The task could be enhanced to minimize the error rate relating to non-specific naming errors.

One way to assess the developmental constraints that can potentially emerge in speech disordered children would be to undertake a longitudinal study looking at the different aspects of speech over the course of development. It might be expected that deficits that originate in particular aspects or levels of the speech and language system, such as speech motor programming, should be evident in early development. It would be of interest to determine if younger infants who go on to develop CAS or PD show similar profiles of deficit in those early stages of development (Highman, Hennessey, Leitão, & Piek, 2013).

Further studies exploring phonological competency in children with CAS needs to be undertaken and follow up analysis using regressions to explore the points of difference that emerged in the current study. The relationship between DPNrt and

PNrt that differentiated between the children with CAS and PD warrants further examination. The significant correlation between the DPNrt and the PhonFreq effect for the children with CAS also warrants further investigation, given the correlation was not significant for the children with PD. More studies exploring the relationships between speech motor measures and measures of phonological competence need to be undertaken in children with CAS and PD, simultaneously to determine if different constraints emerge for these two disorders.

Summary and Conclusions

The main purpose of this study was to assess the extent of shared deficits in CAS and PD at the level of phonological competence and speech motor ability and to explore the interdependencies and associations between these levels of processing. The children with CAS overall showed more severe deficits in speech motor control consistent with their diagnosis. In contrast, the CAS and PD groups show similar deficits at the level of phonological competence in the domains of input phonology and the development of abstract phonemic codes. There was no evidence of impairment for either group in the efficiency of output phonological processing. On finding similar deficits in phonological competence, however, it is not possible to conclude that CAS and PD share a common underlying cause of these difficulties. Consequently, regression analysis was undertaken to determine if the speech motor measures predicted phonological competence in children with CAS to a greater extent than in children with PD, consistent with the view that the underlying deficit in CAS (i.e., a deficit at the level of speech motor control) constrains development of higher-level phonological/linguistic representations. Overall, when the motor measures were analysed as potential predictors in the regression analysis they did not support our hypotheses, in that they failed to predict phonological competence in children with CAS to a greater extent than in children with PD. In addition, the two measures directly targeting a speech motor deficit (CAS trait score and DPNrt) did not predict phonological competence for the group as a whole, except for the NWrep Low accuracy measure where it was shown that less accurate NWR performance was associated with a higher CAS trait score and slower DPNrt. So there was little evidence across all children that measures of the speech motor control system are directly related to measures of phonological competence after taking into account the contribution of age and vocabulary. GFTA, a measure of articulation accuracy, did

uniquely predict measures of phonological competence across all children (freqEff, NWrep Low, and d prime), but as discussed in Chapter 6, conclusions based just on the GFTA should be treated with some caution because the GFTA measure does not clearly discern between articulatory problems due to a speech motor deficit and those due to a deficit at the level of phonological representations.

These results, therefore, fail to support the developmental interaction account of the deficits observed in CAS (e.g., Maassen, 2002; Pennington & Bishop, 2009) whereby a deficit originating at the level of speech motor control, because speech motor skills and phonological knowledge interact during development, will constrain the development of phonological knowledge resulting in a tight coupling between the severity of the speech motor deficit and phonological competence. Furthermore, the results failed to show the predicted difference between CAS and PD in the pattern of relationship between speech motor ability and phonological competence. These findings are inconsistent with the original hypothesis that CAS originates at the level of speech motor control and PD originates at a higher linguistic-phonological level. It is not possible based on these data to conclude that CAS and PD share a single underlying deficit, however. Overall, these conclusions should be seen as tentative because of their basis on null findings. There was also the finding of a significant group by DPNrt interaction when predicting PNrt that indicates a dissociation between the efficiency of initiating speech plans and lexical encoding processes in children with CAS, whereas these processes appear to be associated in PD and TD children. These findings lend some support to the possibility that there are one or more factors linked to the underlying deficit in speech motor control in CAS that determines efficiency of responding in the delayed picture naming task that is independent of higher level lexical development and also not present in children with PD. These issues are discussed further in the General Discussion.

Chapter 6

General Discussion

Overview

Despite the general consensus that the underlying deficit in CAS is a speech motor deficit and the underlying deficit in PD is phonological in nature, there is little evidence to support this proposal. Children with CAS and PD present with similar deficits, and to date empirical evidence that differentiates between these two disorders is limited. Therefore, the main goal of this PhD was to find empirical evidence to determine if CAS and PD have different etiologies, with speech motor control being the core deficit in CAS, which constrains the development of higher-level phonological abilities. Having evidence of different causal origins is a key criterion, as proposed by Meehl (1992), which can be used to determine whether two pathological or disordered groups form different taxons.

A number of tasks were developed and piloted in Study 1 with children with typical development that targeted phonological competence and speech motor ability. These tasks were then implemented in Study 2 to determine the extent of shared deficits in children with CAS, PD and TD. Three experimental tasks focused on different components of phonological competence, these included; the abstractness of phonemic categories (nonword repetition task), output phonological representations and phonological encoding efficiency (picture-naming task with auditory primes) and input phonological representations (speech discrimination task). A delayed picture-naming task was implemented to target speech motor execution processes and a measure of severity of speech motor deficit, the CAS trait score, derived from the exploratory factor analysis was also implemented in study 2 as an additional measure of speech motor ability. Hierarchical regression analysis was then used to determine if speech motor measures predicted measures of phonological competence in children with CAS to a greater extent than in children with PD and TD.

The systematic review reported in Chapter 3 identified the most prevalent features used in published research literature between 1993 and 2013 for classifying children as CAS. These features were then operationally defined and validated using the children recruited for Study 2. By quantifying the CAS related features we could use exploratory factor analysis to determine if those features related to the single (i.e., unidimensional) underlying construct, consistent with a trait-like speech motor deficit. A single factor solution was confirmed and a novel classification protocol using discriminate function analysis was developed. The factor scores

obtained from the factor analysis were used as a measure of severity of CAS, in combination with the usual metric of number of CAS features. Children with CAS in Study 2 were verified as having CAS using this new protocol to ensure that our CAS sample was representative of children classified as such, consistent with the published research literature. This procedure resulted in some children with suspected CAS from the originally clinically ascertained group being excluded from the final CAS sample, resulting in a smaller sample size. This was regarded as the preferred option to test the main research hypothesis of this thesis as it ensured that the children with CAS met strict criteria for inclusion.

General Summary of Findings

The nonword repetition task and the speech discrimination task replicated a number of findings from previous research, in Study 1 and Study 2, validating these tasks as measures of phonological competence. The picture-naming task, although demonstrating robust phonological facilitation effects, failed to reveal a difference between the younger and older children and was amended for implementation in Study 2 with children with CAS, PD and TD. The simple verbal reaction time (SVrt) revealed age differences consistent with our understanding that younger children have less efficient speech motor skills than older children. The SVrt just missed out on reaching significance as a predictor of speech discrimination ability (d') in children with TD in Study 1, but it interacted with group (i.e., younger versus older children), indicating that the groups differed in this relationship. Further analysis revealed that as reaction time increased, speech discrimination ability decreased for the younger children, indicating that speech motor execution efficiency predicts speech discrimination ability in younger children. This was not the case for the older children demonstrating independence of speech discrimination ability and speech motor execution processes for these children. This finding supports our hypothesis that speech motor measures predict phonological development, in this instance input phonological representations, to a greater degree in younger children than older children. The theoretical implication being that the two systems are interdependent in early development and as development proceeds the different levels of processing become more encapsulated (or modularized) over time. This validates the SVrt as a sensitive measure of speech motor execution relevant to speech motor development and for this reason a similar task was implemented in Study 2.

A unique approach was implemented to reclassify the children with SSD in Study 2. Prior to reclassifying these children as CAS and PD, we operationalized the most prevalent features identified in the systematic review. By using these operationalized features to reclassify our children with SSD in Study 2 we maximized the likelihood that the children with CAS were representative of children classified as such in the research literature, but with a more stringent and transparent protocol. Furthermore, by operationalizing these features we had quantifiable data that could be used to assess the underlying construct of these feature variables utilizing exploratory factor analysis. A key finding of this research was that the factor analysis resulted in a single factor solution for five of the CAS related features, demonstrated by a high factor loading for each of these features and indicating that these features relate to the same underlying construct. In the context of CAS, this single factor solution could be reliably interpreted as an underlying speech motor deficit, consistent with the collective perspectives that CAS is a symptom complex with an underlying deficit in speech motor control (ASHA, 2007; Lewis, Freebairn, Hansen, Taylor, et al., 2004; Maassen et al., 2010; Shriberg et al., 2012). Another notable finding from the factor analysis was the clear difference between children with CAS compared to the children with PD and TD, who were different overall in terms of these CAS related features but also in relation to the Guttman Scaling pattern and factor scores. The Guttman scaling method used to examine the CAS related features (detailed in Chapter 4) revealed a pattern consistent with the understanding that CAS is a unidimensional continuum of a praxis type deficit. This was demonstrated by the pattern that emerged for the children with CAS, indicating that children who were more severe in relation to the praxis deficit, had more CAS related features, but these children also presented with less prevalent and unique CAS related features. The distribution of features for the children with CAS was more consistent with a unidimensional scale according to the Guttman Scaling approach compared to the children with PD (Price, 2016). Although some of the children with PD presented with some of the CAS related features the Guttman pattern was not consistent with a unidimensional scale (or the same unidimensional scale as the children with CAS). The factor scores, reflecting severity of deficit, were consistently larger for the children with CAS, compared to the children with PD and TD. Furthermore, the correlation between PCC and sum of features was significant for the children with CAS but not for the children with PD,

which was another point of difference between CAS and PD indicative of a stronger association between PCC and sum of features for the children with CAS.

CAS and PD

A significant portion of the research in CAS has focused on finding differential markers for this disorder, however, the majority of this research failed to consider the interactive and dynamic nature of speech and language development. The interdependency of the different levels of processing within the speech and language system is not a novel proposition. The dynamic systems view of development ratifies this interdependency and has been well documented in the research literature in early speech and language development (Kuhl, 1993; Thelen, 2005; Thelen & Bates, 2003). In addition, connectionist modeling, consistent with the dynamic view of speech and language development further supports the interdependency of the different levels of processing and representation in speech and language development (Guenther, 1994; Westermann & Miranda, 2004). Terband and Maassen (2009) used connectionist modeling to examine the specific levels of breakdown in children with CAS and the possible reason(s) for these deficits. The role of sensorimotor integration and its significance to speech and language development has also been addressed in a number of research papers examining the dynamic nature of speech and language development (Bruderer, Danielson, Kandhadai, & Werker, 2015b; Westermann & Miranda, 2004), and theoretical frameworks, such as Van der Merwe's (1997), further highlights the interdependencies between the different levels of processing and the importance of internal feedback at different stages of development.

According to the dynamic view of speech and language development, in developmental disorders it is the associations between deficits and not the dissociations that provides greater clarity relating to underlying deficits and prognosis (Pennington & Bishop, 2009). Consequently, the patterns and associations that emerge between the different levels of processing during experimental tasks are highly informative of the potential underlying deficits. This is consistent with Meehl's (1992) concept of taxonomy. To reiterate, taxonomy refers to a method of classification that has been used to address questions addressing the categorical discreteness of psychopathological diagnosis with the concept of a taxon relating to a specific aetiology that is associated with a specific pathology (Meehl, 1992).

Consequently, different aetiologies, despite having similar pathologies, can ensue from different development trajectories (Meehl, 1992). Therefore, in the context of CAS and PD, shared deficits at multiple levels of the speech and language system observed does not preclude the possibility these disorders have different aetiologies, with deficits arising from different or distinct constraints within the developing system. By examining CAS and PD in terms of this notion of developmental constraint, permits a more comprehensive understanding of their respective aetiologies. In particular, we hypothesized that if CAS and PD are distinctly different and CAS has an underlying deficit in speech motor control and PD has an underlying phonological deficit, then the regression analysis has the potential to tease out the different constraints on development through differences in the associations between measures of speech motor ability and phonological competence.

The main aim of this PhD was to explore the differences in underlying deficits between children with CAS and PD by comparing both groups, along with a group of TD children, in terms of their speech motor ability and measures of phonological competence and by exploring the relationships between these measures. Our key hypothesis was that speech motor measures would predict phonological competence measures in children with CAS to a greater extent than in children with PD, reflecting the speech motor system as the primary locus of deficit for children with CAS. This result would confirm that there are different constraints on development between children with CAS and PD consistent with these disorders having distinct underlying aetiologies, in keeping with Meehl's (1992) concept of taxonomy.

The children with CAS and PD had a number of shared deficits, as expected. The children with CAS demonstrated poorer accuracy than the children with PD and TD on nonword repetition, poorer speech discrimination ability and poorer picture-naming in relation to reaction time and accuracy. The children with CAS were significantly poorer than PD and TD in relation to NWR accuracy. They did not differ in picture-naming reaction time from the children with PD and TD, although they were numerically faster. The children with CAS were however significantly poorer in speech discrimination ability (d' -prime) from the children with TD, but not significantly poorer from the children with PD, consistent with previous studies that have shown speech perception deficits in children with CAS (Nijland, 2009) and PD (Edwards et al., 1999; Edwards et al., 2002). Numerically the children with PD had a larger phonotactic frequency effect than the children with CAS and TD, although

this difference was not significant. If this study is replicated and the children with PD were shown to have a significantly larger phonotactic frequency effect then this would suggest that the children with PD were more disadvantaged on the low frequency items compared to the other two groups, indicative of a more severe deficit at this level of processing.

In relation to shared deficits at the level of speech motor control, the children with CAS had the slowest delayed picture-naming reaction time, but not significantly slower than the PD, however both were significantly slower than the children with TD. The children with CAS had the highest CAS trait score and all groups differed significantly from one another. Groups also differed significantly in relation to articulation accuracy as measured by the GFTA, with the children with CAS having the most severe deficit at this level of processing.

Overall, the findings from the experimental tasks are consistent with the shared deficits demonstrated in children with CAS and PD discussed in Chapter 1, with both the children with CAS and PD demonstrating deficits at multiple levels of the speech processing system. In addition, the children with CAS and PD also showed some weakness in vocabulary compared to TD in spite of being in the normal range for most of these children, indicating that the speech disorder for both groups does not arise from a higher-level lexical semantic type deficit constraining their phonological and speech motor abilities.

Our key hypothesis was that the speech motor measures would predict phonological competency to a greater extent in children with CAS compared to children with PD and TD. Overall the regression analysis did not support our hypothesis. The motor measures did predict some of the measures of phonological competence, although, largely they did not interact with group, indicating that these relationships did not differ for children with CAS, PD and TD. In relation to the GFTA as a primary predictor, it is plausible that the small sample size for CAS limited the potential to show an enhanced relationship between the GFTA and outcome measures for CAS, possibly explaining the lack of interactions observed. The DPNrt did not predict PNrt but the interaction term with group was significant indicating that the groups differed in this predictive relationship. The DPNrt predicted PNrt for the children with PD and TD but not for the children with CAS, which was a noteworthy point of difference indicating that there is dissociation between these measures for the children with CAS. One plausible explanation is that

higher-level stages of picture-naming develop normally for CAS. The fact that the children with CAS did not differ in the phonological facilitation effect further supports this proposal. If the speech motor system is more selectively impaired in children with CAS then the source of individual differences in performance on the DPNrt measure may be independent of higher levels of development. It may be that individual differences are due to other factors relating to the severity of the speech motor deficit, such as feedforward commands or the degree of noise in the neural network (Terband et al., 2009). In contrast, it appears as though the efficiency of lexical processes in children with PD and TD is related to the efficiency of the speech motor system in executing planned responses, although it is not possible to know the causal direction of this relationship. Although, it is unrelated to age differences and differences in vocabulary size, given these factors were controlled. This finding suggests that there are different constraints in the developing system for children with CAS, compared to children with PD and TD, at least for this task. Although the regression analysis failed to demonstrate a difference between the groups for the DPNrt when predicting the phonotactic frequency effect there were notable differences in the correlations between these measures for the three groups. The correlation was significant for the children with CAS but not for the children with PD or TD, consistent with the view that different constraints emerge in the development of phonemic categories for children with CAS compared to PD and TD. Furthermore, different patterns of associations emerged between measures of vocabulary and measures of phonological competence for the children with CAS compared to the children with PD and TD. The correlations between nonword repetition accuracy and vocabulary measures, and speech discrimination and vocabulary were significant for the children with PD and TD but not for the children with CAS, again suggesting different constraints in the developing speech and language systems of children with CAS compared to the children with PD and TD.

In summary, overall the regression analysis findings indicate that the speech motor deficits in CAS do not directly constrain the phonological system in a way that differentiates them from children with PD or TD. The differences between CAS and PD may be related to severity of deficits, without there being a qualitative difference in the underlying deficit. This could be interpreted that the underlying deficit in CAS is not primarily a motor deficit that impacts on the development of higher-level linguist constructs (Maassen, 2002; Pennington & Bishop, 2009). However, the fact

that the CAS related features loaded onto a single factor solution suggests otherwise. Furthermore, the Guttman scaling method for the CAS related features demonstrated a pattern consistent with a unidimensional construct for the children with CAS but not for the children with PD, indicating a different underlying construct for these two disordered groups. The correlation analysis also revealed different associations between the motor measures and the measures of phonological competence for the children with CAS, PD and TD, thus indicating different relationships between domains of deficit in children with CAS and PD.

An alternative account may be that our proposition is incorrect and speech motor measures do not constrain phonological development to a greater degree in children with CAS compared to children with PD. A longitudinal approach may help test this proposition, although this is difficult given the low prevalence of CAS. There are potentially a range of factors that determine the level of phonological competence, however, we controlled for the main one, lexical development, by incorporating expressive and receptive vocabulary into the regression analysis. It is plausible that other factors undermined the possibility of finding a direct relationship between measures of speech motor control and measures of phonological competence. The predicted constraints may have operated at an earlier stage of development, as demonstrated by the predictive relationship between SVrt and d-prime in Study 1 with younger TD children. For children with SSD, such as CAS and PD, the highly variable nature of these disorders and the complexity of development in a disordered system, further complicates the task of uncovering potential differences in their developmental trajectories. If the constraints are interpreted as a footprint of where the deficits originate then perhaps the footprint may no longer be detectable later in development. Consequently, a design looking at these relationships, at earlier and later stages of development would be useful.

Limitations and Future Directions

One major limitation of this research was the number of participants in each study, thus potentially compromising findings by reducing the power and diminishing the potential to find significant effects. In addition, following the discriminate function analysis and reclassification of CAS participants our groups had unequal numbers, which can also affect power and reduce the likelihood of significant findings. Furthermore, the children with PD in our study included

children with mild and severe deficits, which potentially confounded some of the findings. Better control in research regarding severity of speech impairment is needed to tease out qualitative differences in underlying speech motor deficits. It may be that CAS is on the extreme end of a continuum of speech motor impairments that are an integral part of a broader verbal trait deficit, consistent with Lewis and colleagues (Lewis, Freebairn, Hansen, Taylor, et al., 2004; Lewis et al., 2000). The underlying aetiologies therefore may not be distinct between children with CAS and PD but the speech motor symptoms associated with CAS might be the characteristic features of a child's speech that emerges as severity of impairment to the speech motor system increases (H. Terband et al., 2009). Consequently, controlling for severity of deficit would be an important requirement in future research that compares children with CAS with other sub-types of SSD. In addition, longitudinal research would be of value that includes children with CAS and PD to ascertain if different constraints in the development of the speech and language system are in fact detectable at an earlier stage of development. This might tease out differences in the profile of underlying deficits during development, suggesting different etiology, regardless of severity.

The use of exploratory factor analysis and regression analysis in CAS is under utilized, most probably due to the lack of operationally defined features since quantifiable data are needed for this type of analysis. We, therefore, propose future research use operationally defined features in classification of CAS that will permit exploratory analysis and enhance our understanding of this complex and highly heterogeneous disorder. The protocol developed and validated in this thesis is an example of a more stringent protocol that could help in better classification of children with CAS. This protocol could be utilized in research and in the clinical setting to ensure children are diagnosed with greater accuracy. However, it would also be of value to test the protocol developed in Chapter 4 on another sample of children with SSD.

In addition, this research has highlighted the overlap of phonological and speech motor deficits in CAS and PD, indicating that children with CAS and PD share in the risks associated with such deficits. Consequently, clinicians need to ensure that they focus on higher-level aspects of speech and language development, such as lexical-phonological knowledge, and not just focus on the speech motor deficits associated with this disorder. Treatment protocols predominantly focus on

the lower-level speech motor deficits observed in CAS, using programs such as the Nuffield Dyspraxia program (Willams & Stephens, 2010), which focuses on teaching motor programming skills. However, there are linguistic approaches available that emphasize phonological patterns that also focus on the functionality of speech (Velleman & Vihman, 2002). To ensure that children with CAS obtain the optimal treatment, protocols should include both phonological and speech motor deficits.

It would also be useful to explore the underlying dimensionality of PD using a similar paradigm implemented in the current study (Chapter 4) to determine if features associated specifically with PD also resulted in a singular construct (e.g., a PD trait deficit). This type of analysis has been used in SSD in general, not to explore the underlying dimensionality of SSD, but to determine predictors of language, reading and spelling ability (Lewis et al., 2000). Exploratory factor analysis has also been used to explore the different dimensions of SSD, resulting in two distinct constructs, one consistent with articulation/phonology deficits and the other consistent with semantic/syntactic deficits (Lewis et al., 2006). Using Lewis et al.'s (2006) paradigm, it would be of interest to include children with CAS and PD and to also include measures of speech motor control, such as the CAS trait score, to determine if measures load onto a single factor for the children with CAS, similar to children with SSD in general, or if the speech motor measures dissociate from the phonological measures.

Finally, further research is warranted to investigate the different relationship that emerged between the DPNrt, as a measure of initiation time of preplanned speech, and how that dissociated from the PNrt, as a measure of phonological encoding efficiency, for the children with CAS but not the children with PD and TD. This different pattern indicates dissociation between these levels of processing for the children with CAS but not for the children with PD, which could potentially be used as a differential marker for diagnosis.

Conclusion

Despite our key hypothesis not being supported by our findings from the regression analysis this PhD thesis had made a number of novel contributions to the evaluation of phonological competence and speech motor ability in children with SSD, and to our understanding of CAS and the relationship between CAS and PD.

- We explored phonological facilitation effects during picture-naming and phonotactic frequency effects on NWR accuracy in children with CAS, PD and TD simultaneously, thus broadening the focus to include other measures of phonological competence. Despite neither of these measures showing marked differences between the SSD groups and the TD group, there was a non-significant trend for the frequency effect to be weaker for CAS (but not for PD). Further research is needed to determine if a deficit at this level of the phonological processing is the primary source of deficits observed in CAS or an incidental constraint as a result of an underlying speech motor deficit.
- We attempted to develop a real-time measure of speech motor planning and execution processes. The simple and choice verbal reaction time tasks were developed and implemented in Study 1 and the delayed picture-naming task was developed and implemented in Study 2. The CVrt did not capture speech motor programming as intended, but despite its limitations the SVrt revealed a significant differences in younger and older children with TD, showing some merit. The DPNrt, implemented in Study 2 in place of the SVrt, revealed significant differences for the children with CAS and PD compared to the children with TD. However, it failed to differentiate between the children with CAS and PD and therefore did not differentiate between a speech motor deficit and a phonological deficit. A more refined task is required that specifically targets speech motor ability to differentiate between CAS and PD.
- A review of classification protocols was undertaken and highlighted a number of incongruities and ambiguities with regard to protocols used for classifying children as CAS. Moreover, limitations in the use of operationalized definitions were highlighted. This review resulted in the identification of the most prevalent and consistently used features to classify children as having CAS between 1993 and 2013. This review provides researchers with an opportunity to use a feature based checklist approach to classification that can ensure that the children recruited are representative of the broader population of children classified by researchers as having CAS.
- The most prevalent features identified in the systematic review were operationalized using a readily available clinical tool, the Inconsistency Subtest of the DEAP, not only providing reliable data (as indicated by the high inter-

rater reliability) but also with regard to time efficiency given this tool is regularly used in the classification of CAS and therefore has already been administered.

- The operational definitions developed in Chapter 4 enabled exploratory factor analysis to be undertaken to explore the underlying dimensionality of the CAS related features, revealing a one-factor solution, which was interpreted to be consistent with a praxis-type speech motor deficit. The terminology used in this thesis was to refer to this construct as a CAS trait score.
- Contingent on this one factor solution, a novel protocol for classifying children as CAS was developed using the factor scores (indicating severity of CAS), obtained from the exploratory factor analysis, as well as the frequency of CAS related features. This protocol was then used in Study 2 to reassign children with SSD into either the CAS or PD group.
- Study 2 built on previous findings in relation to the extent of shared deficits in CAS and PD and specifically in relation to the phonotactic frequency effect, phonological facilitation effect and speech discrimination, not previously investigated in children with CAS. Children with CAS did not demonstrate a phonotactic frequency effect, indicating that they do not perform similar to younger children with TD. The fact that they did not demonstrate particular advantage for the high frequency sequences could be that children with CAS have not yet attained a sufficient degree of phonological development to benefit from the high frequency sequences. The groups did not differ from one another in relation to phonological facilitation effects, despite amending the task to include two SOAs and using rimes in place of onset primes in Study 2. This finding is consistent with recent research that showed that children with SSD do not differ in phonological facilitation effects, indicating that deficits in SSD are not related to deficient phonological encoding ability (Munson & Krause, 2017). In relation to speech discrimination ability, the children with CAS were significantly poorer in relation to accuracy for the whole word condition compared to the TD children, and poorer but not significantly from the children with PD, indicating both children with CAS and PD have deficient input phonological representations.

- Finally, Study 2 explored the relationship between speech motor competence and phonological competence in children with CAS, PD and TD. The regression analysis used to unpack these relationships has not been used in CAS. On the whole, however, the regression analysis failed to differentiate between the groups, indicating that the speech motor measures, as hypothesized, did not predict phonological competence in children with CAS to a greater extent than in children with PD or TD. However, some differences emerged in relation to the correlations between the groups that are worthy of further investigation. Further analysis is warranted in order to confirm whether or not differences in the relationships between measures of speech motor skills and phonological competence exist in children with CAS and PD. One specific area of interest is the nature of the relationship between measures of efficiency of execution of speech plans, using a delayed picture naming paradigm, and higher order lexical retrieval, using speeded picture naming, which appears to be dissociated in CAS but not in PD or TD.

In summary, this PhD has added to our knowledge of CAS in a number of ways. The systematic review highlighted the shortcomings of research in the area of CAS resulting in a list of the most pertinent features of CAS and the reliability of these features over the period of time covered by the review. The use of operationally defined features removes the ambiguity associated with how children meet classification criteria and therefore enhances the application and usefulness of research findings. Furthermore, by quantifying features of CAS we were able to explore the underlying dimensionality using factor analysis, which supported the proposal that, despite CAS being multi-deficit in nature, these deficits relate to a singular underlying construct, namely a deficit at the level of speech motor control. The new protocol developed, using the CAS trait score and sum of features, provided a clear and concise method by which to classify children as having CAS using operationally defined features.

The majority of studies exploring CAS have failed to include another specific speech disordered population and only compared CAS with TD, thus limiting the application of findings. This thesis explored phonological competence and speech motor ability in children with CAS, PD and TD simultaneously, thereby providing data specific to children with CAS and PD, that could be compared to children with

TD. Furthermore, the tasks undertaken in this study have not been implemented with children with CAS, therefore providing novel insight relating to the underlying dimensionality of CAS and the relationship between different aspects of phonological competency targeted and speech motor ability.

By exploring the relationships between the different levels of processing concurrently helps to tease out possible differences in the causal origin of deficits in CAS by acknowledging the dynamic nature of speech and language development. Despite our endeavors, as a whole the evidence did not support our proposal that speech motor measures would predict phonological competence in children with CAS to a greater extent than in PD. This was demonstrated largely by the absence of significant interaction terms between the speech motor measures and group that tested for the predictions of phonological competence. The one significant and reliable interaction that did emerge between the DPNrt and PNrt demonstrates there are differences between CAS and PD, although this difference could be related to severity of deficit. Furthermore, there were a number of different associations that emerged from the correlation analysis that differentiated between the children with CAS and the children with PD and TD, thereby suggesting that there are some differences in constraints in the developing speech and language systems of children with CAS, PD and TD.

In closing, there was little evidence from the regression analysis to support our proposal that children with CAS have a core deficit in speech motor control that constrains phonological development, resulting in the overlap of deficits observed in CAS and PD. This poses the question if CAS and PD are in fact distinct disorders with different underlying core deficits. One assumption could be that CAS and PD are part of the same continuum, only varying in relation to severity of deficit within the emerging speech motor and phonological systems. Alternatively, constraints in development may not be detectable at the stage of development investigated.

REFERENCES

- Allen, P., & Bennett, K. (2010). *PASW Statistics by SPSS: A Practical Guide to Version 18.0* (A. Crabb Ed. First ed.). Melbourne, Victoria: Cengage Learning.
- ASHA, A.S.-L.a.H.A. (2007). Childhood Apraxia of Speech Technical Report. In ASHA (Ed.), (ASHA ed., pp. 1-74). <http://www.asha.org/policy>: ASHA.
- Bahr, R.H. (2005). Differential diagnosis of severe speech disorders using speech gestures. *Topics in Language Disorders*, 25(3), 254-265.
- Bahr, R.H., Velleman, S.L., & Ziegler, M.A. (1999). Meeting the challenge of suspected developmental apraxia of speech through inclusion. *Topics in Language Disorders*, 19(3), 19.
- Ballard, K.J., Robin, D.A., & Folkins, J.W. (2003). An integrative model of speech motor control: a response to Ziegler. *Aphasiology*, 17(1), 37-48.
- Ballard, K.J., Robin, D.A., McCabe, P., & McDonald, J. (2010). A treatment for dysprosody in childhood apraxia of speech. *Journal of Speech, Language and Hearing Research*, 53, 1227-1245.
- Beckman, M.E., & Edwards, J. (2000). The ontogeny of phonological categories and the primacy of lexical learning in linguistic development. *Child Development*, 71(1), 240-249.
- Betz, S.K., & Stoel-Gammon, C. (2005). Measuring articulatory error consistency in children with developmental apraxia of speech. *Clinical Linguistics & Phonetics*, 19(1), 53-66.
- Bishop, D. (1997). Cognitive neuropsychology and developmental disorders: uncomfortable bedfellows. *The Quarterly Journal of Experimental Psychology*, 50, 899-923.
- Boersma, P., & Weenink, D. (1995). Praat.
- Bowen, C. (2009). *Children's Speech Sound Disorders*. Chichester, West Sussex: Wiley & Sons.
- Bradford, A., & Dodd, B. (1994). The Motor Planning Abilities of Phonologically Disordered Children. *European Journal of Disorders of Communication*.
- Bradford, A., & Dodd, B. (1996). Do all speech-disordered children have motor deficits? *Clinical Linguistics & Phonetics*, 10(2), 77-101.
- Bradford, A., Murdoch, B., Thompson, E., & Stokes, P. (1997). Lip and tongue function in children with developmental speech disorders: a preliminary investigation. *Clinical Linguistics & Phonetics*, 11(5), 363-387.
- Brooks, P.J., & Mac Whinney, B. (2000). Phonological priming in children's picture naming. *Journal of Child Language*, 27(2), 335-366.
- Browman, C.P., & Goldstein, L. (1988). Some Notes on Syllable Structure in Articulatory Phonology. *Phonetica*, 45(2-4), 140-155.
- Browman, C.P., & Goldstein, L. (1992). Articulatory Phonology - an Overview. *Phonetica*, 49(3-4), 155-180.
- Browman, C.P., & Goldstein, L. (1997). The Gestural Phonology Model. In W. Hulstijn, H. F. M. Peters & P. H. H. M. van Lieshout (Eds.), *Speech Production: Motor Control, Brain Research and Fluency Disorders* (pp. 57-71): Elsevier Science, B.V.
- Bruderer, A., Danielson, D., Kandhadai, P., & Werker, J. (2015a). Sensorimotor influences on speech perception in infancy. *Proceedings of National Academy of Science*, 112(44), 13531-13536.

- Bruderer, A., Danielson, D., Kandhadai, P., & Werker, J. (2015b). Sensorimotor influences on speech perception in infancy. *Proceedings of National Academy of Science*, 112(44), 13531-13536.
- Button, L., Peter, B., Stoel-gammon, C., & Raskind, W.H. (2013). Associations among measures of sequential processing in motor and linguistics tasks in adults with and without a family history of childhood apraxia of speech: A replication study. *Clinical Linguistics & Phonetics*, 2013, Vol.27(3), p.192-212, 27(3), 192-212. doi: 10.3109/02699206.2012.744097
- Caruso, A.J., & Strand, E. (1999). Motor speech disorders in children: definitions, background, and a theoretical framework. In A. J. Caruso & E. Strand (Eds.), *Clinical Management of Motor Speech Disorders in Children*. (pp. 1-27). New York: Thieme.
- Caspari, S.S., Strand, E.A., Kotagal, S., & Bergqvist, C. (2008). Obstructive sleep apnea, seizures, and childhood apraxia of speech. *Pediatr. Neurol.*, 38(6), 422-425. doi: 10.1016/j.pediatrneurol.2008.03.002
- Coady, J.A., & Aslin, R.N. (2004). Young children's sensitivity to probabilistic phonotactics in the developing lexicon. *Journal of Experimental Child Psychology*, 89(3), 183-213.
- Cohen, J., & Cohen, P. (1983). *Applied multiple regression/ correlational analysis for behavioural sciences*. (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.
- Coltheart, M. (1981). The MRC Psycholinguistic Database. *Quarterly Journal of Experimental Psychology*, 33, 497-505.
- Crary, M.A. (1995). Clinical evaluation of developmental motor speech disorders. *Seminars in Speech and Language*, 16(2), 111-125.
- Cutting, J.C., Ferreira, V.S., Damian, M.F., & Martin, R.C. (1999). Semantic and Phonological Information Flow in the Production Lexicon. [Article]
- Semantic and Phonological Codes Interact in Single Word Production. [Article].
- Dale, P.S., & Hayden, D.A. (2013). Treating speech subsystems in childhood apraxia of speech with tactual input; the PROMPT approach. *American Journal of Speech - Language Pathology*, 644-661.
- Damian, M.F., & Martin, R.C. (1999). Semantic and Phonological Codes Interact in Single Word Production. .
- Davis, B.L., Jacks, A., & Marquardt, T.P. (2005). Vowel patterns in developmental apraxia of speech: three longitudinal case studies. *Clinical Linguistics & Phonetics*, 19(4), 249-274.
- Davis, B.L., Jakielski, K.J., & Marquardt, T.P. (1998). Developmental apraxia of speech: determiners of differential diagnosis. *Clinical Linguistics & Phonetics*, 12(1), 25-45.
- Dell, G.S., Burger, L.K., & Svec, W.R. (1997). Language Production and Serial Order: A Functional Analysis and a Model. *Psychological Review*, 104(1), 123-147.
- Dewey, D. (1995). What is Developmental Dyspraxia. *Brain and Cognition*, 29, 254-274.
- Dodd, B. (2005). *Differential Diagnosis and Treatment of Children with Speech Disorders*. West Sussex UK: Whurr Publishers Ltd.
- Dodd, B., Holm, A., Crosbie, S., & McCormack, P. (2005). Differential diagnosis of phonological disorders. In B. Dodd (Ed.), *Differential diagnosis and treatment of children with speech disorders*. West Sussex: Whurr Publishers Ltd.

- Dodd, B., Hua, Z., Crosbie, S., Holm, A., & Ozanne, A.E. (2002). Diagnostic Evaluation of Articulation and Phonology: Pearson Assessment.
- Dunn, L. (2007). *Peabody Picture Vocabulary Test-4*. Circle Pines, MN: American Guidance Services.
- Dunn, L., & Dunn, L. (1997). *Peabody Picture Vocabulary Test-III*. Circle Pines, MN: American Guidance Services.
- Edeal, D.M., & Gildersleeve-Neumann, C.E. (2011). The importance of production frequency in therapy for Childhood Apraxia of Speech. *American Journal of Speech-Language Pathology*, 20, 95-110.
- Edwards, J. (1992). Compensatory and speech motor abilities in normal and phonologically disordered children. *Journal of Phonetics*, 20, 189-207.
- Edwards, J., Beckman, M.E., & Munson, B. (2004). The Interaction Between Vocabulary Size and Phonotactic Probability Effects on Children's Production Accuracy and Fluency in Nonword Repetition. *Journal of Speech, Language, and Hearing Research*, 47(2), 421.
- Edwards, J., Fourakis, M., Beckman, M.E., & Fox, R.A. (1999). Characterizing knowledge deficits in phonological disorders. *Journal of Speech Language and Hearing Research*, 42(1), 169-186.
- Edwards, J., Fox, R.A., & Rogers, C.L. (2002). Final consonant discrimination in children: Effects of phonological disorder, vocabulary size, and articulatory accuracy. *Journal of Speech Language and Hearing Research*, 45(2), 231-242.
- Forrest, K. (2003). Diagnostic criteria of developmental apraxia of speech used by clinical speech-language pathologists. *American Journal of Speech-Language Pathology*, 12(3), 376-380.
- Forrest, K., & Elbert, M. (2001). Treatment for phonologically disordered children with variable substitution patterns. *Clinical Linguistics & Phonetics*, 15(1-2), 41-45.
- Forster, J.C. (1997). DMDX. University of Arizona.
- Fowler, A.E. (1991). How early phonological development might set the stage for phoneme awareness. In S. A. Brady & D. P. Shankweiler (Eds.), *Phonological Processes in Literacy* (pp. 97-117). New Jersey: Lawrence Erlbaum Associates.
- Fox, A., Dodd, B., & Howard, D. (2002). Risk factors for speech disorders in children. *International Journal of Language & Communication Disorders*, 37(2), 117-131.
- Froud, K., & Khamis-Dakwar, R. (2012). Mismatch negativity responses in children with a diagnosis of childhood apraxia of speech (CAS).(Research Article)(Report). *American Journal of Speech-Language Pathology*, 21, 302.
- Gathercole, S.E. (2006). Nonword repetition and word learning: The nature of the relationship. *Applied Psycholinguistics*, 27(4), 513-543.
- Gibbon, F.E. (1999). Undifferentiated lingual gestures in children with articulation/phonological disorders. *Journal of Speech Language and Hearing Research*, 42(2), 382-397.
- Gibbon, F.E. (2002). Features of impaired tongue control in children with Phonological Disorder. In F. Windsor, M. L. Kelly & N. Hewlett (Eds.), *Investigations in Clinical Phonetics and Linguistics*. Mahway, New Jersey: Lawrence Erlbaum Associates.
- Gierut, J.A. (1998a). Production, conceptualization and change in distinctive featural categories. *Journal of Child Language*, 25(2), 321-341.

- Gierut, J.A. (1998b). Treatment efficacy: Functional phonological disorders in children. *Journal of Speech Language and Hearing Research*, 41(1), S85-S100.
- Goldman, R., & Fristoe, M. (2000). *The Goldman-Fristoe Test of Articulation 2* (2 ed.). Circle Pines, NM: American Guidance Service.
- Grigos, M.I., & Kolenda, N. (2010). The relationship between articulatory control and improved phonemic accuracy in childhood apraxia of speech: a longitudinal case study. *Clinical Linguistics & Phonetics*, 24(1), 17-40.
- Groenen, P., Maassen, B., Crul, T., & Thoonen, G. (1996). The specific relation between perception and production errors for place of articulation in developmental apraxia of speech. *Journal of Speech and Hearing Research*, 39, 468-482.
- Guenther, F. (1994). A Neural-Network Model of Speech Acquisition and Motor Equivalent Speech Production. *Biological Cybernetics*, 72(1), 43-53.
- Guenther, F., & Perkell, J. (2004). A neural model of speech production and its application to studies of the role of auditory feedback in speech. In B. Maassen, R. Kent, H. Peters, P. van Lieshout & W. Hulstijn (Eds.), *Speech motor control in normal and disordered speech* (pp. 29-50). Oxford: Oxford University Press.
- Guyette, T.W., & Diedrich, W.M. (1981). A Critical Review of Developmental Apraxia of Speech. In N. J. Lass (Ed.), *Speech and Language: Advances in Basic Research and Practice* (Vol. 5, pp. 1-49). New York: Academic Press.
- Hall, P.K., Jordon, L.S., & Robin, D.A. (1993). *Developmental Apraxia of Speech*. Austin, TX: Pro-Ed.
- Hesketh, A., Adams, C., Nightingale, C., & Hall, R. (2000). Phonological awareness therapy and articulatory training approaches for children with phonological disorders: a comparative outcome study. *International Journal of Language & Communication Disorders*, 35(3), 337-354.
- Highman, C., Hennessey, N., Leitão, S., & Piek, J. (2013). Early development in infants at risk of childhood apraxia of speech: A longitudinal investigation. *Developmental Neuropsychology*, 38(3), 197-210. doi: 10.1080/87565641.2013.774405
- Highman, C., Hennessey, N., Sherwood, M., & Leitao, S. (2008). Retrospective parent report of early vocal behaviours in children with suspected Childhood Apraxia of Speech (sCAS). *Child Language Teaching and Therapy*, 24(3), 285-306.
- Highman, C., Leitao, S., Hennessey, N., & Piek, J. (2012). Prelinguistic communication development in children with childhood apraxia of speech: A retrospective analysis. *International Journal of Speech-Language Pathology*, 14(1), 35-47.
- Hosom, J., Shriberg, L., & Green, J. (2004). Diagnostic assessment of childhood apraxia of speech using automatic speech recognition (ASR) methods. *Journal of Medical Speech-Language Pathology*, 12(4), 167-171.
- Iuzzini, J.a.F., K. (2010). Evaluation of a combined treatment approach for childhood apraxia of speech. *Clinical Linguistics & Phonetics*, 24(4-5), 335-345.
- Jacks, A., Marquardt, T., & Davis, B. (2006). Consonant and syllable structure patterns in childhood apraxia of speech: Developmental change in three children. *Journal of Communication Disorders*, 39(6), 424-441.

- Jacks, A., Mathes, K.A., & Marquardt, T.P. (2010). Vowel acoustics in adults with apraxia of speech. *Journal of Speech , Language and Hearing Research*, 53, 61-74.
- Jacks, A., & Robin, D. (2013). Apraxia of Speet. In J. S. Damico, N. Muller & M. J. Ball (Eds.), *The handbook of language and speech disorders* (pp. 391-409). West Sussex: Blackwell Publishing.
- Jamieson, D.G., & Rvachew, S. (1992). Remediating speech production errors with sound identification training. *Journal of Speech-Language Pathology*, 16, 201-210.
- Jerger, S., Martin, R.C., & Damian, M.F. (2002). Semantic and phonological influences on picture naming by children and teenagers. *Journal of Memory and Language*, 47(2), 229-249.
- Jescheniak, j.D., & Levelt, W.J.M. (1994). Word frequency effects in speech production: Retrieval of syntactic information and of phonological form. *Journal of Experimental Psychology-Learning Memory and Cognition*, 20(4), 824-843.
- Kawamoto, A.H., Qiang Liu, K.M., & Sanchez, A. (2008). Articulatory preparation in delayed naming task. *Journal of Memory and Language*, 58, 347-365.
- Kent, R.D. (2000). Research on speech motor control and its disorders: A review and prospective. *Journal of Communication Disorders*, 33(5), 391-428.
- Kent, R.D., & Kim, Y.J. (2003). Toward an acoustic typology of motor speech disorders. *Clinical Linguistics & Phonetics*, 17(6), 427-445.
- Klapp, S.T. (1995). Motor response programming during simple and choice reactive time: the role of practice. *Journal of Experimental Psychology: Human Perception and Performance*, 21(5), 1015-1027.
- Klapp, S.T. (2003). Reaction time analysis of two types of motor preparation for speech articulation: Action as a sequence of chunks. *Journal of Motor Behavior*, 35(2), 135-150.
- Krishnan, S., Alcock, K.J., Mercure, E., Leech, R., Barker, E., Karmiloff-Smith, A., & Dick, F. (2013). Articulating novel words; childrens oromotor skills predict nonword repetition abilities. *Journal of Speech , Language and Hearing Research*, 56, 1800-1812.
- Kučera, H., & Francis, W.N. (1967). *Computational analysis of present-day American English*. Providence: Brown University Press.
- Kuhl, P.K. (1993). Early Linguistic Experience and Phonetic Perception - Implications for Theories of Developmental Speech-Perception. *Journal of Phonetics*, 21(1-2), 125-139.
- Kuhl, P.K., & Meltzoff, A.N. (1996). Infant vocalizations in response to speech: Vocal imitation and developmental change. *Journal of the Acoustical Society of America*, 100(4), 2425-2438.
- Laffin, J., Raca, G., Jackson, C., Strand, E., Jakielski, K., & Shriberg, L. (2012). Novel candidate genes and regions for childhood apraxia of speech identified by array comparative genomic hybridization (vol 14, pg 928, 2012). *Genet. Med.*, 15(7), 587-588. doi: 10.1038/gim.2013.85
- Larrivee, L.S., & Catts, H.W. (1999). Early reading achievement in children with expressive phonological disorders. *American Journal of Speech-Language Pathology*, 8(2), 118-128.
- Le Normand, M.T., Vaivre-Douret, L., Payan, C., & Cohen, H. (2000). Neuro motor development and language processing in developmental dyspraxia: A follow-

- up case study. *Journal of Clinical and Experimental Neuropsychology*, 22(3), 408-417.
- Levelt, W.J.M. (1989). *Speaking: From Intention to Articulation*. London: MIT Press.
- Levelt, W.J.M. (1999). Models of word production. *Trends in Cognitive Sciences*, 3(6), 223-232.
- Levelt, W.J.M., Roelofs, A., & Meyer, A.S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, 22(1), 1-75.
- Lewis, B., Avrich, A., Freebairn, L., Taylor, H., Iyengar, S., & Stein, C. (2011). Subtyping children with speech sound disorders by endophenotypes. *Topics in Language Disorders*, 31(2), 112-127.
- Lewis, B., Freebairn, L., Hansen, A., Iyengar, S., & Taylor, H. (2004). School-age follow-up of children with childhood apraxia of speech. *Language Speech and Hearing Services in Schools*, 35(2), 122-140.
- Lewis, B., Freebairn, L., Hansen, A., Stein, C., Shriberg, L., Iyengar, S., & Taylor, H. (2006). Dimensions of early speech sound disorders: a factor analytical study. *Journal of Communication Disorders*, 36, 139-157.
- Lewis, B., Freebairn, L., Hansen, A., Taylor, H., Iyengar, S., & Shriberg, L. (2004). Family Pedigrees of children with suspected childhood apraxia of speech. *Journal of Communication Disorders*, 37, 157-175.
- Lewis, B., Freebairn, L., & Taylor, H. (2000). Academic outcomes in children with histories of speech sound disorders. *Journal of Communication Disorders*, 33, 11-30.
- Lewis, B., Short, E., Iyengar, S., Taylor, H., Freebairn, L., Tag, J., . . . Stein, C. (2012). Speech-sound disorders and attention-deficit hyperactivity disorder symptoms. *Topics in Language Disorders*, 32(3), 247-263.
- Lindblom, B. (2000). Developmental origins of adult phonology: The interplay between phonetic emergents and the evolutionary adaptations of sound patterns. *Phonetica*, 57(2-4), 297-314.
- Locke, J.L. (1980). The inference of phoneme perception in the phonologically disordered child. Part 2: Clinically novel procedures, their use, some findings. *Journal of Speech and Hearing Disorders*, 45, 445-468.
- Lum, C. (2002). *Scientific Thinking in Speech and Language Therapy*. Mahwah, New Jersey Lawrence Erlbaum Associates.
- Lundeborg, I., & McAllister, A. (2007). Treatment with a combination of intra-oral sensory stimulation and electropalatography in a child with severe developmental dyspraxia. *Logop. Phoniatr. Vocology*, 32(2), 71-79. doi: 10.1080/14015430600852035
- Lunganaro, M., & Xavier Alario, F. (2006). On the locus of syllable frequency effect in speech production. *Journal of Memory and Language*, 55, 178-196.
- Maas, E., Butalla, C.E., & Farinella, K.A. (2012). Feedback frequency in treatment for childhood apraxia of speech. *American Journal of Speech - Language Pathology*, 21, 239-257.
- Maas, E., & Farinella, K.A. (2012). Random Versus Blocked Practice in Treatment for Childhood Apraxia of Speech. *J. Speech Lang. Hear. Res.*, 55(2), 561-578. doi: 10.1044/1092-4388(2011/11-0120)
- Maassen, B. (2002). Issues contrasting adult acquired versus developmental apraxia of speech. *Seminars in Speech and Language*, 23(4), 257-266.

- Maassen, B., Groenen, P., & Crul, T. (2003). Auditory and phonetic perception of vowels in children with apraxic speech disorders. *Clinical Linguistics & Phonetics*, 17(6), 447-467.
- Maassen, B., Nijland, L., & Terband, H. (2010). Developmental Models of Childhood Apraxia of Speech. In B. Maassen & P. van Lieshout (Eds.), *Speech Motor Control: New developments in basic and applied research*. NY: Oxford University Press.
- Maassen, B., Nijland, L., & Van der Meulen, S. (2001). Coarticulation within and between syllables by children with developmental apraxia of speech. *Clinical Linguistics & Phonetics*, 15(1-2), 145-150.
- MacWhinney, B. (1998). Models of the emergence of language. *Annual Review of Psychology*, 49, 199-227.
- Marignier, S., Lesca, G., Marguin, J., Bussy, G., Sanlaville, D., & des Portes, V. (2012). Childhood apraxia of speech without intellectual deficit in a patient with cri du chat syndrome. *Eur. J. Med. Genet.*, 55(6-7), 433-436. doi: 10.1016/j.ejmg.2012.03.008
- Marion, M., Sussman, H., & Marquardt, T. (1993). The perception and production of rhyme in normal and developmentally apraxic children. *Journal of Communication Disorders*, 26(3), 129-160.
- Markus, D.F., Bowers, J.S., Stadthagen-Gonzalez, H., & Spalek, K. (2010). Does word length affect speech onset latencies when producing single words? *Journal of Experimental Psychology-Learning Memory and Cognition*, 36(4), 892-905.
- Marquardt, T., Jacks, A., & Davis, B. (2004). Token-to-token variability in developmental apraxia of speech: three longitudinal case studies. *Clinical Linguistics & Phonetics*, 18(2), 127-144.
- Marquardt, T., Sussman, H., Snow, T., & Jacks, A. (2002). The integrity of the syllable in developmental apraxia of speech. *Journal of Communication Disorders*, 35(1), 31-49.
- Martikainen, A.L., & Korpilahti, P. (2011). Intervention for childhood apraxia of speech: A single-case study. *Child Lang. Teach. Ther.*, 27(1), 9-20. doi: 10.1177/0265659010369985
- Mc Neil, B.C., Gillon, G.T., & Dodd, B. (2009). Phonological awareness and early reading development in childhood apraxia of speech (CAS). *International Journal of Language and Communications Disorders*, 44(2), 175-192.
- McCabe, P., Rosenthal, J.B., & McLeod, S. (1998). Features of developmental dyspraxia in the general speech-impaired population? *Clinical Linguistics & Phonetics*, 12(2), 105-126.
- McCauley, R.J., & Strand, E. (2008). Treatment of childhood apraxia of speech: clinical decision making in the use of nonspeech oral motor exercises. *Seminars in Speech and Language*, 29(4), 284-293.
- McCune, L., & Vihman, M.M. (2001). Early phonetic and lexical development: A productivity approach. *Journal of Speech Language and Hearing Research*, 44(3), 670-684.
- McLeod, S.a.H., L.J. (2009). Epidemiology of speech and language impairment in a nationally representative sample of 4-5 year old children. *Journal of Speech, Language and Hearing Research*, 52(5), 1213-1229.
- McNeil, M., Pratt, S., & Fossett, T. (2004). The differential diagnosis of Apraxia of Speech. In B. Maassen, R. D. Kent, H. Peters, P. van Lieshout & W. Hulstijn

- (Eds.), *Speech Motor Control in Normal and Disordered Speech* (pp. 51-81). NY: Oxford University Press.
- McNeil, M., Robin, D., & Schmidt, R. (1997). Apraxia of speech: definition, differentiation and treatment. In M. R. Mc Neil (Ed.), *Clinical Management of Sensorimotor Speech Disorders* (pp. 311-344). New York: Thieme.
- McNeill, B., Gillon, G., & Dodd, B. (2009a). Effectiveness of an integrated phonological awareness approach for children with childhood apraxia of speech (CAS). *Child Language Teaching and Therapy*, 25(3), 341-366.
- McNeill, B.C., Gillon, G.T., & Dodd, B. (2009b). Phonological awareness and early reading development in childhood apraxia of speech. *International Journal of Language & Communication Disorders*, 44(2), 175-192.
- Meehl, P.E. (1992). Factors and Taxa, Traits and Types, Differences of Degree and Differences in Kind. *Journal of Personality*, 60(1), 117-.
- Menn, L., & Stoel-Gammon, C. (1995). Phonological Development. In P. Fletcher & B. Mac Whinney (Eds.), *The handbook of child language* (pp. 335-359). Oxford: Blackwell.
- Metsala, J.L. (1997). An examination of word frequency and neighborhood density in the development of spoken-word recognition. *Memory & Cognition*, 25(1), 47-56.
- Metsala, J.L. (1999). Young children's phonological awareness and nonword repetition as a function of vocabulary development. *Journal of Educational Psychology*, 91(1), 3-19.
- Meyer, A.S. (1991). The time course of phonological encoding in language production: Phonological encoding inside a syllable. *Journal of Memory and Language*, 30(1), 69-89.
- Miccio, A.W., Elbert, M., & Forrest, K. (1999). The relationship between stimulability and phonological acquisition in children with normally developing and disordered phonologies. *American Journal of Speech-Language Pathology*, 8(4), 347-363.
- Moher, D., Liberati, A., Tetzlaff, J., & Altman, D.G. (2009). Preferred reporting items for systematic reviews and meta-analysis: The PRISMA Statement. *Annals of Internal Medicine*, 151(4).
- Monsell, S. (1987). On the relation between lexical input and output pathways for speech. In A. Allport, D. Mac Kay, W. Prinz & E. Sheerer (Eds.), *Language perception and production* (pp. 273-311). London: Academic Press.
- Morgan, A.T., & Vogel, A.P. (2009). A Cochrane review of treatment of childhood apraxia of speech. *European Journal of Physical Rehabilitation*, 45(1), 103.
- Moriarty, B.C., & Gillon, G.T. (2006). Phonological awareness intervention for children with childhood apraxia of speech. *International Journal of Language and Communications Disorders*, 41(6), 713-734.
- Munson, B. (2001a). Phonological pattern frequency and speech production in adults and children. *Journal of Speech Language and Hearing Research*, 44(4), 778-792.
- Munson, B. (2001b). Relationships between vocabulary size and spoken word recognition in children aged 3 to 7. *Contemporary Issues in Communication Science and Disorders*, 28, 20-29.
- Munson, B., & Babel, M. (2005). The sequential cueing effect in children's speech production. *Applied Psycholinguistics*, 26(2), 157-174.
- Munson, B., Bjorum, E., & Windsor, J. (2003). Acoustic and perceptual correlates of stress in nonwords produced by children with suspected developmental

- apraxia of speech and children with phonological disorder. *Journal of Speech Language and Hearing Research*, 46(1), 189-202.
- Munson, B., Edwards, J., & Beckman, M. (2005a). Phonological knowledge in typical and atypical speech-sound development. *Topics in Language Disorders*, 25(3), 190-206.
- Munson, B., Edwards, J., & Beckman, M. (2005b). Relationships Between Nonword Repetition Accuracy and Other Measures of Linguistic Development in Children With Phonological Disorders. *Journal of Speech, Language, and Hearing Research*, 48(1), 61.
- Munson, B., & Krause, M. (2017). Phonological encoding in speech-sound disorder: evidence from a cross-modal priming experiment. *International Journal of Language & Communication Disorders*, 52(3), 285-300.
- Munson, B., Kurtz, B., & Windsor, J. (2005). The influence of vocabulary size, phonotactic probability, and wordlikeness on nonword repetitions of children with and without specific language impairment. *Journal of Speech Language and Hearing Research*, 48(5), 1033-1047.
- Munson, B., Swenson, C., & Manthei, S. (2005). Lexical and Phonological Organization in Children: Evidence From Repetition Tasks. *Journal of Speech, Language, and Hearing Research*, 48(1), 108.
- Murray, E., McCabe, P., & Ballard, K.J. (2012). A comparison of two treatments for childhood apraxia of speech: methods and treatment protocol for a parallel group randomised control trial. *BMC Pediatrics*, 12, 112.
- Newman, R.S., & German, D.J. (2002). Effects of lexical factors on lexical access among typical language-learning children and children with word-finding difficulties. *Language and Speech*, 45, 285-317.
- Newmeyer, A.J., Aylward, C., Akers, R., Ishikawa, K., Grether, S., deGrauw, T., . . . White, J. (2009). Results of the sensory profile in children with suspected childhood apraxia of speech. *Physical and Occupational Therapy in Pediatrics*, 29(2), 203-218.
- Newmeyer, A.J., Grether, S., Grasha, C., White, J., Akers, R., Aylward, C., . . . deGrauw, T. (2007). Fine motor function and oral motor imitation skills preschool-age children with speech sound disorders. *Clinical Pediatrics*, 46(7), 604-611.
- Nijland, L. (2009). Speech perception in children with speech output disorders. *Clinical Linguistics and Phonetics*, 23(3), 222-239.
- Nijland, L., & Maassen, B. (2005). Syllable planning and motor programming deficits in developmental apraxia of speech. In R. J. Hartsuiker, R. Bastiaanse, A. Postma & F. Wijnen (Eds.), *Phonological Encoding and Monitoring in Normal and Pathological Speech*. New York: Psychological Press.
- Nijland, L., Maassen, B., & van der Meulen, S. (2003). Evidence of motor programming deficits in children diagnosed with DAS. *Journal of Speech, Language and Hearing Research*, 46(2), 437-450.
- Nijland, L., Maassen, B., van der Meulen, S., Gabreels, F., Kraaimaat, F.W., & Schreuder, R. (2002). Coarticulation patterns in children with developmental apraxia of speech. *Clinical Linguistics and Phonetics*, 16(6), 461-483.
- Nijland, L., Maassen, B., van der Meulen, S., Gabreels, F., Kraaimaat, F.W., & Schreuder, R. (2002). Coarticulation patterns in children with developmental apraxia of speech. *Clinical Linguistics and Phonetics*, 16(6), 461-483.

- Nijland, L., Maassen, B., van der Meulen, S., Gabreels, F., Kraaimaat, F.W., & Schrueder, R. (2003). Planning of syllables in children with developmental apraxia of speech. *Clinical Linguistics and Phonetics*, 17(1), 1-24.
- Nijland, L., Terband, H., & Maassen, B. (2015). Cognitive functions in Childhood Apraxia of Speech. *Journal of Speech, Language and Hearing Research*, 58, 550-565.
- Nittrouer, S. (1993). The emergence of mature gestural patterns is not uniform: Evidence from an acoustic study. *Journal of Speech and Hearing Research*, 36(5), 959.
- Nittrouer, S. (2002). From ear to cortex: a perspective on what clinicians need to understand about speech perception and language processing. *Language, Speech and Hearing Services in Schools*, 33, 237-252.
- Nittrouer, S., Studdert-Kennedy, M., & Neely, S.T. (1996). How children learn to organize their speech gestures: Further evidence from fricative-vowel syllables. *Journal of Speech and Hearing Research*, 39(2), 379.
- Odell, K., & Shriberg, L. (2001). Prosody-voice characteristics of children and adults with apraxia of speech. *Clinical Linguistics & Phonetics*, 15(4), 275-307.
- Ozanne, A.E. (1995). The search for developmental verbal dyspraxia. In B. Dodd (Ed.), *Differential diagnosis and treatment of children with speech disorders* (pp. 91-109). London: Whurr.
- Ozanne, A.E. (2005). Childhood Apraxia of Speech. In B. Dodd (Ed.), *Differential Diagnosis and Treatment of Children with Speech Disorders*. West Sussex UK: Whurr Publishers Ltd.
- Pennington, B., & Bishop, D. (2009). Relations among speech, language and reading disorders. *Annual Review of Psychology*, 60, 283-306.
- Perkell, J.S. (1980). Phonetic features and the physiology of speech production. In B. Butterworth (Ed.), *Language Production Volume 1: Speech and Talk* (Vol. 1, pp. 337-372). London: Academic Press.
- Perkell, J.S., Guenther, F.H., Lane, H., Matthies, M.L., Perrier, P., Vick, J., . . . Zandipour, M. (2000). A theory of speech motor control and supporting data from speakers with normal hearing and with profound hearing loss. *Journal of Phonetics*, 28(3), 233-272.
- Perkell, J.S., Matthies, M., Lane, H., Guenther, F.H., Wilhelms-Tricarico, R., Wozniak, J., & Guiod, P. (1997). Speech motor control: Acoustic goals, saturation effects, auditory feedback and internal models. *Speech Communication*, 22(2-3), 227-250.
- Peter, B., Button, L., Stoel-Gammon, C., Chapman, K., & Raskind, W.H. (2013). Deficits in sequential processing manifest in motor and linguistic tasks in a multigenerational family with childhood apraxia of speech. *Clin Linguist Phon*, 27(3), 163-191. doi: 10.3109/02699206.2012.736011
- Peter, B., & Stoel-Gammon, C. (2005). Timing errors in two children with suspected childhood apraxia of speech (sCAS) during speech and music-related tasks. *Clinical Linguistics & Phonetics*, 19(2), 67-87.
- Peter, B., & Stoel-Gammon, C. (2008). Central timing deficits in subtypes of primary speech disorders. *Clinical Linguistics & Phonetics*, 22(3), 171 - 198.
- Piske, T. (1997). Phonological organization in early speech production: Evidence for the importance of articulatory patterns. *Speech Communication*, 22(2-3), 279-295.

- Plunkett, K., Karmiloff-Smith, A., Bates, E., Elman, J.L., & Johnson, M.H. (1997). Connectionism and developmental psychology. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 38(1), 53-80.
- Portney, L.G., & Watkins, M.P. (2011). *Foundations of Clinical Research: Application to Practice* (3rd ed.). New Jersey: Pearson Prentice Hall.
- Price, L.R. (2016). *Psychometric Methods: Theory into Practice*: Guildford Publications.
- Raca, G., Baas, B., Kirmani, S., Laffin, J., Jackson, C., Strand, E., . . . Shriberg, L. (2012). Childhood Apraxia of Speech (CAS) in two patients with 16p11.2 microdeletion syndrome. *European Journal of Human Genetics*. doi: 10.1038/ejhg.2012.165
- Ratcliff, R. (1993). Methods for Dealing With Reaction Time Outliers. *Psychological Bulletin*, 114(3), 510-532.
- RCSLT, R.C.o.S.a.L.T. (2011). RCSLT Policy Statement Developmental Verbal Dyspraxia. In RCSLT (Ed.), (pp. 1-60). <http://www.rcslt.org>: RCSLT.
- Roelofs, A. (1997a). Syllabification in speech production: evaluation of WEAVER. *Language and Cognitive Processes*, 12(5/6), 657-693.
- Roelofs, A. (1997b). The WEAVER model of word-form encoding in speech production. *Cognition*, 64(3), 249-284.
- Rueda, M.R., Fan, J., McCandliss, B.D., Halparin, J.D., Gruber, D.B., Lercari, L.P., & Posner, M.I. (2004). Development of attentional networks in childhood. *Neuropsychologia*, 42(8), 1029-1040.
- Ruscello, D.M. (2012). Is the amount of practice an important variable in the treatment of Childhood Apraxia of Speech (CAS): The importance of production frequency in therapy for childhood apraxia of speech. *Evidence-Based Communication Assessment and Intervention*, 6(2), 66-75. doi: 10.1080/17489539.2012.692966
- Rvachew, S. (2005). Stimulability and treatment success. *Topics in Language Disorders*, 25(3), 207-219.
- Rvachew, S., Hodge, M., & Ohberg, A. (2005). Obtaining and interpreting maximum performance tasks from Children: a tutorial. *Journal of Speech-Language Pathology and Audiology*, 29(4), 146-157.
- Rvachew, S., & Jamieson, D.G. (1989). Perception of Voiceless Fricatives by Children with a Functional Articulation Disorder. *Journal of Speech and Hearing Disorders*, 54(2), 193-208.
- Schriefers, H., Meyer, A.S., & Levelt, W.J.M. (1990). Exploring the time course of lexical access in language production; picture-word interference studies. *Journal of Memory and Language*, 29, 86-102.
- Sealey, L.R., & Giddens, C.L. (2010). Aerodynamic indices of velopharyngeal function in childhood apraxia of speech. *Clin. Linguist. Phon.*, 24(6), 417-430. doi: 10.3109/02699200903447947
- Seidenberg, M.S. (1997). Language acquisition and use: learning and applying probabilistic constraints. *Science*, 275, 1599-1603.
- Seiger-Gardner, L., & Brooks, P.J. (2008). Effects of onset and rhyme related distractors on phonological processing in children with Specific Language Impairment. *Journal of Speech, Language and Hearing Research*, 51(5), 1263-1281.
- Shriberg, L. (2003). Diagnostic markers for child speech-sound disorders: introductory comments. *Clinical Linguistics and Phonetics*, 17(7), 501-505.

- Shriberg, L. (2010). A neurodevelopmental framework for research in childhood apraxia of speech. In B. Maassen & P. van Lieshout (Eds.), *Speech Motor Control: New developments in basic and applied research* (pp. 259-270). Oxford, England: Oxford University Press.
- Shriberg, L., Aram, D., & Kwiatkowski, J. (1997a). Developmental apraxia of speech: I. Descriptive and theoretical perspectives. *Journal of Speech, Language and Hearing Research*, 40(2), 273-285.
- Shriberg, L., Aram, D., & Kwiatkowski, J. (1997b). Developmental apraxia of speech: II. Toward a diagnostic marker. *Journal of Speech, Language, and Hearing Research*, 40(2), 286-312.
- Shriberg, L., Aram, D., & Kwiatkowski, J. (1997c). Developmental apraxia of speech: III. A subtype marked by inappropriate stress. *Journal of Speech, Language, and Hearing Research*, 40(2), 313-337.
- Shriberg, L., Austin, D., Lewis, B., McSweeney, J., & Wilson, D. (1997). The Speech Disorders Classification System (SDCS): Extensions and lifespan reference data. *Journal of Speech Language and Hearing Research*, 40(4), 723-740.
- Shriberg, L., Ballard, K., Tomblin, J., Duffy, J., Odell, K., & Williams, C. (2006). Speech, prosody, and voice characteristics of a mother and daughter with a 7;13 translocation affecting FOXP2. *Journal of Speech, Language and Hearing Research*, 49, 500-525.
- Shriberg, L., Campbell, T., Karlsson, H., Brown, R., McSweeney, J., & Nadler, C. (2003). A diagnostic marker for childhood apraxia of speech: the lexical stress ratio. *Clinical Linguistics and Phonetics*, 17(7), 549-574.
- Shriberg, L., Fourakis, M., Hall, M., Karlsson, H., Lohmeier, H., McSweeney, J., . . . Wilson, D. (2010a). Perceptual and acoustic reliability estimates for the Speech Disorders Classification System (SDCS). *Clinical Linguistics & Phonetics*, 24(10), 825-846.
- Shriberg, L., Fourakis, M., Hall, S., Karlsson, H., Lohmeier, H., McSweeney, J., . . . Wilson, D. (2010b). Extensions to the Speech Disorders Classification System (SDCS). *Clinical Linguistics & Phonetics*, 24(10), 795-824.
- Shriberg, L., Green, J., Campbell, T., McSweeney, J., & Scheer, A. (2003). A diagnostic marker for childhood apraxia of speech: the coefficient of variation ratio. *Clinical Linguistics and Phonetics*, 17(7), 575-595.
- Shriberg, L., Jakielski, K., & El-Shanti, H. (2008). Breakpoint localization using array-CGH in three siblings with an unbalanced 4q;16q translocation and childhood apraxia of speech (CAS). *Am. J. Med. Genet. A*, 146A(17), 2227-2233. doi: 10.1002/ajmg.a.32363
- Shriberg, L., & Kwiatkowski, J. (1982). Phonological Disorders III: A Procedure for Assessing Severity of Involvement. *Journal of Speech and Hearing Disorders*, 47(3), 256-270.
- Shriberg, L., Lohmeier, H., Campbell, T., Dollaghan, C., Green, J., & Moore, C. (2009). A nonword repetition task for speakers with misarticulations: the syllable repetition task (SRT). *Journal of Speech, Language and Hearing Research*, 52, 1189-1212.
- Shriberg, L., Lohmeier, H., Strand, E., & Jakielski, K. (2012). Encoding, memory and transcoding deficits in childhood apraxia of speech. *Clinical Linguistics & Phonetics*, 26(5), 445-482.
- Shriberg, L., Paul, R., Black, L., & van Santen, J. (2011). The Hypothesis of Apraxia of Speech in Children with Autism Spectrum Disorder. *J. Autism Dev. Disord.*, 41(4), 405-426. doi: 10.1007/s10803-010-1117-5

- Shriberg, L., Potter, N., & Strand, E. (2011). Prevalence and phenotype of Childhood Apraxia of Speech in Youth with galactosemia. *Journal of Speech, Language and Hearing Research*, 54, 487-519.
- Skinder, A., Connaghan, K., Strand, E., & Betz, S. (2000). Acoustic correlates of perceived lexical stress errors in children with developmental apraxia of speech. *Journal of Medical Speech-Language Pathology*, 8(4), 279-284.
- Skinder, A., Strand, E., & Mignerey, M. (1999). Perceptual and acoustic analysis of lexical and sentential stress in children with developmental apraxia of speech. *Journal of Medical Speech-Language Pathology*, 7(2), 133-144.
- Stackhouse, J., & Snowling, M. (1992). Developmental Verbal Dyspraxia .2. A Developmental Perspective on 2 Case-Studies. *European Journal of Disorders of Communication*, 27(1), 35-54.
- Stackhouse, J., & Wells, B. (1993). Psycholinguistic assessment of developmental speech disorders. *European Journal of Cognitive Psychology*, 25(3-22).
- Stackhouse, J., & Wells, B. (1997). *Children's Speech and Literacy Difficulties* (Vol. 1). London: Whurr Publishers.
- Sternberg, S., Knoll, R.L., Monsell, S., & Wright, C.E. (1988). Motor programs and hierarchical organisation in the control of rapid speech. *Phonetica*, 45, 175-197.
- Storkel, H. (2001). Learning new words: Phonotactic probability in language development. *Journal of Speech, Language, and Hearing Research*, 44(6), 1321.
- Storkel, H. (2002). Restructuring of similarity neighbourhoods in the developing mental lexicon. *Journal of Child Language*, 29(2), 251-274.
- Storkel, H., & Rogers, M. (2000). The effect of probabilistic phonotactics on lexical acquisition. *Clinical Linguistics & Phonetics*, 14(6), 407-425. doi: doi:10.1080/026992000415859
- Strand, E., & Debertine, P. (2000). The efficacy of integral stimulation intervention with developmental apraxia of speech. *Journal of Medical Speech-Language Pathology*, 8(4), 295-300.
- Sussman, H., Marquardt, T., & Doyle, J. (2000). An acoustic analysis of phonemic integrity and contrastiveness in developmental apraxia of speech. *Journal of Medical Speech-Language Pathology*, 8(4), 301-313.
- Swan, D., & Goswami, U. (1997). Picture naming deficits in developmental dyslexia: The phonological representations hypothesis. *Brain and Language*, 56(3), 334-353.
- Tabachnick, B., & Fidell, L. (2007). *Using Multivariate Statistics*. Boston MA: Allyn and Bacon.
- Templin, M., & Darley, F. (1969). *The Templin-Darley Tests of Articulation*. City of Iowa: University of Iowa Press.
- Terband, H., & Maassen, B. (2010a). Speech Motor Development in Childhood Apraxia of Speech: Generating Testable Hypotheses by Neurocomputational Modeling. *Folia Phoniatri. Logop.*, 62(3), 134-142. doi: 10.1159/000287212
- Terband, H., & Maassen, B. (2010b). Speech motor development in childhood apraxia of speech: generating testable hypotheses by neurocomputational modeling. *Folia Phoniatrica Et Logopaedica*, 62(3), 134-142.
- Terband, H., Maassen, B., Guenther, F., & Brumberg, J. (2009). Computational Neural Modeling of Speech Motor Control in Childhood Apraxia of Speech (CAS). *J. Speech Lang. Hear. Res.*, 52(6), 1595-1609. doi: 10.1044/1092-4388(2009/07-0283)

- Terband, H., Maassen, B., van Lieshout, P., & Nijland, L. (2011). Stability and composition of functional synergies for speech movements in children with developmental speech disorders. *Journal of Communication Disorders*, 44, 59-74.
- Terband, H., Maassen, B., Guenther, F.H., and Brumberg, J. (2009). Computational neural modeling of speech motor control in childhood apraxia of speech CAS. *Journal of Speech, Language and Hearing Research*, 52(6), 1595-1609.
- Terband, H.R., & Maassen, B. (2010c). Speech motor development in childhood apraxia of speech: generating testable hypotheses by neurocomputational modeling. *Folia Phoniatrica Et Logopaedica*, 62, 134-7762.
- Teverovsky, E., Bickel, J., & Feldman, H. (2009a). Functional characteristics of children diagnosed with Childhood Apraxia of Speech. *Disability and Rehabilitation*, 31(2), 94-102.
- Teverovsky, E.G., Bickel, J.O., & Feldman, H.M. (2009b). Functional characteristics of children diagnosed with Childhood Apraxia of Speech. *Disability and Rehabilitation*, 31(2), 94-102. doi: 10.1080/09638280701795030
- Thelen, E. (2005). Dynamic systems theory and the complexity of change. *Psychoanalytic Dialogues*, 15(2), 255-283.
- Thelen, E., & Bates, E. (2003). Connectionism and dynamic systems: are they really different? *Developmental Science*, 6(4), 378-391.
- Thevenon, J., Callier, P., Andrieux, J., Delobel, B., David, A., Sukno, S., . . . Faivre, L. (2013). 12p13.33 microdeletion including ELKS/ERC1, a new locus associated with childhood apraxia of speech. *Eur. J. Hum. Genet.*, 21(1), 82-88. doi: 10.1038/ejhg.2012.116
- Thoonen, G., Maassen, B., Gabreels, F., & Schreuder, R. (1994). Feature Analysis of Singleton Consonant Errors in Developmental Verbal Dyspraxia (DVD). *Journal of Speech and Hearing Research*, 37(6), 1424-1440.
- Thoonen, G., Maassen, B., Gabreels, F., & Schreuder, R. (1999). Validity of maximum performance tasks to diagnose motor speech disorders in children. *Clinical Linguistics & Phonetics*, 13(1), 1-23.
- Thoonen, G., Maassen, B., Gabreels, F., Schreuder, R., & deSwart, B. (1997). Towards a standardised assessment procedure for developmental apraxia of speech. *European Journal of Disorders of Communication*, 32(1), 37-60.
- Thoonen, G., Maassen, B., Wit, J., Gabreels, F., & Schreuder, R. (1996). The integrated use of maximum performance tasks in differential diagnostic evaluations among children with motor speech disorders. *Clinical Linguistics & Phonetics*, 10(4), 311-336.
- Tourville, J., & Guenther, F. (2011). The DIVA model: a neural theory of speech acquisition and production. *Language and Cognitive Processes*, 26(7), 952-981.
- Trochim, W., & Donnelly, J. (2007). *The Research Methods Knowledge Base* (3rd ed.). Cincinnati, OH: Atomic Dog Publishing Inc.
- Truman, A., & Hennessey, N. (2006). The locus of naming difficulties in children with dyslexia: Evidence of inefficient phonological encoding. *Language and Cognitive Processes*, 21(4), 361-393.
- Tsao, F., Liu, H., & Kuhl, P. (2004). Speech perception in infancy predicts language development in the second year of life: A longitudinal study. *Child Development*, 75(4), 1067-1084.

- Van der Merwe, A. (1997). A theoretical framework for the characterisation of pathological speech sensorimotor control. In M. R. Mc Neil (Ed.), *Clinical Management of Sensorimotor Speech Disorders* (pp. 1-25). New York: Thieme.
- Velleman, S., & Shriberg, L. (1999). Metrical analysis of the speech of children with suspected developmental apraxia of speech. *Journal of Speech, Language, and Hearing Research*, 42(6), 1444-1460.
- Velleman, S., & Vihman, M. (2002). Whole-word phonology and templates: Trap, bootstrap, or some of each? *Language Speech and Hearing Services in Schools*, 33(1).
- Vihman, M.M., & Velleman, S.L. (2000). The construction of a first phonology. *Phonetica*, 57(2-4), 255-266.
- Vitevitch, M.S., & Sommers, M.S. (2003). The facilitative influence of phonological similarity and neighborhood frequency in speech production in younger and older adults. *Memory and Cognition*, 31(4), 491-504.
- Walley, A.C. (1993). The Role of Vocabulary Development in Childrens Spoken Word Recognition and Segmentation Ability. *Developmental Review*, 13(3), 286-350.
- Walley, A.C., Michela, V.L., & Wood, D.R. (1995). The Gating Paradigm - Effects of Presentation Format on Spoken Word Recognition by Children and Adults. *Perception & Psychophysics*, 57(3), 343-351.
- Waring, R., & Knight, R. (2013). How should children with speech sound disorders be classified? A review and critical evaluation of current classification systems. *International Journal of Language and Communications Disorders*, 48(1), 25-40.
- Westermann, G., & Miranda, E.R. (2004). A new model of sensorimotor coupling in the development of speech. *Brain and Language*, 89(2), 393-400.
- Williams, P., & Stephens, H. (2010). Nuffield Dyspraxia Program. In A. Williams, S. McLeod & R. McCauley (Eds.), *Interventions for speech sound disorders*. Baltimore:MD: Brooks.
- Williams, K. (1997). *Expressive Vocabulary Test*. Circle Pines, MN: American Guidance Services.
- Worthey, E., Raca, G., Laffin, J., Wilk, B., Harris, J., Jakielski, K., . . . Shriberg, L. (2013). Whole-exome sequencing supports genetic heterogeneity in childhood apraxia of speech.(Research). *Journal of Neurodevelopmental Disorders*, 5, 29.
- Zaretsky, E., Velleman, S.L., & Curro, K. (2010). Through the magnifying glass: Underlying literacy deficits and remediation potential in Childhood Apraxia of Speech. *Int. J. Speech-Lang. Pathol.*, 12(1), 58-68. doi: 10.3109/17549500903216720
- Ziegler, W., & Maassen, B. (2004). The role of the syllable in disorders of spoken language production. In B. Maassen, R. D. Kent, H. F. M. Peters, P. H. H. M. van Lieshout & W. Hulstijn (Eds.), *Speech Motor Control in Normal and Disordered Speech* (pp. 415-447). New York: Oxford University Press.
- Ziegler, W., Staiger, A., & Aichert, I. (2010). Apraxia of speech: what the deconstruction of phonetic plans tells us about the construction of articulate language. In B. Maassen & P. van Lieshout (Eds.), *Speech Motor Control: New developments in basic and applied research*. NY: Oxford University Press.

“Every reasonable effort has been made to acknowledge the owners of copyright material. I would be pleased to hear from any copyright owner who has been omitted or incorrectly acknowledged.”

Appendix A

Tables from Chapter 2: Study 1

Table A1

Stimuli for nonword repetition task with two and three syllable pairs with low versus high frequency target sequences underlined.

Length	Low frequency	High frequency
Two syllable	jugoin	bogib
	<u>moip</u> əd	<u>mæb</u> ɛp
	<u>vug</u> im	<u>vid</u> æg
	nəf <u>æmb</u>	mɪn <u>æmp</u>
	<u>pw</u> ɑgəb	<u>tw</u> ɛkɛt
	bʊf <u>k</u> it	kɪf <u>t</u> en
	dɒg <u>d</u> et	tæk <u>t</u> ut
	motau <u>k</u>	pɛt <u>i</u> k
	dɒn <u>u</u> g	bɛd <u>æ</u> g
	tɛdau <u>m</u>	pɒdau <u>d</u>
	<u>aup</u> təd	ɪp <u>t</u> ən
Three syllable	dʊgn <u>ə</u> təd	tʌgn <u>ə</u> dɪt
	<u>auk</u> pəde	ɪk <u>b</u> əni
	<u>auf</u> təɡɑ	<u>aunt</u> əko
	bɒdə <u>yau</u>	mɛdə <u>ju</u>
	<u>vuk</u> ətɛm	<u>vit</u> əɡɑp
	<u>gaun</u> əpek	<u>git</u> əmo:k
	<u>nɒb</u> əmən	<u>nɪd</u> əbɪp
	kɛdəw <u>əmb</u>	fɪkət <u>æmp</u>
	<u>pw</u> ɛnətɛp	<u>tw</u> ɛdəmin
	næf <u>k</u> ətu	ɡʌft <u>ə</u> dɑɪ
	dɛɡd <u>ə</u> ne	tɪkt <u>ə</u> po

Note. Target sequences are underlined and vary in phonotactic frequency. Stimuli are a replication from Munson et al. (2005).

Table A2

Stimuli for picture-naming task with related and unrelated auditory distractor rimes.

Item No.	Target	Freq	Related IW	Unrelated IW
Test Items:				
1	Ant	6	Add	Bed
2	Kite	1	Kind	Food
3	Sock	4	Song	Pit
4	Button	1	Butler	Skipper
5	Turtle	8	Turnip	Motor
6	Pedal	4	Pencil	Wafer
7	Cage	9	Case	Shoe
8	Nail	6	Name	Den
9	Hose	9	Hold	Pin
10	Puppy	2	Puddle	Table
11	Bucket	7	Bundle	Waddle
12	Hammer	9	Handle	Jumper
13	Goat	6	Ghost	Tool
14	Hook	5	Hood	Bun
15	Peg	4	Pet	Rat
16	Kitten	5	Kitchen	Paddle
17	Pillow	8	Pistol	Medal
18	Camel	1	Castle	Finger
Practice Items:				
1	Bin	9	Bit	Walk
2	Duck	9	Dump	Time

Note. Low frequency rating 1-4 (inclusive), high frequency rating 5-9 (inclusive).

Table A3

Stimuli for simple and choice verbal reaction time tasks.

	Stimuli			
	1 Syllable	KF	2 Syllable	KF
Pair 1	Foot	70	Football	36
Pair 2	News	102	Newspaper	65
Pair 3	Cart	5	Cartwheel	-
Pair 4	Pig	8	Piglet	-
Pair 5	Doll	10	Dollhouse	-
Pair 6	Pea	-	Peanut	6
Pair 7	Bed	127	Bedroom	52
Pair 8	Cow	29	Cowboy	16

Note. KF = Kucera Francis Frequency rating

Table A4
Unstandardized (B) and standardized (β) Regression Coefficients, and Squared Part Correlations (sr^2) for a Hierarchical Multiple Regression Model Predicting Mean Nonword Repetition Accuracy for the Low Frequency Sequences (NWRepLow) from Simple Reaction Time (SVRT), Chronological Age, and the Chronological Age x SVRT Interaction ($N = 53$).

Predictors (IVs)	B	95% CI ¹	β	sr^2	p-value ¹
DV: NWRepLow					
Step 1					
EVT raw	-0.048	-0.180, 0.084	-.126	.006	.468
PPVT raw	0.132	0.037, 0.226	.640	.138	.007*
SVRT	0.000	-0.013, 0.014	.010	.001	.948
$R^2 = .295, p = .002^{**}$					
Step 2					
EVT raw	-0.092	-0.257, 0.072	-.244	.017	.263
PPVT raw	0.114	0.012, 0.217	.556	.090	.029*
SVRT	0.001	-0.013, 0.014	.018	.000	.902
ChronAge	0.061	-0.073, 0.195	.230	.015	.363
ChronAge: $F(1, 42) = 0.85, p = .363$					
$\Delta R^2 = .015, p = .337$					
$R^2 = .310, p = .003^{**}$					
Step 3					
EVT raw	-0.094	-0.267, 0.079	-.249	.014	.277
PPVT raw	0.138	0.012, 0.217	.556	.090	.030*
SVRT	7.943	-0.078, 0.076	.018	.000	.986
ChronAge	0.033	-0.072, 0.197	.235	.013	.356
ChronAge: $F(1, 41) = 0.87, p = .356^2$					
ChronAge x SVRT ³	-0.176	-0.079, 0.083	.006	.000	.966
ChronAge x SVRT: $F(1, 41) = 0.00, p = .966^4$					
$\Delta R^2 = .000, p = .968$					
$R^2 = .310, p = .008^{**}$					

1: These are the GLMM adjusted values.

2: This is the overall F -value for the group effect.

3: The interaction term is computed using centered SVRTs.

4: This is the overall F -value for the Chronological Age x SVRT interaction effect.

Table A5

Unstandardized (B) and standardized (β) Regression Coefficients, and Squared Part Correlations (sr^2) for a Hierarchical Multiple Regression Model Predicting Mean Nonword Repetition Accuracy for the Low Frequency Sequences (NWRepLow) from Choice Verbal Reaction Time (CVRT), Chronological Age (ChronAge), and the ChronAge x CVRT Interaction (N = 53).

Predictors (IVs)	B	95% CI ¹	β	sr^2	p-value ¹
DV: NWRepLow					
Step 1					
EVT raw	-0.054	-0.170, 0.061	-.143	.008	.347
PPVT raw	0.127	0.042, 0.213	.619	.144	.004**
CVRT	-0.002	-0.010, 0.006	-.070	.004	.608
$R^2 = .298, p = .002^{**}$					
Step 2					
EVT raw	-0.091	-0.248, 0.066	-.240	.017	.247
PPVT raw	0.112	0.017, 0.208	.547	.010	.023*
CVRT	0.000	-0.009, 0.999	.001	.000	.996
ChronAge	0.061	-0.092, 0.213	.229	.011	.425
ChronAge: $F(1, 42) = 0.65, p = .425$					
$\Delta R^2 = .011, p = .408$					
$R^2 = .310, p = .003^{**}$					
Step 3					
EVT raw	-0.097	-0.254, -0.061	-.255	.018	.222
PPVT raw	0.112	0.015, 0.208	.543	.095	.024*
CVRT	-0.010	0.060, 0.041	.004	.000	.708
ChronAge	0.064	-0.087, 0.215	-.241	.013	.398
ChronAge: $F(1, 41) = 0.73, p = .398^2$					
ChronAge x CVRT ³	-0.017	-0.068, 0.103	.054	.003	.684
ChronAge x CVRT: $F(1, 41) = 0.17, p = .684^4$					
$\Delta R^2 = .003, p = .686$					
$R^2 = .313, p = .007^{**}$					

1: These are the GLMM adjusted values.

2: This is the overall F -value for the group effect.

3: The interaction term is computed using centered CVRTs.

4: This is the overall F -value for the Chronological Age x ChRT interaction effect.

Table A6

Unstandardized (B) and standardized (β) Regression Coefficients, and Squared Part Correlations (sr^2) for a Hierarchical Multiple Regression Model Predicting the Frequency Effect from Simple Reaction Time (SimRT), Chronological Age, and the Chronological Age x SimRT Interaction (N = 53).

Predictors (IVs)	B	95% CI ¹	β	sr^2	p-value ¹
DV: FreqEff					
Step 1					
PPVT raw	-0.076	-0.139, -0.013	-.448	.163	.019
SVRT	-0.001	-0.012, 0.010	-.029	.001	.836
$R^2 = .190, p = .010^*$					
Step 2					
PPVT raw	-0.062	-0.156, 0.032	-.365	.048	.190
SVRT	-0.001	-0.012, 0.009	-.029	.001	.834
ChronAge	-0.023	-0.116, 0.069	-.106	.004	.615
ChronAge: $F(1, 43) = 0.26, p = .615$					
$\Delta R^2 = .004, p = .630$					
$R^2 = .195, p = .024^*$					
Step 3					
PPVT raw	-0.069	-0.159, 0.021	-.406	.056	.130
SVRT	-0.024	-0.080, 0.033	-.035	.001	.408
ChronAge	-0.016	-0.105, 0.072	-.075	.002	.711
ChronAge: $F(1, 42) = 0.14, p = .711^2$					
ChronAge x SVRT ³	0.025	-0.033, 0.083	-.106	.010	.383
ChronAge x SVRT: $F(1, 42) = 0.78, p = .383^4$					
$\Delta R^2 = .011, p = .456$					
$R^2 = .206, p = .042^*$					

1: These are the GLMM adjusted values.

2: This is the overall F -value for the group effect.

3: The interaction term is computed using centered SVRTs.

4: This is the overall F -value for the Group x simple reaction time interaction effect.

Table A7

Unstandardized (B) and standardized (β) Regression Coefficients, and Squared Part Correlations (sr^2) for a Hierarchical Multiple Regression Model Predicting the Frequency Effect from Choice Reaction Time (ChRT), Chronological Age (ChronAge), and the ChronAge x ChRT Interaction (N = 53).

Predictors (IVs)	B	95% CI ¹	β	sr^2	p-value ¹
DV: FreqEff					
Step 1					
PPVT raw	-0.076	-0.140, -0.011	-.445	.160	.022*
CVRT	-0.001	-0.006, 0.005	-.021	.000	.849
$R^2 = .190, p = .010^*$					
Step 2					
PPVT raw	-0.058	-0.156, 0.040	-.340	.045	.240
CVRT	-0.002	-0.008, 0.004	-.082	.003	.492
ChronAge	-0.037	-0.141, 0.068	-.168	.008	.483
ChronAge: $F(1, 43) = 0.50, p = .483$					
$\Delta R^2 = .008, p = .516$					
$R^2 = .198, p = .022^*$					
Step 3					
PPVT raw	-0.080	-0.157, -0.002	.340	.044	.044*
CVRT	-0.000	-0.030, 0.030	-.082	.003	.990
ChronAge	0.307	-2.353, 1.739	-.168	.008	.763
ChronAge: $F(1, 42) = 0.09, p = .763^2$					
ChronAge x CVRT ³	0.000	-0.053, 0.054	-.002	.000	.996
ChronAge x CVRT : $F(1, 42) = 0.00, p = .996^4$					
$\Delta R^2 = .000, p = .988$					
$R^2 = .198, p = .050^*$					

1: These are the GLMM adjusted values.

2: This is the overall F -value for the group effect.

3: The interaction term is computed using centered CVRTs.

4: This is the overall F -value for the Group x choice reaction time interaction effect.

Table A8

Unstandardized (B) and standardized (β) Regression Coefficients, and Squared Part Correlations (sr^2) for a Hierarchical Multiple Regression Model Predicting Speech Discrimination Ability (D-prime) from GFTA Raw Scores, Chronological Age (ChronAge), and the Chronological Age x GFTA Interaction (N = 53)

Predictors (IVs)	B	95% CI ¹	β	sr^2	p-value ¹
DV: d-prime					
Step 1					
EVT raw	0.002	-0.004, 0.008	.093	.003	.450
PPVT raw	0.004	0.001, 0.008	.287	.029	.110
GFTA raw	-0.002	-0.042, 0.038	-.013	.000	.909
$R^2 = .136, p = .095$					
Step 2					
EVT raw	-0.002	-0.008, 0.004	-.069	.001	.583
PPVT raw	0.002	-0.003, 0.007	.176	.009	.350
GFTA raw	0.001	-0.042, 0.044	.007	.000	.953
ChronAge	0.005	-0.001, 0.012	.315	.029	.111
ChronAge: $F(1, 42) = 2.66, p = .111$					
$\Delta R^2 = .029, p = .237$					
$R^2 = .165, p = .102$					
Step 3					
EVT raw	-0.002	-0.008, 0.005	-.071	.001	.604
PPVT raw	0.002	-0.003, 0.007	.176	.009	.356
GFTA raw	0.007	-0.224, 0.237	.000	.000	.954
ChronAge	0.005	-0.001, 0.012	.313	.027	.123
ChronAge: $F(1, 41) = 2.48, p = .123^2$					
ChronAge x GFTA ³	-0.000	-0.006, 0.005	-.008	.000	.964
ChronAge x GFTA : $F(1, 41) = 0.00, p = .964^4$					
$\Delta R^2 = .000, p = .981$					
$R^2 = .165, p = .177$					

1: These are the GLMM adjusted values.

2: This is the overall F -value for the group effect.

3: These are the centered GFTA raw scores.

4: This is the overall F -value for the Chronological Age x GFTA interaction effect.

* $p < .05$; ** $p < .01$

Appendix B

Tables from Chapter 3

Table B1

Studies Excluded from Systematic Review

	<i>Study</i>	<i>Reason for exclusion</i>	<i>Code</i>
1.	Hodge, 1993	Not specific to CAS – assessment and treatment of a child with a developmental speech disorder	1
2.	Crary, 1995	Clinical evaluation of developmental motor speech disorders; no experimental groups.	2
3.	Strand, 1995	Review of treatment of motor speech disorders – no experimental data and not specific to CAS	2
4.	Ozanne (1995)	Participants suspected of having motor speech disorder	1
5.	Shriberg, Aram & Kwiatkowski (1997a)	Descriptive and theoretical perspectives on sCAS – no specific list of features provided in relation to classification of subjects.	3
6.	Shriberg, Aram & Kwiatkowski (1997b)	No specific features nominated in relation to classification of CAS to experimental group.	3
7.	Shriberg, Austin, Lewis, McSweeney and Wilson 1997	Speech disorder classification system: no experimental groups – not specific to CAS.	2
8.	McCabe, Rosenthal and McLeod, 1998	Clinical population was speech impairment – no assignment to specific groups such as CAS	1
9.	Forrest and Morrisette, 1999	Experimental group - PD	1

10.	Hall, 2000	Speech characteristics of CAS: no experimental groups / data	2
11.	Bahr and Velleman (1999)	No experimental groups – no participant criteria	2
12.	Knock, Ballard, Robin and Schmidt, 2000	Acquired apraxia of speech	4
13.	Lewis et al. (2000)	Broad analysis of outcomes of children with SSD not specific to CAS	1
14.	Strand, 2001	Synopsis of Darley's contribution to CAS; no experimental groups/ data.	2
15.	Fox, Dodd & Howard (2002)	Not specific to CAS	1
16.	Forrest 2003	Diagnostic criteria of CAS: participants were SLPs	2
17.	Shriberg (2003)	Diagnostic markers for SSD – not specific to CAS	1
18.	Rvachew, Hodge and Ohberg (2005)	A tutorial on obtaining and interpreting maximum performance tasks from children.	2
19.	Lewis et al. (2006)	Factor analysis used to classify as SSD – no reference to CAS	1
20.	Lewis et al. (2006)	Classification of SSD not specific to CAS	1
21.	Shriberg et al. (2006)	Adults with AOS (50 y.o and 18 y.o.)	4

22.	Caspari et al. (2008)	No experimental data – discussion on relationship between sleep apnea, tonsillectomy and speech	2
23.	Newmeyer et al. (2007)	Not specific to CAS – no classification no features	1
24.	Gildersleeve- Neumann 2007	No experimental data – a description of integral stimulation and motor learning	2
25.	McCauley and Strand (2008)	No experimental data included in study	2
26.	Shriberg, Jakielski and El-Shanti (2008)	Genetic study on family members – all children pre-diagnosed and no mention of classification criteria.	3
27.	McLeod and Harrison 2009	Epidemiology study of Speech and Language impairment: not specific to CAS	1
28.	Shriberg et al. (2009)	Speech disorders – not specific to CAS	1
29.	Teverovsky, Bickel and Feldman (2009b)	Pre-diagnosed participants – no features reported.	3
30.	Shriberg et al. (2010a)	SSD not specific to CAS	1
31.	Shriberg et al. (2010b)	SSD not specific to CAS	1
32.	Terband and Maassen (2010b)	Authors advocated a modeling approach not consistent with classification.	3

33.	Jacks, Mathes & Marquardt (2010)	Adults with AOS	4
34.	Sealy and Giddens (2010)	Participants pre-diagnosed with CAS, verified by Kaufman Speech Praxis Test for Children but no features reported.	3
35.	Zaretsky, Velleman and Curro (2010)	Longitudinal study of one subject with CAS and borderline IQ. Deficits not specific to CAS and features of speech not classified as such.	3
36.	Shriberg, Paul, Black and van Santen (2011)	Pre-diagnosed as CAS – specific features not reported.	3
37.	Lewis et al. (2011)	Classified participants based on severity of SSD not presence of features.	3
38.	Shriberg et al. (2011)	Participants not classified according to features.	3
39.	Raca et al. (2012)	Participants not classified according to features. Madison speech assessment used but no features reported.	3
40.	Lewis et al. (2012)	Not specific to CAS – no classification other than severity of SSD	1
41.	Highman, Leitao, Hennessey & Piek (2012)	Participants diagnosed as CAS – features not specified.	3
42.	Maas and Farinella (2012)	Participants diagnosed as CAS – features not specified.	3
43.	Marignier et al. (2012)	Single subject study of child with cri du chat syndrome. History of delayed speech but no specific characteristics reported in relation to CAS.	3
44.	Button, Peter, Gammon and Raskind (2013)	Genetic study on family members – all participants prediagnosed and classification criteria not reported.	3

45.	Worthy et al. (2013)	Genetic study on family members – all participants pre=diagnosed and classification criteria not reported.	3
46.	Thevenon et al. (2013)	Genetic study on family members – all participants pre-diagnosed and classification criteria not reported.	3
47.	Peter et al. (2013)	Participants not specified according to features.	3
48.	Waring and Knight (2013)	Evaluation of current classification system for SSD; not specific to CAS and no experimental groups.	2

Note. 1 = not specific to CAS; 2 = no experimental groups; 3 = no features described to assign children to CAS; 4 = Apraxia Of Speech.

Table B2

Details of Studies Included in Systematic Review

	Author(s)	Features	Operational Definition	Y/N	No. of Features	No of Subjects	Age Range	Control Groups
1	Marion, Sussman & Marquardt, (1993) MSM	Diagnosis based on the following: 95% probability correct assignment to diagnosis of DAS Normal receptive language (within 1/1.5 SD) 2SD below mean on articulation test Consonant and syllable omission Vowel errors Oral peripheral exam indicating difficulty sequencing	Screening Test for DAS (Blakeley, 1983) TACL-R (Carrow-Woolfolk, 1985) PPVT-R (Dunn, 1981) Templin-Darley Screen Test of Articulation (Templin and Darley, 1969) OME (Ekelman and Aram, 1984)	Y	4 /4	4	5-7	TD

2	Thoonen, Maassen, Gabreels and Schreuder (1994) TMGSa	Clinical diagnosis determined by SLP plus persistent speech difficulty; Deviant rather than immature articulatory behavior; Poor production of consonants; Speech proficiency dependent on length; Inconsistent patterns of errors; Inability to produce complex phonemic sequences.		N	4/5	11	6.2-7.9 (6.11)	TD
3	Bradford and Dodd (1996) BD	DVD group had breakdown in three levels of speech motor programming: (Ozanne 1995)	Spontaneous speech sample OME (Robins and Klee, 1987)	Y	3/11	51	3.2-6.7	Speech Delay DevCon DevIncon

Phonological planning	Word Inconsistency test
Vowel errors	(Dodd, 1995) $\geq 40\%$
Polysyllabic errors	
Phrasal errors	
Poor phonotactics	
Inconsistent articulation	
Phonetic planning	
Groping	
Consonant deletion	
Voluntary v's involuntary	
Oro/speech motor	
DDK rate	
DDK sequence	
Poor oromotor	

4	Groenen, Maassen, Crul & Thoonen (1996) GMCT	Diagnosis based on presentation of following features (with 100% agreement between two SLPs):	Audio recordings of spontaneous speech and sentence imitations.	N	X/8	17	6.11-11.6 (M=8.9)	TD (16)
		Periods of highly unintelligible speech;						
		Difficulty to produce complex sequences;						
		High incidence of context related sound substitutions (metathetic errors);						
		Inconsistent speech;						
		Normal OME;						
		Normal IQ;						
		No ADD;						
		Normal hearing.						

5	Thoonen, Maassen, Wit, Gabreels and Schreuder (1996) TMWGS	Evaluation made by SLP based on Hall (1992)	NO	6/6	11 (CAS) 9 (DYS) 11 (TD)	6.3-7.9 6.4-10.3 6.0-8.3	DYS TD
		High rate of speech sound errors; Groping; Periods of highly unintelligible speech; Difficulty/ inability to produce complex sequences; High incidence of context related errors; Inconsistent speech performance					

6	Thoonen et al. (1997) TMGSD	Participants with clear diagnosis of DAS by SLPs based on features: Deviant speech; Poor production of consonants and vowels; Sequencing difficulties; Inconsistent error patterns; Inability to produce complex sequences.		N	4/5	11	6.2-7.9 (6.11)	TD
7	Shriberg, Aram & Kwiatkowski (1997c) SAK	Children suspected of having CAS were further analyzed via speech samples sent from clinics. Data from samples showed: Inconsistent productions; Numerous vowel errors; Unusual and persisting errors;	SSD – features reported per subject PCC from speech sample	N	X/10	19	4.7-14.4	-

Intonation and stress
 inconsistencies;
 Poor performance on DDK tasks; DDK
 Gap between EVT and PPVT; EVT/ PPVT
 Inappropriate loudness;
 Unable to imitate oral
 movements;
 Groping movements;
 Poor progress in therapy.

8	Davis Jakielski and Marquardt (1998) DJMa	Subjects recruited following DAS workshop using criteria below:	SSD methodology Diagnostic protocol: Spontaneous speech sample	N	8/11	5	3.2-5.7	-
		Limited consonant and vowel phonemic repertoire; Frequent omission errors; High incidence of vowel errors;	GFTA DDK OME Normal oral and limb praxis					

Inconsistent articulation errors;
 Suprasegmental errors;
 Difficulty imitating words;
 Increased errors with increased
 complexity;
 Simple syllable shapes.

9	Thoonen, Maassen, Gabreels and Schreuder (1999) TMGSb	See Thoonen et al., 1996 (above)		N	6/6	11 CAS 9 DYS 11TD 11 nSD	6.3-7.9 6.4-10.3 6.0-8.2 4.4-10.11	DYS TD nSD
	Skinder, Strand and Mignerey (1999) SSM	Participants diagnosed as CAS based on:	Assessments included: PPVT OME	N	8 /10	5	5.9-8.8	TD

Limited consonant and vowel
inventory;
Frequent omission errors;
High incidence of vowel errors;
Inconsistent errors;
Altered supra-segmental
characteristics;
Increased error with increased
length;
Difficulty imitating words;
Use of simple syllable shapes;
Expressive skills less than
receptive;
Reduced DDK.

TACL -R

10	Velleman and Shriberg (1999)	See Shriberg et al. 1997. Subjects subsample of previous study (data	Scores on inappropriate stress PVSP and PCC reported per subject.	N	X/10	15	5.8(M)	SD
						8	9.2(M)	SD-DASi
						7	8.9(M)	SD-DASa

	VS	on subjects with sDAS reported individually)						
11	Skinder, Connaghan, Strand and Betz (2000) SCSB	Diagnosis by SLP based on following features (8/11 features needed for diagnosis as having CAS): Limited consonant and vowel inventory; Frequent omission errors; High incidence of vowel errors; Inconsistent errors; Supra-segmental characteristics; Increased error with increased length; Difficulty imitating words; Use of simple syllable shapes;	Additional assessments included: PPVT TACL-R Preschool Language Scale-3	N	8/11	5	4.2-8.2	-

		Expressive skills less than receptive; Reduced DDK.						
12	Strand and Debertine, (2000) SD	SSD	N	X/5	1	5.9	-	
		Glottal stops Vowel distortions Inconsistency Low intelligibility Limited consonant use p, b, t, d, f, v and h (but not in all contexts)	20% to familiar listener					
13	Sussman, Marquardt and Doyle (2000)	Diagnosis based on detailed analysis of speech and language:	Y	X	5	5.6-6.9	TD	

SMD

Receptive vocab >45 percentile
for CAS and TD

Percentile ranking <=5;
Screening test for DAS
(99% probability correct
assignment)

PPVT

GFTA
(Blakely, 1980)

14	Odell and Shriberg (2001) OS	See Shriberg et al.1997; Used subsample from Shriberg 1997 above	N	X	14	4.7-14.4 Adults	Adults with AOS
15	Maassen, Nijland and Van der Meulen (2001)	Clinical criteria from Hall et al. (1993) and Thoonen et al. (1996)		6	6	5.0-5.11	TD

MNV

16	Marquardt, Sussman, Snow & Jacks (2002) MSSJ	Diagnosis made based on symptoms consistent with disorder:	Additional tests: PPVT-R TACL-R (Test of comprehension of Language, Carrow and Woolfolk, 1985)	Y	X	3 (results reported per subject – single subject design)	6-8	TD
		Phonemic repertoire reduced;	Templin-Darley Screening Test of Articulation (Templin and Darley, 1969) GFTA (Goldman and Fristoe, 1984)					
		Inconsistent errors;	Screening Test for Developmental Apraxia of Speech (Blakeley, 1980)		No score provided			
		Difficulty with syllable sequencing;	DDK (Ekelman & Aram, 1984)					

		Orofacial screening normal	OME (Robbins and Klee, 1987)					
		Normal Hearing						
		Intelligibility rating	M = 89.1					
			SD = 11.6					
			Range = 59-99					
		PCC	M = 79.3					
			SD = 13.0					
			Range = 53-95					
19	Shriberg, Green, Campbell, McSweeney and Scheer (2003) SGCMS	Inclusionary criteria were that transcriber perceived speech-timing deficit. Other features included: Excessive/ equal stress;	No operational measures reported.	N	6	15	3-6	SD sDYS TD

Inappropriate timing consistent
with syllable segregation,
Inconsistent errors;
Groping;
Post articulatory repetitions/
revisions;
Metathetic/ sequencing errors.

20	Maassen, Groenen & Crul (2003) MGC	Each child diagnosed by SLP as having apraxic speech problems based on following criteria:	Diagnosis based on spontaneous speech sample and speech and sentence limitations.	N	7	11	6.9-9.5	TD
		High rate of speech sound errors; Inadequate DDK profile for multisyllabic words; Posturing and groping of articulators;						

Periods of highly unintelligible speech;
 Difficulty producing complex sequences;
 Inconsistent speech performance;
 Unequivocally diagnosed as having apraxic speech problems by SP.

21	Munson, Bjorum and Windsor (2003) MBW	Diagnostic features from Davis et al. 1998, included:	Scores reported per subject: GFTA for CAS $\leq 5\%$ GFTA for PD $\leq 25\%$	Y	X	5	3.9-8.10	PD
		Difficulty with volitional movements; Slow DDK rate; Increased error with length						

In addition other features reported in sDAS included compared to PD;
 No operational measures reported.

PD;

Decreased quantity of spontaneous speech;

Awareness that speaking was difficult;

Inconsistent errors;

Decreased accuracy with increased complexity,

Groping;

False starts,

Preponderance of simple syllable shapes;

Misplaced stress in multisyllabic words.

22	Nijland, Maassen, van	As above (Nijland et al. 2002)	N	6	6	4.11-6.10	TD (19)
----	--------------------------	--------------------------------	---	---	---	-----------	---------

der Meulen,
 Gabreels,
 Kraaimaat
 and
 Schreuder
 (2003)
 NMVGKSb

23	Nijland, Maassen and van der Meulen (2003) NMV	Features described by Hall et al. (1993) and Thoonen et al. (1996)	N	5	5	5.0-6.10	TD
		Complete phoneme repertoire with many phonemic errors; High frequency of consonant substitutions and omissions in clusters;					

		Sequencing difficulties; Inconsistent error patterns; Difficulty producing complex sequences.						
24	Hoson, Shriberg, & Green (2004) HSG	Subsample of participants from Shriberg, Campbell et al, 2003 and Shriberg, Green et al., 2003.	N		4	3-8	-	
25	Lewis et al. (2004) LFHTIS	Diagnosis based on: Severely restricted phonemic repertoire; Vowel errors; < 5 th percentile Presence of 3 phonological processes;	Y	5	22 (CAS) 51 (SSD) 42 (SL)	3.0-10.11	SSD SL	
		GFTA < 5 th percentile Khan Lewis Phonological Analysis - 3 processes						

		Phonological analysis rating of 4 (severe); DDK 2 SD below mean	Khan Lewis Phonological Analysis – severity rating 4 Oral speech motor control (Robbins & Klee, 1987)						
26	Marquardt, Jacks and Davis (2004) MJD	Cluster of features including; Test scores for expressive and receptive language reported per subject.	SSD Longitudinal study	N	4	3	4.6-7.7	-	
		Prosodic abnormalities; Vowel errors; High frequency of consonant and syllable omissions; Segmental variability							
27	Lewis et al. (2004) LFHIT	Clinical diagnosis; plus 4 features suggestive of motor programming deficit;		N	4	10 (CAS) 15 (S) 14 (SL)	4.0-6.0	S SL	

Difficulty sequencing sounds/
syllables;

Oral and Speech motor control
Protocol (Robbins and Klee,
1987)

Groping;
Decreased DDK;
Prosodic disturbances;
Metathetic errors;
Consonant deletions;
Increased errors on polysyllabic
words;
Inconsistency on consonants and
vowels;
Normal OME.

DDK task

28	Davis, Jacks and Marquardt (2005) DJMb	Referred by SLP and diagnosis confirmed by 3 SPLs in university setting confirmed diagnosis based on cluster of features:	SSD Longitudinal study	N	All	3	4.6-7.5	-
----	--	--	-------------------------------	---	-----	---	---------	---

Prosodic abnormalities;
 Vowel errors;
 High frequency of consonant and
 syllable omissions;
 Segmental variability.

29	Nijland and Maassen (2005) NM	Clinical criteria described Thoonen et al (1996). See above Nijland et al. 2002 and 2003	N	6	6	5-10	TD
30	Peter and Stoel- Gammon (2005) PSa	Diagnosis by first author based on presence of 8/11 features named by Davis et al 1998. Additional criteria for CAS: Normal cognitive functioning as estimated by receptive vocabulary;	N	8	2	4.3 and 9.5	TD

PPVT

		Normal receptive language;	TACL-3 (Carrow-Woolfolk, 1999)						
		Normal OME;	VMPAC (Hayden and Square, 1999)						
		Normal hearing.							
31	Betz and Stoel- Gammon (2005) BS	Participants a subset of previous study. See Davis et al. 1998	CAS and PD <16% GFTA SSD	N	X	1	4.2 (CAS) 5.10 (PD) 5.0 (TD)	PD TD	
32	Bahr (2005) B	CAS group based on: Slow response in treatment; Groping or struggle; Motor sequencing deficits including speech.	APP-R (Hodson, 1986) LOPT	N	3	5	4.0-7.0	PD TD	

33	Jacks, Marquardt and Davis (2006) JMD	Children referred for differential diagnosis by 3 SLPs based on cluster of speech and language features consistent with diagnosis		N	X	3 (Results reported on individual basis)	4.6-7.7 (longitudin al study over 3 years)	
		Prosodic abnormalities; Vowel errors; High frequency of consonant and syllable omissions Segmental variability						
		OME normal DDK normal.						
34	Moriarty and Gillon (2006) MG	Referred by SLPS as sCAS and classification of CAS based on Ozanne's 1995 Diagnostic	SSD – reported features per subject (including scores on standardized tests)	Y	X	3	6.3-7.3	-

Model, i.e. deficit in three
linguistic levels
phonological planning,
phonetic programming and
motor programming.

Features:

Inconsistency

Low PCC

Vowel errors

Sequencing deficit

Reduced rate

Testing included:

Phonology

BBTOP, Bankson-Bernthal

Test of Phonology, 1990

Consistency of speech

25 Word Consistency Test
(Dodd, 1995)

No score
provided

Severity of speech

PCC<50

			Vowel errors					
			Unintelligibility in connected speech.					
		Receptive language	PPVT-R (Dunn & Dunn, 1981)					
		Expressive language	CELF-4 (Semel et al, 2003)					
		Normal Non verbal intelligence	Test of Nonverbal Intelligence (TONI-2, Brown et al. 1990)					
		Reading ability	Burt Word reading Test					
		Phonological awareness	Phonological awareness Skills Program Test (Rosner, 1999); Preschool Battery of Phonological Awareness, Letter Sound Knowledge (Dodd et al, 2000)					
		Poor sequencing	DDK					
35	Lundeborg and	Single subject treatment study. Child diagnosed as CAS. Presentation of features included:	SSD	N	X	1	5 years	n/a

	McAllister (2007) LM	Restricted sound repertoire Deviant articulation Groping Difficulty imitating sounds Normal receptive grammar and vocab						
36	Peter and Stoel- Gammon (2008) PS	Diagnosis based on CAS characteristics below (11 in total) Limited phoneme inventory; Frequent omission errors; High incidence of vowel errors; Inconsistent articulation errors; Altered suprasegmental characteristics;	SSD – reported features per subject. – not all children had all features (range of features 3-9)	N	X	11	4.7-6.6	TD

Increase errors in longer utterances;
 Difficulty imitating words/ sequences;
 Predominant use of simple syllable shapes;
 Impaired volitional movements;
 Reduced expressive versus receptive language;
 Reduced DDK rate.

37	Highman, Hennessey, Sherwood and Leitao (2008) HHSL	Retrospective study – questionnaire (parent report). Clinical diagnosis based on: Spontaneous speech sample; single word naming; OME; DDK; stimulability of sounds in isolation and syllables. Features present included:	N	X	20	Parents of mothers with clinical diagnosis	SLI TD
----	---	--	---	---	----	--	--------

Limited consonant and vowel
 phonetic inventory;
 Predominant use of simple
 syllable shapes;
 Frequent omission errors;
 High incidence of vowel errors;
 Altered supra-segmental
 characteristics;
 Variability/lack of consistent
 patterns of output;
 Increased error on longer
 sequences;
 Groping/ lack of willingness to
 imitate.

38	McNeill, Gillon and Dodd (2009a)	Treatment Study – 15 SLP administered battery to children with sCAS. Battery included: MGDa	Y	6	12	4-7
----	--	--	---	---	----	-----

Receptive vocabulary within 1.5 SD of mean	PPVT-III (Dunne and Dunn, 1997)
Articulation test SS below 1.5 SD from mean	BBTOP (Bankson –Bernthal Test of Phonology, Bankson and Bernthal, 1990)
Oromotor SS below 8 on all three oromotor subtests or SS below 8 on DDK subtest	DEAP subtest
Inconsistency 40% or greater PCC	DEAP subtest Analysed from 1 st trial of
PVC (percent vowels correct)	DEAP inconsistency subtest
PPU (percent processes usage)	using PROPH (Computerized profiling software, Long and Fey, 2005)
Prosody (stress, loudness, resonance and pitch) – informally evaluated.	Personal narrative collected (following protocol by Westerveld and Gillon, 2002)
Presence of groping	

39	McNeill, Gillon and Dodd (2009b) MGDb	Diagnosis based on CAS features:	Y	3	12	4-8	TD ISD
		Inconsistent speech					
		Oro-motor skill - SS below 8;					
		Presence of groping;					
		DDK ability - SS below 8					
40	Newmeyer et al. (2009) NAAIGDGW	Clinical characteristics associated with CAS:	N	5	38	3-10 years M=58m	TD
		Inconsistent sound production;					
		oral motor difficulties;					
		Inability to imitate sounds;					
		Groping articulation patterns;					

Increased difficulty with longer utterances;
Poor sequencing.

41	Ballard, Robin, McCabe and McDonald (2010) BRMM	Diagnosis based on presence of core perceptual features (ASHA, 2007) – observed during following assessments:	SSD No scores reported on assessments	Y	X	3 (siblings)	10.10 (m) 9.2 (f) 7.8 (m)
		Articulation	GFTA-2				
		Production of mono and multi-syllabic words, DDK task	Motor Speech Examination (Duffy, 2005)				
		Non-word repetition	Nonword Repetition Task (Gathercole and Baddeley, 1996)				
		Inconsistent speech	DEAP – Inconsistent subtest (Dodd et al., 2002)	No score reported			

		Normal language skills	CELF-4 (Semel, Wiig & Secord, 2006)					
42	Iuzzini and Forrest (2010) IF	Children persistent in exhibiting severely disordered speech. CAS classified based on: low PCC highly variable sound substitutes.	Single subject design – scores reported on individual basis	N	X	4	3.7-6.10	TD PD
		Testing included						
		Articulation	GFTA (SS 64 for 1 and <40 for other 3 subs)					
		Receptive vocab within 1/1.5 SD from mean	PPVT-3					
		Nonword repetition task						
		Speech perception task						
		Language	CELF-P (Wiig, Secord and Semel, 1992)					

		CSIP >25%	CSIP (consonant substitute inconsistency percentage)					
43	Grigos and Kolenda (2010) GK	Diagnosed based on presence of 8 features, 5 segmental and 3 suprasegmental (Shriberg, 2003) Groping Metathetic errors Inconsistent productions Sound and syllable deletions Vowel errors Inconsistent stress placement Reduced temporal variation Inconsistent oral-nasal gestures	Assessments included: GFTA TELD VMPAC Sequencing VMPAC oromotor control	N	8	1	3.2	TD (3)
44	Terband, Maassen, van Lieshout and	Clinical judgment based on:		N	6	5	6.2-8.9	TD (6) SSD/PD (5)

Nijland
(2011)
TMVN

Unintelligible speech for parents
and others;

Inconsistency in articulation
errors;

Slow progress in therapy;

Articulation errors comprising
simplifications;

Inability to produce /pataka/ DDK task
Groping.

45 Shriberg,
Potter and
Strand (2011)
SPS

Diagnosis based on:

N

4/10

25

3-6

TD

Vowel distortions,
Voicing errors,

Distorted substitutions;
 Difficulty achieving initial
 articulatory movement gestures;
 Groping;
 Intrusive schwa;
 Increased difficulty with
 increased complexity;
 Syllable segregations;
 Slow rate,
 Slow DDK;
 Equal stress or lexical stress
 errors.

46	Edeal and Gildersleeve- Neumann (2011) EG	Diagnostic features: Limited phonetic inventory; Inconsistent errors; Difficulty sequencing sounds.	SSD - Subject profiles reported on individual basis.	N	X	2	6.2 and 3.4	-
----	---	--	---	---	---	---	-------------	---

47	Martikainen and Korpilahti (2011) MK	Single subject treatment study. Diagnosis based on subject presenting with 10/11 inclusionary criteria for CAS. Describes speech but does not list features. Vowel inventory complete, Missing consonants from repertoire, Omissions and vowel errors (especially substitutions and distortions) Glottal stops used frequently, Inconsistency of articulation, Poor intelligibility, Overuse of simple syllable shapes, Nasalization of vowels,	Assessments included: Reynell Developmental Language Scales III (receptive language) Wechsler (IQ) Finnish word finding test	N	10	1	4.7	n/a
----	--	--	--	---	----	---	-----	-----

Difficulty producing rapid speech
movements (e.g. pataka),
Receptive language in normal
limits,
Normal IQ,
Naming ability poor.

48	Ruscello (2012) R	Participants pre-diagnosed as CAS. Features of speech for both participants included: Reduced sound inventory, Inconsistent errors Vowel errors Difficulty sequencing Glottal replacement Sound omissions (FCD/ICD) Syllable deletion Prosody differences Voicing errors.	SSD	N	X	2	6.2 3.4	N/A
----	-------------------------	--	-----	---	---	---	------------	-----

49	Shriberg, Lohmeier, Strand and Jakielski (2012) SLSJ	Diagnosis based on CAS features (conversational speech sample): Vowel distortions, Voicing errors, Distorted substitutions; Difficulty achieving initial articulatory movement gestures; Groping; Intrusive schwa; Increased difficulty with increased complexity; Syllable segregations; Slow rate,	PEPPER (programs to examine phonetic and phonological evaluation records; Shriberg, Allen, McSweeny & Wilson, 2001)	N	4/10	18	2 age groups: 3-6yrs and 7+	TD, Speech Delay (SD) with and without Lang Impair (LI)
----	---	--	---	---	------	----	--------------------------------------	---

Slow DDK;
Equal stress or lexical stress
errors.

50	Murray, McCabe and Ballard (2012) MMB	Eligibility includes; Clinical diagnosis of sCAS based on ASHA (2007) criteria and Strand's 10 point check list (see Shriberg et al., 2012) . Eligibility assessment included: Questionnaire Hearing screen; CELF; DEAP Inconsistency Subtest; OME.	No OD inclusion/ exclusion or standard scores reported. Additional assessments used included: CTOPP (phonological processing) NWRep PEPS-C (prosody) GFTA (artic) PPVT (verbal cognitive ability) Scores note reported.	Y (DEAP score not reported)	4+	30	4-12	n/a
----	---	---	---	--------------------------------------	----	----	------	-----

51	Maas and Farinella (2012) MF	Participants pre-diagnosed as CAS based on ASHA features (sub 2 and 3 had dysarthria): Inconsistent vowels and consonants Difficulty transitioning between sounds and syllables Prosodic disturbances		N	3	4	7.9 5.0 (+dys) 6.11 (+dys) 5.3	n/a
52	Maas, Butalla and Farinella (2012) MBF	Diagnosis based on 3 features proposed by ASHA 2007: (Scores reported case by case)	SSD Features determined on spontaneous speech sample by experienced SLP. GFTA-2 Dynamic Evaluation of Motor Speech Skills (DEMSS; Strand, McCauley & Stoeckel, 2006) Inconsistent errors	Y	3	4	5.4-8.4	-

Difficulty transitioning between sounds (presence of inter/intra syllabic pauses) Prosodic errors.	Judged by inter/intra syllabic pauses	
PCC and PVC Receptive language	Repetition of lists of words PPVT-4 CELF-4 Concepts and Directions and Word Structure PLS-4	
Expressive language	CELF-4 core language subtests	
≥ 1.5 SD below Mean ($< 12^{\text{th}}$ Percentile)	GFTA-2 reported individually not as operation definition	Scores reported per sub
Below 85% on sequencing /pataka/	Sequencing Subtest of VMPAC (Hayden and Square, 1999)	

		Word inconsistency	DEAP not reported	Score not reported				
		PCC	Scores reported per subject					
		And PCC-Late 8						
53	Laffin et al, (2012) LRJSJS	Participant eligibility based on Madison Speech Assessment Protocol;		N	3/4	24	8.7 (M)	-
		Transcoding deficits– i.e. planning / programming deficits evidenced by: Vowel distortions, Voicing errors, Distorted substitutions; Difficulty achieving initial articulatory movement gestures; Groping; Intrusive schwa;						

Increased difficulty with
increased complexity;
Acoustic/ perceptual deficits
evidenced by:
Syllable segregations;
Slow rate,
Slow DDK;
Equal stress or lexical stress
errors.

54	Froud and Khamis- Dakwar (2012)	Participants recruited from Childhood Apraxia of Speech Association North America. Participants screened for idiopathic CAS using parent report, SLP report and Apraxia profile. Apraxia profiles lists 10 characteristics of CAS:	Apraxia Profile (Hickman, 1997) used to assess features. SSD Scores on apraxia Profile reported per subject.	N	Features reported per subject	5	5.1-8.3	TD
----	--	---	--	---	--	---	---------	----

		Oral movement (verbal and non-verbal) WNL			
		Prosody % accurate (ranged from 25%-85%)			
		DDK rate below norm by 1 SD			
		Inconsistent speech			
55	Preston, Brick & Landi (2013) PBL	Long-standing diagnosis of CAS. Verified by following assessments: Articulation Sequencing Metathetic errors (switching sounds in words) Migration errors (sounds moving to other positions in words)	Y	6	9.10-15.10
		Below 1.5SD below mean on GFTA VMPAC Sequencing Subtest <85% Sentence imitation and picture-naming			

		Sequencing errors (omissions or additions)						
		PCC	Sentence imitation and picture-naming					
		PCC late 8						
		Inconsistency of errors	DEAP	Score not reported				
56	Dale and Hayden (2013) DH	Diagnosis by SLP based on following;	SSD	Y	6	4	3.6-4.8	n/a
		Minimum criterion Global motor subtest of VMPAC relevant for age;	85% for both 3 and 4 yo.					
		Sequencing Subtest of VMPAC minimum criteria for age	<43% for 3 yo <56% for 4 yo					
		>= 1.5 SD on articulation	DEAP Subtest >=1.5 SD					

Consistency of speech production below 50% on DEAP inconsistency subtest;	DEAP inconsistency subtest < 50%
Receptive skills <1.5 SD below mean;	Auditory Comprehension Scale of Preschool Language – 4 th Ed (PLS-4)
Hearing and orofacial structures normal.	

Note. Age is in years and months; TD = typically developing; LOPT = Limb and Oral Praxis Test; SSD = speech sound disordered without language disorder; SL = speech sound disordered with language disorder; S = speech sound disordered; SL = combined speech and language disorder; DYS = spastic dysarthria; sDYS = suspected dysarthria; SD = non-specific speech disorder; ISD = Inconsistent Speech Disorder; SS = Standard Score; DEAP = Diagnostic Evaluation of Articulation and Phonology; Deviant Consistent = DevCon; Deviant Inconsistent = DevIncon

Table B3

Most Prevalent Diagnostic Features of CAS used from 1993 to 2013

Feature	Authors (initials)	No. of Refs
1 Sequencing deficit	MSM; TMGSa; TMWGS; BD; GMCT; SAK; TMGSb; MNV; OS; NMVGKSa; MGC; NMV; B; NM; MSSJ; TMVN; NMVGKSb; SCKBMN; LFHIT; HSG; MG; EG; R; HHSL; SLSJ; MMB; TMGSD; DJMa; SSM; SCSB; SGCMS; MBW; BS; PSa; NAAIGDGW; BRMM; MK; VS; MF; SPS; PSb; LRJSJS; MGDa; MGDb; PBL; MBF; DH	47
2 Inconsistent speech	OS; GMCT; SAK; TMGSD; NMVGKSa; MGC; NMVGKSb; TMWGS; TMGSa; SSM; SCSB; SD; MNV; NMV; MBW; HSG; LFHIT; NM; PSa; GK; MF; MK; EG; DJMa; PSb; NAAIGDGW; SGCMS; R; TMGSb; BS; HHSL; TMVN; FK; VS; MG; BRMM; MMB; MBF; PBL; MSSJ; IF; DH; BD; MGDa; MGDb	45
3 Limited phonetic inventory	TMGSa; DJMa; SSM; SCSB; BS; PSa; HHSL; LM; BRMM; IF; MK; EG; R; TMGSD; TMGSb; PSb; NMVGKSa; NMVGKSb; NMV; MSSJ; NM; TMWGS; MNV; MGC; SCKBMN; MG; MGDa; DH; PBL; MBF; SMD; LFHTIS; MBW; MSM	34

4	Prosody/ Stress Errors	SSM; MJD; LFHIT; DJMb; JMD; MGDa; MF; R; MBF; SAK; OS; GK; DJMa; MSSJ; SCSB; MBW; BS; PSa; HHSL; SGCMS; LRJSJS; HSG; SPS; SLSJ; MMB; PSb; VS; FK	28
5	Vowel errors	MSM; BD; DJMa; SAK; SSM; VS; SCSB; OS; MBW; MJD; LFHTIS; BS; PSa; DJMb; JMD; MG; HHSL; MGDa; GK; MK; PSb; R; SD; SPS; SLSJ; MMB; LRJSJS	27
6	Omissions/ simplifications	MSM; BD; PSa; R; DJMa; SSM; SCSB; MBW; BS; DJMb; HHSL; MK; MJD; LFHIT; JMD; NM; GK; PBL; TMVN; NMVGKSa; NMVGKSb; NMV; PSb	23
7	Groping	TMWGS; SAK; TMGSb; OS; MNV; SGCMS; HSG; MGC; LFHIT; B; HHSL; LM; MGDb; TMVN; GK; SPS; SLSJ; MMB; BD; NAAIGDGW; VS; LRJSJS; MGDa	23
8	Reduced DDK rate	BD; MG; SPS; SLSJ; MBW; LFHIT; MMB; LRJSJS; SSM; PSb; PSa; BS; LFHTIS; FK	14
9	Deviant errors	TMGSa; SAK; SPS; VS; MMB; TMGSD; SD; OS; LM; MK; R; SLSJ; LRJSJS	13
10	Difficulty imitating sounds/ words	DJMa; SAK; SSM; SCSB; MBW; BS; PSa; LM; NAAIGDGW; OS; VS; PSb	12
11	Unintelligible speech	TMWGS; TMGSb; MNV; GMCT; SD; MGC; SCKBMN; TMVN; MG; MK	10

12	Gap between receptive and expressive language	SAK; DJMa; VS; SSM; SCSB; OS; MSSJ; PSa; PSb; BS	10
13	Context related errors	TMWGS; TMGSb; MNV; GMCT; LFHIT; GK; PBL	7
14	Voicing errors	SPS; SLSJ; R; MMB; LRJSJS	5
15	Slow response to treatment	B; SAK; VS; OS; TMVN	5
16	Slow speaking rate	SPS; SLSJ; MMB; LRJSJS	4
17	Inappropriate loudness	SAK; OS; VS	3
18	Poor phonotactics	BD	1

Table B4

Merged Diagnostic Terminology of CAS from Systematic Review

Final feature	Original Label	Author(s)
Limited phonetic inventory	Limited consonant/ vowel repertoire	SSM; HHSL
	Poor production of consonants	TMGSA
	Many phonemic errors	NMVGKSa; MGC
	Low PCC	SCKBMN
	Phonemic repertoire reduced	MSSJ
	Severely restricted phonemic repertoire	LFHTIS; LM
Groping	Posturing of articulators	MGC
	Difficulty with volitional movements	MBW; PSb
	Difficulty achieving initial articulatory movement	SPS; SLSJ; LRJSJS
Context related errors	Metathetic errors	GMCT; LFHIT; PBL
	Migration errors	PBL
Vowel errors	Nasalization of vowels	MK
	Vowel distortions	
Sequencing problems	Difficulty or inability to produce complex sequences	TMGSD; TMWGS; GMCT; MSSJ; MBW; NMVGKSa; MGC; NMV; SLSJ; LRJSJS
	Difficulty sequencing sounds and syllables	LFHIT

Deviant articulation errors	Predominant/ over use of simple syllable shapes	SSM; SCSB; MBW; PSb; HHSL; MK
	Increased errors on longer units	SSM; SCSB; MBW; LFHIT; NAAIGDGW
	Sequencing errors	PBL
	Difficulty transitioning between sounds and syllables	MF;MBF
	Distorted substitutions	LRJSJS
	Glottal replacement	SD; MK
	Intrusive schwa	SPS; SLSJ; LRJSJS
Prosody-Stress Errors (incl. intonation and timing)	Unusual and persisting errors	SAK
	Syllable segregation	SPS; LRJSJS
	Segmental variability	MJD; DJMb; JMD
	Suprasegmental errors	DJMa;
	Altered suprasegmental characteristics	SSM; SCSB; PSb; HHSL
	Prosodic deficits/ disturbances/ abnormalities/ differences	MSSJ; LFHIT; MJD; DJMb; JMD; R; MF
	Misplaced stress in multi-syllabic words	MBW
	Intonation and stress inconsistencies	GK
	Reduced temporal variation	GK
	Inappropriate timing	SGCMS
	Phrasal errors	BD; MBF
	Equal stress or lexical stress errors	SGCMS; SPS; SLSJ; LRJSJS

Table B5

Features of CAS including Features with Operational Definitions

<i>Feature</i>	<i>Feature</i>	<i>Operational Definition</i>	<i>Papers</i>
Inconsistent Speech	Inconsistent speech	Score $\geq 40\%$ on DEAP IS	B&D (1996); MGD (2009a); MGD (2009b)
	Inconsistent speech	Score $\geq 50\%$ on DEAP IS	DH (2013)
	Inconsistent substitutions	CSIP $> 25\%$	IF (2010)
	Inconsistent errors	Not reported	OS (2001); GMCT (1996); SAK (1997); TMGSD (1997); OS (2001); MSSJ (2002); NMVGKS (2002); MGC (2003); NMVGKS (2003); TMWGS (1996); TMGS (1999); SSM (1999); SCSB (2000); MNV (2001); NMV (2003); MBW (2003); SGCMS (2003); HSG (2004); LFHIT (2004); NM (2005); PS (2008); BRMM (2010); GK (2010); MF (2012); MK (2011); EG (2011); NAAIGDGW (2009); R (2012); MBF (2012); TMGS (1994); DJM (1998); BS (2005); PS (2005); TMVN (2010); FK (2012)
	Periods of highly unintelligible speech	Not reported	TMWGS (1996); GMCT (1996); TMGS (1999); MNV (2001);
	Intelligibility rating poor	Not reported	SCKBMN (2003); HSG (2004);

	Variability and lack of consistent patterns of speech output	Not reported	HHSL (2008);
	Poor intelligibility	Not reported	SD (2000); MGC (2003); TMVN (2010); MK (2011);
Articulation	Severely impaired articulation ability	2 SD below or more below mean on TDSTA	MSM (1993)
	Articulation on GFTA	5 th percentile or less on GFTA	SMD (2000); LFHTIS (2004)
	Articulation on GFTA	≥ 1.5 SD below mean on GFTA	PBL (2013)
	Articulation on subtest of DEAP	≥ 1.5 SD below mean on DEAPAS	DH (2013)
	Articulation on BBTOP	≥ 1.5 SD below mean on BBTOPAS	MGD (2009a); MGD (2009b)
	Limited consonant and vowel repertoire/ inventory	Not reported	DJM (1998); SSM (1999); SCSB (2000); MBW (2003); BS (2005); PS (2005); HHSL (2008); LM (2007); PS (2008); BRMM (2010); EG (2011);

Reduced sound inventory	Not reported	IF (2010); R (2012)
Consonant and Syllable Omissions/ deletions	Not reported	B&D (1996); MJD (2004); LFHIT (2004); MG (2006); JMD (2006); GK (2010);
Articulation errors comprising simplifications	Not reported	TMVN (2010);
PCC reported on individual basis	Not reported	VS (1999);
PCC reported per group (M, SD and range)	Not reported	SCKBMN (2003);
Many phonemic errors	Not reported	SAK (1997); OS (2001); NMVGKS (2002); NMVGKS (2003); NMV = (2003); NM (2005);
High rate of speech sound errors	Not reported	TMWGS (1996); TMGS (1999); MNV (2001); MGC (2003);
Frequent sound omissions (ICD and FCD)	Not reported	SCSB (2000); PS (2008); R (2012); DJM (1998); SSM (1999); SCSB (2000); MBW (2003); BS (2005); DJM (2005); PS (2005); HHSL (2008); MK (2011);

High substitute of consonant substitutions (omissions in clusters)	Not reported	NMVGKS (2002); NMVGKS (2003); NMV = (2003); NM (2005);
Missing consonants	Not reported	MG (2006); MK (2011)
Voicing errors	Not reported	SAK (1997); OS (2001); SPS (2011); SLSJ (2012); MG (2006); R (2012); MMB (2012); LRJSJS (2012)
Glottal replacement	Not reported	SD (2000); MK (2011); R (2012)
Distorted substitutions	Not reported	SPS (2011); SLSJ (2012); MMB (2012);
Intrusive schwa	Not reported	SPS (2011); SLSJ (2012); MMB (2012);
Metathetic/ sequencing errors (switching sounds in words)	Not reported	GMCT (1996); SGCMS (2003); HSG (2004); LFHIT (2004); GK (2010); PBL (2013)
Migration errors (sounds moving to other positions in words)	Not reported	PBL (2013)
Sequencing errors (omissions and/or additions)	Not reported	MG (2006); PBL (2013)

	Reduced PCC and PCC Late-8	Not reported	PBL (2013)
	Poor production of consonants	Not reported	TMGS (1994); TMGSD (1997)
	High incidence of context related errors	Not reported	TMWGS (1996); TMGS (1999); MNV (2001);
	Deviant speech	Not reported	TMGS (1994); TMGSD (1997); LM (2007).
Motor sequencing	DDK rate	TFS ≥ 2 SD below mean on OSMCP	LFHTIS (2004);
	Sequencing difficulties	Age related criterion on VMPACSS	DH (2013)
	Sequencing deficit	VMPACSS $< 85\%$	PBL (2013)
	Reduced sequencing ability	SS below 8 on DDK subtest of DEAP	MGD (2009a); MGD (2009b)
	Sequencing difficulties	Not reported	TMGSD (1997); NMVGKS (2002); NMV = (2003); NMVGKS (2003); LFHIT (2004); NM (2005); MG (2006); SCKBMN (2003); HSG (2004); NAAIGDGW (2009); EG (2011); R (2012)

Difficulty /Inability to produce complex sequences	Not reported	TMWGS (1996); GMCT (1996); TMGS (1999); MNV (2001); NMVGKS (2002); MGC (2003); NMV = (2003); NMVGKS (2003); NM (2005); MG (2006); TMVN (2010); MK (2011);
Poor performance on DDK	Not reported	SAK (1997); OS (2001); MGC (2003); LFHIT (2004);
Difficulty transitioning between sounds and syllables	Not reported	MF (2012); MBF (2012);
Difficulties with motor sequencing as evidenced on LOPT	Not reported	B (2005);
Increased errors on longer units	Not reported	TMGS (1994); DJM (1998); SSM (1999); SCSB (2000); MBW = Munson, Bjorum & Windsor (2003); BS (2005); PS (2008); NAAIGDGW (2009);
Predominant use of simple syllable shapes	Not reported	DJM (1998); SSM (1999); SCSB (2000); MBW (2003); BS (2005); HHSL (2008); PS (2008); MK (2011);
Syllable deletion	Not reported	R (2012)

	Polysyllabic errors	Not reported	B&D (1996);
	Increased difficulty with multisyllabic words	Not reported	LFHIT (2004); SPS (2011); HHSL (2008); SLSJ (2012); MMB (2012);
	Inability to produce complex sequences	Not reported	TMGSD (1997)
	Reduced DDK rate	Not reported	B&D (1996); MG (2006); PS (2008); SPS (2011); SLSJ (2012); FK (2012); MMB (2012); LRJSJS (2012);
	DDK sequencing ability poor	Not reported	B&D (1996);
Prosody	Inappropriate Stress	80% Percentage of Appropriate Stress scores for Prosody-Voice Code 15; Excessive/ Equal/ Misplaced Stress on PVSP.	VS (1999);
	Prosodic abnormalities/ disturbances	Not reported	MJD (2004); LFHIT (2004); DJM (2005); JMD (2006); MG (2006);MGD (2009a); MF (2012); R (2012); MBF (2012); FK (2012);

Altered suprasegmental characteristics (rate, pitch and loudness)	Not reported	DJM (1998); SSM (1999); SCSB (2000); MBW (2003); BS (2005); (2005); PS (2005); HHSL (2008); PS (2008);
Syllable segregation	Not reported	SPS (2011); SLSJ (2012); MMB (2012); LRJSJS (2012);
Segmental variability	Not reported	DJM (2005);
Intonation and stress inconsistencies	Not reported	SAK (1997); OS (2001);
Inappropriate loudness	Not reported	SAK (1997); OS (2001);
Excessive / equal stress	Not reported	SGCMS (2003);
Inappropriate timing (syllable segregation)	Not reported	SGCMS (2003);
Reduced temporal variation	Not reported	GK (2010);
Inconsistent stress placement	Not reported	GK (2010);
Slow rate	Not reported	SPS (2011); SLSJ (2012); MMB (2012); LRJSJS (2012);
Phrasal errors	Not reported	B&D (1996);

	Equal stress or lexical stress errors	Not reported	SPS (2011); SLSJ (2012); MMB (2012); LRJSJS (2012);
Vowel Errors	Vowel Errors	Not reported	SAK (1997); OS (2001); MJD(2004); LFHTIS (2004); DJM (2005); JMD (2006); B&D (1996); MG (2006); GK (2010); MK (2011); R (2012);
	Percent Vowel Correct (PVC) calculated using PROPH	Not reported	MGD (2009a);
	High incidence of vowel errors	Not reported	DJM (1998); SSM (1999); SCSB (2000); MBW = Munson, Bjorum & Windsor (2003); BS (2005); PS (2005); HHSL (2008); PS (2008);
	Nasalization of vowels	Not reported	MK (2011);
	Vowel distortions	Not reported	SD (2000); BRMM (2010); SPS (2011); SLSJ (2012); MMB (2012);
Groping	Oromotor skill	SS below 8 on subtest of DEAP	MGD (2009a)
	Groping	Not reported	TMWGS (1996); SAK (1997); TMGS (1999); OS (2001); MNV (2001); SGCMS (2003); MGC (2003);

Groping / struggle evidenced on LOPT	Not reported	LFHIT (2004); MG (2006); MG (2006); HHSL (2008); NAAIGDGW (2009); LM (2007); MGD (2009a); MGD (2009b); TMVN (2010); B (2005)
Oro motor proficiency as assessed by OSMCP	Not reported	B&D (1996);
Inability to imitate sounds	Not reported	SAK (1997); NAAIGDGW (2009)
Oral motor difficulties	Not reported	NAAIGDGW = (2009)
Impaired volitional oral movements	Not reported	PS (2008);
Difficulty achieving initial articulatory movement gesture	Not reported	SPS (2011); SLSJ (2012); MMB (2012);
Difficulty imitating words/sounds	Not reported	DJM (1998); SSM (1999); SCSB (2000); MBW = Munson, Bjorum & Windsor (2003); BS (2005); PS (2005); LM (2007);

Alternative Classification Method:

Screen for CAS	Probability of correct assignment	95% Probability of correct assignment on STDAS	MSM (1993)
	Probability of correct assignment	99% Probability of correct assignment on STDAS	SMD (2000)
Global Motor	Global Motor skill normal	GMVMPAC age related criterion – 85% for 3 and 4 year olds	DH (2013)
Progress in therapy	Slow response to treatment	Not reported	B (2005);
Phonological Processing	Presence of phonological processes	Presence of 3 phonological processes on KLPA	LFHTIS (2004);
	Poor phonological analysis	Severity rating of 4 (severe) on KLPA	LFHTIS (2004)
	Percent Process Usage calculated using PROPH	Not reported	MGD (2009)
	Poor phonotactics	Not reported	B&D (1996); MG (2006);

Note. BBTOPAS = Bankson-Bernthal Test of Phonology Articulation Subtest (Bankson and Bernthal, 1990); CSIP = Consonant Substitute Inconsistency Percentage (Iuzzini & Forrest, 2010); DDK = diachokinetic; DEAP IS = Diagnostic Evaluation of Articulation and Phonology Inconsistency Subtest (Dodd et al., 2002); FCD = Final Consonant Deletion; GFTA = Goldman Fristoe Test of Articulation; GMVMPAC = Global Motor Subtest of Verbal Motor Production Assessment (Hayden & Square, 1999); ICD = Initial Consonant Deletion; KLPA = Khan Lewis Phonological Analysis; LOPT = Limb and Oral Praxis Test; OSMCP = Oral and Speech Motor Control Protocol (Robbins and Klee, 1987). PCC = Percent Consonants Correct; PROPH = Profile in Phonology Computerised

Profiling Software (Long & Fey, 2005); SS = Standard Score; STDAS = Screening Test for Developmental Apraxia of Speech (Blakeley, 1983); TDSTA = Templin Darley Screening Test of Articulation; TFS = Total Function Score; VMPACSS = Verbal Motor Production Assessment Sequencing Subtest (Hayden & Square, 1999);

Table B6

Coding for Papers Included in the Systematic Review

Study	Author	Year	Code
1.	Marion Sussman & Marquardt	1993	MSM
2.	Thoonen, Maassen, Gabreels & Schreuder	1994	TMGSa
3.	Bradford & Dodd	1996	BD
4.	Groenen, Maassen, Crul & Thoonen	1996	GMCT
5.	Thoonen, Maassen, Wit, Gabreels & Schreuder	1996	TMWGS
6.	Thoonen, Maassen, Gabreels, Schreuder & DeSwart	1997	TMGSD
7.	Shriberg, Aram & Kwiatkowski	1997	SAK
8.	Davis, Jakielski & Marquardt	1998	DJMa
9.	Thoonen, Maassen, Gabreels & Schreuder	1999	TMGSb
10.	Skinder, Strand & Mignerey	1999	SSM
11.	Velleman & Shriberg	1999	VS
12.	Skinder, Connaghan, Strand & Betz	2000	SCSB
13.	Strand & Debertine	2000	SD
14.	Sussman, Marquardt & Doyle	2000	SMD
15.	Odell & Shriberg	2001	OS
16.	Maassen, Nijland & Van der Meulen	2001	MNV

17.	Marquardt, Sussman, Snows & Jacks	2002	MSSJ
18.	Nijland, Maassen, Van der Meulen, Gabreels, Kraaimaat & Schrueder	2002	NMVGKSa
19.	Shriberg, Campbell, Karlsson, Brown, McSweeny & Nadler	2003	SCKBMN
20.	Shriberg, Green, Campbell, McSweeny & Scheer	2003	SGCMS
21.	Maassen, Groenen & Crul	2003	MGC
22.	Munson, Bjorum & Windsor	2003	MBW
23.	Nijland, Maassen, Van der Meulen, Gabreels, Kraaimaat & Schrueder	2003	NMVGKSb
24.	Nijland, Maassen & Van der Meulen	2003	NMV
25.	Hosom, Shriberg & Green	2004	HSG
26.	Lewis, Freebairn, Hansen, Taylor, Iyengar & Shriberg	2004	LFHTIS
27.	Marquardt, Jacks & Davis	2004	MJD
28.	Lewis, Freebairn, Hansen, Iyengar & Taylor	2004	LFHIT
29.	Davis, Jacks & Marquardt	2005	DJMb
30.	Nijland & Maassen	2005	NM
31.	Peter & Stoel-Gammon	2005	PSa
32.	Betz & Stoel-Gammon	2005	BS
33.	Bahr	2005	B
34.	Jacks, Marquardt & Davis	2006	JMD
35.	Moriarty & Gillon	2006	MG

36.	Lundeborg & McAllister	2007	LM
37.	Peter & Stoel-Gammon	2008	PSb
38.	Highman, Hennessey, Sherwood & Leita	2008	HHSL
39.	McNeill, Gillon & Dodd	2009	MGDa
40.	McNeill, Gillon & Dodd	2009	MGDb
41.	Newmeyer, Aylward, Akers, Ishikawa, Grether, deGrauw, Grasha & White	2009	NAAIGDGW
42.	Ballard, Robin, McCabe & McDonald	2010	BRMM
43.	Iuzzini & Forrest	2010	IF
44.	Grigos & Kolenda	2010	GK
45.	Terband, Maassen, van Lieshout & Nijland	2011	TMVN
46.	Shriberg, Potter & Strand	2011	SPS
47.	Edeal & Gildersleeve-Neumann	2011	EG
48.	Martikainen & Korpilahti	2011	MK
49.	Ruscello	2012	R
50.	Shriberg, Lohmeier, Strand & Jakielski	2012	SLSJ
51.	Murray, McCabe & Ballard	2012	MMB
52.	Maas & Farinella	2012	MF
53.	Maas, Butella & Farinella	2012	MBF
54.	Laffin, Racca, Jackson, Strand, Jakielski & Shriberg	2012	LFJSJS

55.	Froud & Khamis-Dakwar	2012	FK
56.	Preston, Brick & Landi	2013	PBL
57.	Dale & Hayden	2013	DH

APPENDIX C

Tables from Chapter 4

Table C1

Prosody Perceptual Rating Scale

Subject ID:							Date:			
Trial	Shark	Boat	Birthday Cake	Elephant	Slippery Slide	Umbrella	Kangaroo	Thank you	Helicopter	Dinosaur
1										
2										
3										
Total										
%										

Note. 0 = normal; 1 = distorted; 2 = severely distorted

APPENDIX D

Tables from Chapter 5: Study 2

Table D1.

Stimuli for Picture-naming Task with Related and Unrelated Auditory Distractor Rimes.

Item No.	Target	Freq	Related Rime	Unrelated Rime
1	frog	1	_og	_ark
2	doll	10	_oll	_og
3	sock	4	_ock	_um
4	torch	2	_orch	_oll
5	shark	0	_ark	_oot
6	drum	11	_um	_oe
7	duck	9	_uck	_og
8	Stamp	8	_amp	_orch
9	Balloon	10	_oon	_le
10	Biscuit	2	_uit	_oom
11	Camel	1	_el	_oot
12	Whistle	4	_le	_oon
13	Apple	9	_le	-et
14	Bucket	7	_et	_en
15	carrot	1	_ot	_oom

Table D2

Stimuli for Delayed Picture-naming Task with Frequency Rating and Practice Items

Test Items	Target	KF Frequency
1	Bag	42
2	Bulb	7
3	Boat	72
4	Bean	5
5	Dog	75
6	Dart	0
7	Fish	35
8	Frog	16
9	Hat	56
10	Hose	9
11	Key	88
12	Cot	-
13	Lamb	7
14	Leg	58
15	Nail	6
16	Nose	60
17	Pear	6
18	Plane	114
19	Snake	44
20	Snail	1

Note: KF Frequency = Kucera & Francis Frequency Rating (Kučera & Francis, 1967)

Table D3

Unstandardized (B) and standardized (β) Regression Coefficients, and Squared Part Correlations (sr^2) for a Hierarchical Multiple Regression Model Predicting the Phonotactic frequency effect from Delayed Picture-naming Reaction Time (DPNrt), Group, and the Group x DPNrt Interaction (N = 53)

Predictors (IVs)	B	95% CI ¹	β	sr^2	p-value ¹
DV: FreqEff					
Step 1					
EVT raw	0.068	-0.241, 0.376	.107	.003	.661
PPVT raw	-0.123	-0.297, 0.0051	-.362	.030	.161
DPNrt	0.006	-0.005, 0.017	.169	.020	.257
$R^2 = .102, p = .041^*$					
Step 2					
EVT raw	0.054	-0.245, 0.352	.085	.002	.718
PPVT raw	-0.136	-0.298, 0.025	-.402	.036	.096
DPNrt	0.012	0.000, 0.023	.321	.063	.041*
Group (D1) ²	-7.502	-13.217, -1.787	-.443	.115	.011*
Group (D2) ²	-0.879	-5.150, 3.392	-.058	.002	.681
Group: $F(2, 47) = 4.03, p = .024^3$					
$\Delta R^2 = .138, p = .015^*$					
$R^2 = .232, p = .005^{**}$					
Step 3					
EVT raw	0.025	-0.272, 0.322	.039	.000	.867
PPVT raw	-0.141	-0.305, -0.023	-.415	.038	.091
DPNrt	0.009	-0.036, 0.018	-.249	.006	.508
Group (D1) ²	-6.830	-13.038, -0.621	-.404	.083	.032*
Group (D2) ²	0.641	-4.084, 5.365	.042	.001	.786
Group: $F(2, 45) = 3.79, p = .030^3$					
Group (D1)xDPNrt ⁴	5.350	-0.330, 11.029	.519	.040	.064
Group (D2)xDPNrt ⁴	3.598	-1.976, 9.173	.244	.016	.200
Group x DPNrt : $F(2, 45) = 1.94, p = .156^5$					
$\Delta R^2 = .041, p = .259$					
$R^2 = .244, p = .008^{**}$					

1: These are the GLMM adjusted values.

2: These are dummy variables; in combination, they reflect the group effect.

3: This is the overall F -value for the group effect.

4: DPNrt has been centered.

5: This is the overall F -value for the Group x DPNrt interaction effect.

* $p < .05$; ** $p < .01$

Table D4

Unstandardized (B) and standardized (β) Regression Coefficients, and Squared Part Correlations (sr^2) for a Hierarchical Multiple Regression Model Predicting Picture-naming Mean Reaction Time from GFTA Raw Scores, Group, and the Group x GFTA Interaction (N = 53)

Predictors (IVs)	B	95% CI ¹	β	sr^2	p-value ¹
DV: PNrt					
Step 1					
EVT raw	-2.366	-11.330, 6.598	-.110	.003	.598
PPVT raw	-1.612	-6.801, 3.577	-.139	.004	.535
Chron Age	-9.677	-15.401, -3.953	-.492	.101	.001**
GFTA raw	-2.384	-5.177, 0.408	-.144	.018	.093
$R^2 = .449, p = .000^{**}$					
Step 2					
EVT raw	-1.066	-9.780, 7.647	-.050	.001	.807
PPVT raw	-0.895	-6.443, 4.653	-.077	.001	.747
Chron Age	-11.889	-21.900, -1.879	-.605	.089	.021*
GFTA raw	-4.572	-11.538, 2.394	-.277	.019	.193
Group (D1) ²	119.419	-175.843, 414.681	.207	.007	.420
Group (D2) ²	91.048	-119.390, 301.486	.175	.009	.388
Group: $F(2, 46) = 0.47, p = .630$					
$\Delta R^2 = .010, p = .665$					
$R^2 = .459, p = .000^{**}$					
Step 3					
EVT raw	-1.856	-10.678, 6.966	-.086	.002	.674
PPVT raw	-0.838	-6.235, 4.588	-.073	.001	.756
Chron Age	-10.579	-20.034, -1.124	-.538	.067	.029*
GFTA raw	-6.175	-19.895, 7.544	-.374	.005	.369
Group (D1) ²	252.343	-13.463, 518.150	.438	.020	.062
Group (D2) ²	107.315	-176.284, 390.915	.207	.005	.450
Group: $F(2, 44) = 1.96, p = .152^3$					
Group (D1)xGFTA ⁴	-61.585	-308.786, 185.617	-.159	.001	.618
Group (D2)xGFTA ⁴	112.583	-177.914, 403.081	.155	.005	.439
Group x GFTA: $F(2, 44) = 1.47, p = .242^5$					
$\Delta R^2 = .025, p = .359$					
$R^2 = .483, p = .000^{**}$					

1: These are the GLMM adjusted values.

2: These are dummy variables; in combination, they reflect the group effect.

3: This is the overall F -value for the group effect.

4: These are the centered GFTA raw scores.

5: This is the overall F -value for the Group x GFTA interaction effect.

Table D5

Unstandardized (B) and standardized (β) Regression Coefficients, and Squared Part Correlations (sr^2) for a Hierarchical Multiple Regression Model Predicting Speech Discrimination Ability (d -prime) from Delayed Picture-naming Reaction Time (DPNrt), Group, and the Group x DPNrt Interaction ($N = 53$)

Predictors (IVs)	B	95% CI ¹	β	sr^2	p-value ¹
DV: D-prime					
Step 1					
EVT raw	0.018	-0.014, 0.050	.230	.013	.259
PPVT raw	0.008	-0.011, 0.027	.181	.007	.417
Chron Age	0.014	-0.004, 0.033	.200	.016	.128
DPNrt	-0.000	-0.002, 0.001	-.060	.002	.664
$R^2 = .360, p = .000^{**}$					
Step 2					
EVT raw	0.016	-0.016, 0.049	.207	.009	.312
PPVT raw	0.006	-0.015, 0.027	.143	.004	.570
Chron Age	0.017	-0.010, 0.044	.233	.013	.217
DPNrt	0.000	-0.001, 0.002	.031	.001	.836
Group (D1) ²	-0.550	-1.251, 0.150	-.260	.033	.121
Group (D2) ²	-0.090	-0.664, 0.482	-.048	.001	.750
Group: $F(2, 46) = 1.33, p = .275$					
$\Delta R^2 = .045, p = .188$					
$R^2 = .405, p = .000^{**}$					
Step 3					
EVT raw	0.019	-0.012, 0.051	.244	.012	.225
PPVT raw	0.007	-0.014, 0.027	.157	.004	.520
Chron Age	0.015	-0.012, 0.042	.207	.010	.277
DPNrt	0.002	-0.000, 0.004	.357	.011	.108
Group (D1) ²	-0.601	-1.352, 0.150	-.284	.036	.114
Group (D2) ²	-0.182	-0.786, 0.422	-.095	.003	.547
Group: $F(2, 44) = 1.30, p < .282^3$					
Group (D1)xDPNrt ⁴	-0.370	0.911, 0.172	-.287	.012	.176
Group (D2)xDPNrt ⁴	-0.297	-0.779, 0.185	-.161	.007	.221
Group x DPNrt : $F(2, 44) = 1.10, p = .341^5$					
$\Delta R^2 = .012, p = .638$					
$R^2 = .417, p = .001^{**}$					

1: These are the GLMM adjusted values.

2: These are dummy variables; in combination, they reflect the group effect.

3: This is the overall F -value for the group effect.

4: These are the centered reaction times for the delayed picture-naming task.

5: This is the overall F -value for the Group x DPNrt interaction effect.

* $p < .05$; ** $p < .01$