

Expressive language development in
minimally verbal autistic children:
exploring the role of speech
production

Joanne Elizabeth Saul

A thesis submitted for the degree of
Doctor of Philosophy

Department of Language and Cognition
Division of Psychology and Language Sciences
UCL

2020

Declaration

I, Joanne Elizabeth Saul, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Signed

Date

Abstract

Trajectories of expressive language development are highly heterogeneous in autism. I examine the hypothesis that co-morbid speech production difficulties may be a contributing factor for some minimally verbal autistic individuals. **Chapters 1 and 2** provide an overview of language variation within autism, and existing intervention approaches for minimally verbal autistic children. These chapters situate this thesis within the existing literature. **Chapter 3** describes a longitudinal study of expressive language in minimally verbal 3-5 year olds (n=27), with four assessment points over 12 months. Contrary to expectations, initial communicative intent, parent responsiveness and response to joint attention did not predict expressive language growth or outcome. Speech skills were significant predictors. **Chapter 4** describes the design, development and feasibility testing of the BabbleBooster app, a novel, parent-mediated speech skills intervention, in which 19 families participated for 16 weeks. Acceptability feedback was positive but adherence was variable. I discuss how this could be improved in future iterations of the app and intervention protocol. **Chapter 5** details how BabbleBooster's efficacy was evaluated. For interventions with complex or rare populations, a randomized case series design is a useful alternative to an under-powered group trial. There was no evidence that BabbleBooster improved speech production scores, likely due to limited dosage. Future research using this study design could determine optimal treatment intensity and duration with an improved version of the app. Taken together, these studies underscore the contribution of speech production abilities to expressive language development in minimally verbal autistic individuals. I argue that this reflects an additional condition, and is not a consequence of core autism features. The intervention piloted here represents a first step towards developing a scalable tool for parents to support speech development in minimally verbal children, and illustrates the utility of randomized single case series for testing treatment effects in small, heterogeneous cohorts.

Impact Statement

My thesis contributes to our understanding of minimal language development in autism, and highlights the need to explore and identify co-morbid speech-motor difficulties for minimally verbal autistic individuals. It tests the feasibility of a speech-focussed, app-based intervention and demonstrates the use of randomized single case series as a robust alternative to large group studies for intervention research with rare and complex populations.

Key beneficiaries of my research are minimally verbal autistic individuals and their families, for whom improved expressive language abilities contribute to quality of life, reduce stress, facilitate self-advocacy, emotional regulation, and access to leisure, academic and social opportunities. This impact can be delivered by publishing lay summaries and mainstream media articles about my research, carrying out public engagement activities and making use of autism charity networks. Raising public awareness about variability in autistic language development is also impactful, as this can increase understanding when alternative forms of communication are needed and reduce stigma for individuals and their families.

The evidence base for interventions targeting expressive language in minimally verbal autistic individuals is currently weak (Brignell, Chenausky, Song, Zhu, Suo & Morgan, 2018). It is important to investigate the application of digital tools to empower parents to play a more active role in supporting their children's language development. New interventions are developed following iterative loops of design, feedback and revision (Craig et al. 2006) and are unlikely to have immediate impact when they are first created. The feasibility trial of BabbleBooster provides a first step proof of concept. Future impact can be delivered by developing it for larger, longer trials, and leveraging commercial partnerships to improve it. A crucial aspect of this impact delivery is the ongoing involvement of autistic individuals and their families as co-designers (Brosnan, Parsons, Good & Yuill, 2016).

Dissemination of findings to clinicians and practitioners who interact regularly with minimal verbal autistic populations can also deliver impact. This may include staff in residential homes, special schools, occupational therapists, speech-language therapist and teachers. This impact can be delivered via social media, professional networks and organisations. In the long term, my findings could contribute to informing good practice guidelines and intervention delivery. Engagement with public policy makers is necessary to deliver this impact.

Beneficiaries of my research within academia include those interested in predicting and enhancing language skills in autism. This research contributes to the ongoing debate on the nature of language impairment in autism and the role of speech-motor difficulties within it.

One element of my research is to explain and demonstrate the benefit of using specific analysis techniques with single case series designs. This methodological contribution could enhance non-academic clinicians' ability to objectively study the interventions they implement, which could result in more robust studies being published in the field, informing future meta-analyses. This impact could extend beyond autism to any field where the target population is rare or heterogeneous, and could be enhanced through future engagement activities such as workshops or webinars.

Table of Contents

Declaration.....	2
Abstract	3
Impact Statement.....	4
Table of Contents	6
List of Figures	10
List of Tables	12
List of Abbreviations.....	13
Preface.....	14
Acknowledgements.....	15
Dedication.....	16
1. General Introduction	17
1.1. Introduction.....	17
1.1.1. Aims of this chapter	17
1.1.2. Language Development in Autism	19
1.1.3. Causal explanations for language impairment in autism	22
1.1.4. Comorbidity.....	23
1.2. Language impairment as secondary effect of autism core features.....	26
1.2.1. Social orienting.....	26
1.2.2. Social motivation.....	31
1.2.3. Joint attention	32
1.2.4. Non-social features of autism	33
1.2.5. Problems for core social account	34
1.3. Language impairment as result of co-morbidity with DLD	36
1.3.1. Behavioural evidence of shared pathways.....	37
1.3.2. Neurobiological evidence of shared pathways	49
1.4. Multifactorial explanation for language impairment in autism and DLD	55
1.4.1. Protective factors	61
1.4.2. Procedural learning	63

1.4.3.	Attention	66
1.4.4.	Perception	66
1.4.5.	Speech	69
1.5.	Conclusion	70
2.	Spoken language interventions in autism.....	74
2.1.	Introduction.....	74
2.1.1.	Rationale for the review.....	74
2.2.	Spoken language in autism interventions.....	77
2.2.1.	Background	77
2.2.2.	Systematic Reviews and Meta-analyses.....	79
2.2.3.	Individual studies	82
2.3.	Other interventions for minimally verbal participants not focussed on spoken language	95
2.4.	Speech interventions for other minimally verbal populations	97
2.5.	Conclusion	99
3.	Study 1: Longitudinal Study.....	103
3.1.	Introduction.....	103
3.2.	Methods	108
3.2.1.	Participants	108
3.2.2.	Variables.....	110
3.2.3.	Procedure	112
3.2.4.	Data Analysis	115
3.2.5.	Planned Confirmatory Analyses	117
3.3.	Results.....	118
3.3.1.	Preliminary Results.....	118
3.3.2.	Putative Predictors	120
3.3.3.	Confirmatory analyses.....	121
3.4.	Discussion	124
3.4.1.	Expressive Language Growth	124
3.4.2.	Confirmatory Analyses	125
3.4.3.	Limitations.....	128
3.4.4.	Conclusion	129
4.	Study 2: Intervention Feasibility.....	131

4.1.	Introduction.....	131
4.1.1.	Aims.....	135
4.2.	Method.....	136
4.2.1.	Intervention design	136
4.2.2.	BabbleBooster Description.....	137
4.2.3.	Consultation Stage 1	140
4.2.4.	Consultation Stage 2	141
4.2.5.	Intervention Pilot Study	142
4.3.	Results.....	149
4.3.1.	Acceptability: parent satisfaction.....	149
4.3.2.	Usability.....	153
4.4.	Discussion	157
4.4.1.	Limitations.....	161
4.4.2.	Conclusion	162
5.	Study 2: Intervention Efficacy	163
5.1.	Introduction.....	163
5.1.1.	Single Case Design	164
5.1.2.	Randomization Test.....	167
5.1.3.	Effect sizes.....	171
5.1.4.	Between-Case Effect Size	172
5.1.5.	Aims.....	173
5.2.	Methods	176
5.2.1.	Study Design.....	176
5.2.2.	Intervention.....	177
5.2.3.	Analysis plan.....	183
5.3.	Results.....	186
5.3.1.	Reliability of parent ratings of speech production attempts	186
5.3.2.	Can weekly scores be used to evaluate efficacy using the Randomization Test?	186
5.3.3.	Efficacy results from BabbleBooster	187
5.4.	Discussion.....	193
5.4.1.	Limitations.....	197
5.4.2.	Conclusion	197
6.	General discussion	198

6.1.	Summary of Findings.....	200
6.2.	Implications	203
6.2.1.	Expressive language trajectories for minimally verbal autistic children	203
6.2.2.	Importance of speech skills	205
6.2.3.	Social variables were not significant predictors	208
6.2.4.	Novel contributions from intervention study.....	211
6.2.5.	Lack of intervention effect	215
6.3.	Limitations.....	217
6.4.	Concluding comments.....	218
7.	References.....	220
8.	Appendices	261
	Appendix A: CSBS Coding information	261
	Appendix B: Phoneme Questionnaire	268
	Appendix C: Alphabet Stimuli	269
	Appendix D: Family Background Questionnaire	271
	Appendix E: Therapy Report Questionnaire	273
	Appendix F: Parent-Child Interaction Coding information	276
	Appendix G: Stage 1 Consultation	281
	Appendix H: Stage 2 Consultation.....	286
	Appendix I: Sound Target Protocol	289
	Appendix J: App User Questionnaire	293
	Appendix K: Score Breakdown of Acceptability Questionnaire	296
	Appendix L: Simulation.....	297
	Appendix M: Educational Setting and Communication System Questionnaire	305
	Appendix N: Code to Calculate Randomisation Test	308

List of Figures

Figure 1. Comorbidity model where causal factors overlap in some individuals .	23
Figure 2. Bidirectional influences on behaviour, cognitive and brain level in autism and language impairment.....	25
Figure 3. Phenomimicry and co-morbidity explanations of language impairment in autism.....	36
Figure 4. Phenomimicry and co-morbidity explanations for social cognition risk factors in DLD	38
Figure 5. Multi-factor interactive model of shared and non-shared factors	56
Figure 6. Opportunities for disruption in language learning.....	57
Figure 7. Key elements of the intervention development and evaluation process (adapted from Craig et al., 2006, p.8).	101
Figure 8. Recruitment Flow Chart Study 1	109
Figure 9. Data Collection for Study 1	113
Figure 10. Expressive language (parent report of number of words child reliably produces) at each of the four time-points (each separated by four months).	119
Figure 11. Individual expressive language trajectories (parent report of number of words child reliably produces) at each of the four time-points (each separated by four months).	119
Figure 12. Learning stimuli (left: customisable; right: mandatory).....	139
Figure 13. Depiction of one intervention trial.....	140
Figure 14. Recruitment flow chart for Study 2 feasibility analysis	143
Figure 15. All possible permutations of baseline (A) and intervention (B) weeks.	145
Figure 16. Summary of data collected from Study 1 and Study 2	146
Figure 17. Steps needed to calculate a Randomisation Test.....	168
Figure 18. Probability distribution of all possible mean differences.....	169
Figure 19. Calculation of unadjusted Between-Case Effect Size	173
Figure 20. Intervention schedule for Study 2.....	176
Figure 21. Recruitment flow chart for Study 2 efficacy analysis	178
Figure 22. Reliability of parent-rated versus clinician-rated weekly scores.....	186

Figure 23. Weekly scores on elicited phoneme test, by participant (as %)	187
Figure 24. Expressive language over 16 months, by participant.....	191
Figure 25. Consonant inventory over 16 months, by participant	191

List of Tables

Table 1: Summary of individual intervention studies aiming to improve spoken expressive language in minimally verbal autistic children	92
Table 2: Variables measured in Study 1	111
Table 3: Descriptive Measures in Study 1	118
Table 4: Expressive Language Correlations	120
Table 5: Correlations	121
Table 6: Model Summary	123
Table 7: BabbleBooster parent rating buttons	139
Table 8: Descriptive Statistics for whole sample	150
Table 9: Usability Data for 'high' user group	154
Table 10: Descriptive statistics describing the demographic features of the High vs. Low user groups	156
Table 11. Comparison of A and B week elicited phoneme scores	189
Table 12. Distribution of correct and incorrect scores for trained and untrained phonemes	189
Table 13. Descriptive variables	190
Table 14. Average change in secondary outcome measures for 12 months before Visit 1 and the 4 months before Visit 2.....	192

List of Abbreviations

AAC:	Alternative and Augmentive Communication
ADHD:	Attention Deficit Hyperactivity Disorder
ALI:	Autism + Language Impairment
ALN:	Autism + Typical Language Development
DLD:	Developmental Language Disorder
PECS:	Picture Exchange Communication System
RCT:	Randomized Control Trial
SES:	Socio-Economic Status
SGD:	Speech Generating Device
TD:	Typically Developing / Typical Development

Preface

The longitudinal study in Chapter 3 of this thesis was published in the academic journal *Autism* (Saul & Norbury, 2020) and the feasibility study in Chapter 4 is under review by *Pilot and Feasibility Studies*.

Acknowledgements

I would like to thank my wonderful supervisor Courtenay, for her endless patience and support. It has been a privilege to be part of her lab over the past four years and I could not have wished for a more astute advisor, enthusiastic cheerleader or generous provider of cake. I'm grateful to James Cusack for putting us in touch, and for all those who encouraged me to pursue this project and held my hand at the various difficult parts. I could not have got to the end without the unconditional love and support of my amazing husband Jeff and my wonderful mum Lesley. One of the trickiest parts of this project was travelling all over the UK. Special thanks to Anne and Dan for enabling me to zigzag the North East come rain, shine or snow to visit participants!

I am hugely fortunate that our family is kept afloat by a veritable army of creative and caring people. I would like to thank 'Team Freddie' members, past and present, for enabling me to follow my passion: Emily, George, Naz, Karishma, Rebecca, Caitlyn, Vicki, Katie, Vicky, Evelyn, Marlana, Kasia, Karolina and Albena. Special mention goes to Dawn, who has been there from the beginning, always goes above and beyond, and has superhuman levels of tact and insight.

I would also like to thank my second supervisor John Swettenham for his thoughtful input, and the many others who helped and advised me along the way: Mark Huckvale, Aida Suarez-Gonzalez, Simon Wallace, Elizabeth Hill, Sarah Griffiths, Gwen Brekelmans, Cat Silvey, the team at Therapy Box, Ashley Peacock, David Stepanovich, the ASD-UK team, fellow PhD passengers in Room 202 and lovely LiLaC lab-mates. Huge thanks to the amazing second-coders: Chatrin, Ria, Chanelle and Anna. I'm also very thankful that I accidentally found my way into Liz Wonnacott's lab during my Masters, and am grateful for all the support I received from both her and Anna Samara.

Finally, the biggest thanks of all go to the children who took part in my study and their families, I thank you for your patience and generosity from the bottom of my heart.

Dedication

For Max and Freddie.

1. General Introduction

Explanatory models for individual differences in structural language acquisition in autism

1.1. Introduction

1.1.1. Aims of this chapter

The core characteristics associated with autism centre on differences in social engagement and behavioural rigidity (APA, 2013). Among the behavioural features observed in young autistic¹ children are a failure to orient to social stimuli (voices, faces), reduced joint attention to actions or objects, and a tendency to engage alone in repetitive actions rather than interactive or imaginative play. Given that language acquisition is assumed to be an inherently socially facilitated process, this combination of features should heavily derail language learning. Indeed, many autistic children experience temporary or permanent difficulties with structural language (i.e. grammar, morphology, phonology). These range from subtle subclinical differences to a complete lack of functional communication. These language impairments are often thought to be a direct consequence of core autism features. However, this hypothesis does not explain why some individuals nevertheless do acquire language. A key aspect to understanding language variation in autism, and goal of this chapter, is to fully explain the broad range of language outcomes, including autistic individuals who acquire intact structural language. In doing so, I argue that when language impairment is observed, it is necessary to look beyond core autism features to fully understand the profile of

¹ In this thesis, I use identity-first language (e.g. “autistic individual”) rather than person-first language (e.g. “individual with autism”), as this has been highlighted as the preference of the majority of autistic individuals and their families (Kenny et al., 2016).

language deficits, and therefore which intervention programmes may be most suitable.

The aim of this chapter is to examine the heterogeneous range of structural language trajectories in autism, and to evaluate which causal models best account for the body of evidence gathered to date from behavioural and physiological sources. Explanatory models must account for the four broad language profiles of autistic individuals (who in reality present along a continuous spectrum rather than categorically): 1) those who remain minimally verbal; 2) those who acquire language but persistently display features akin to those experienced in Developmental Language Disorder (DLD)²; 3) those with an atypical early language trajectory but whose structural language difficulties resolve by middle childhood and 4) those who have not experienced detectible structural language impairment at any stage, despite autism core features. Explanatory frameworks must therefore address not only variability in end-state reached but also in developmental trajectory and thus language acquisition mechanisms (Paterson, Brown, Gsodl, Johnson, & Karmiloff-Smith, 1999) as well as sub-clinical individual differences.

Two competing theories currently aim to explain language variation: one suggests it merely represents a secondary effect of core autism features (Williams, Botting & Boucher, 2008), and the other posits that some autistic individuals have an additional co-occurring language disorder (co-morbid DLD, Tomblin, 2011). The secondary effect hypothesis has some valid evidence but fails to explain the full range of language outcomes across the autism spectrum. Evidence of shared causal pathways for language impairment on the other hand, is gaining empirical support from behavioural, genetic and neurobiological approaches. Existing behavioural evidence of co-morbidity has been open to conflicting interpretations: autistic children with impaired language (ALI) appear to share fundamental areas

² Previously known as Specific Language Impairment but DLD is the term adopted by the UK research and SLT community, for more on this and other historic terms used see Bishop (2017).

of weakness with children with DLD (e.g. Tager-Flusberg, 2006), however within those weak areas the two groups also appear to have systematic qualitative differences (Whitehouse, Barry & Bishop, 2008).

Experimental data has thus far been skewed to underrepresent both the least able autistic children and the more mildly impaired children with DLD, leading to potential confounds. I conclude that neither explanation is fully satisfactory; a multifactorial model comprising elements of both but also giving consideration to additional risk and protective factors and the nature of co-morbidity, may improve our understanding. A key goal of further research should be to specify this multifactorial model, and thus identify those within-child and environmental aspects that could eventually be therapy targets, in order to reduce the wide-ranging and lifelong effects of language impairment in this population.

1.1.2. Language Development in Autism

Pragmatic communication difficulties are considered a unifying core feature of autism and thus necessary for diagnosis (APA, 2013; Lord & Paul, 1997). Pragmatic skills describe how language is used on-line in social contexts and includes use and interpretation of non-verbal communication, ability to adapt communication to the audience, inferring speaker state of mind and a host of other verbal and non-verbal skills necessary to communicate effectively. In contrast, trajectories of structural language development (phonology, lexicon, syntax) are highly heterogeneous in autism (Boucher, 2012). Development of functional speech by age five is one of the strongest predictors of positive outcome in autism (Howlin, 2005; Szatmari, Bryson, Boyle, Streiner & Duku, 2003), so understanding the causal mechanisms behind such variation is both theoretically and practically important. Language trajectories over childhood are characterised by remarkable stability in typical and atypical populations from school entry (e.g. Norbury et al., 2017; Bornstein & Putnick, 2012; Bornstein, Hahn, Putnick & Suwalsky, 2014; Bornstein, Hahn, Putnick & Pearson, 2018). However, Pickles, Anderson & Lord (2014), highlighted that whilst this is also true for autism, there may be a window

of heightened plasticity in early language development. Latent class analysis of 192 participants with longitudinal data from age two to 17 indicated that each child followed one of seven trajectories. The trajectories varied widely in the early years, with some starting low and rapidly catching up to typical language levels, whilst others started higher but plateaued to reveal a more fundamental delay over time. Nevertheless, each trajectory stabilised by age six and tracked in parallel for the remaining 11 years. This may indicate that there could be a limited time window for language interventions to maximise their impact. Alternatively, these data could be accounted for by increased measurement error in the early stage of language learning. Profiles resulting in better language outcomes had milder average autism symptom severity and non-verbal IQ and vice versa for those whose language skills remained low, however the strength of these correlations was not tested in the paper.

'Minimally verbal' describes those with very limited "useful" speech (i.e. speech used in a frequent, communicative, non-imitative and referential way, Yoder & Stone, 2006). The number of autistic individuals who remain minimally verbal is estimated to be 14-29% (Anderson et al., 2007; Bal, Katz, Bishop & Krasileva, 2016; Norrelgen et al., 2014; Rose, Trembath, Keen & Paynter, 2016). Many but not all minimally verbal individuals also display limited receptive language skills (Gernsbacher, Geye & Ellis Weismar, 2005; Plesa Skwerer, Jordan, Brukilacchio & Tager-Flusberg, 2015). Tager-Flusberg et al. (2009) depicts a 'pre-verbal' phase of language development and multi-dimensional criteria to define the transition to the 'first words' phase. Kasari, Brady, Lord and Tager-Flusberg (2013) suggest that for research purposes the definition of minimally verbal should be a vocabulary size of 20 words or fewer.

Some autistic individuals do develop spoken language but experience difficulties with language acquisition (evidenced by late onset of words and phrases). These difficulties either persist or fully resolve. Others have no history of structural language impairment. Autistic individuals whose structural language is age appropriate are often referred to in group studies as "Autism Language Normal (ALN)". The extent of this group's sub-clinical differences with typically developing

profiles is a matter of ongoing enquiry (Shriberg, Paul, McSweeny, Klin, Cohen & Volkmar, 2001; Boucher, 2012).

In between the extreme of minimally verbal and ALN presentation, many autistic children remain impaired in multiple aspects of structural language such as morphology, syntax, lexicon and tend to be described in group studies as “ALI – Autism Language Impaired”. I believe these subgroupings to be oversimplifications of vastly multidimensional heterogeneity, however due to their ubiquity in past studies, they will be used in this review.

Phonology is an aspect of structural language thought to be largely spared (Naigles & Tek, 2017) and thus a relative strength across all groups, although evidence for typical trajectories is lacking and phonological impairments cannot be excluded as a possible causal factor for those who have no/limited spoken language (Boucher, 2012). I will return to the role of phonology and speech production skills in language acquisition at the end of this chapter as it is a focus of this thesis.

There are several epiphenomenal reasons why the relationship between language impairment and autism is still poorly understood. Diagnostic criteria for autism have widened over time, making older and newer findings hard to reconcile (new studies often comprise participants with IQ and language in the normal range, older studies comprise more participants with language impairments and a broader range of non-verbal cognitive impairment). It is important to examine the validity of grouping variables, assuming the commonly used dichotomy (ALI / ALN), rather than continuous measures of language impairment (Ingram, Takahashi & Miles, 2008). When drawing comparisons with other disorders, there may be a sampling bias, since more profoundly impaired autistic children are often omitted from research (either by exclusion criteria or due to the difficulty of taking part), whereas clinical DLD samples are more representative of severely affected individuals, yet subject to more arbitrary non-verbal IQ criteria. Language impairment in less able autistic individuals has been understudied and is often assumed to be caused by general learning difficulties, a claim which requires further investigation (e.g., McGonigle-Chalmers, Alderson-Day, Fleming & Monsen, 2013; Mirenda, 2008). A general

limitation to understanding the broad diversity in language trajectories and outcomes has been the dearth of research on those who are more severely impaired and fail to acquire any functional spoken communication, however techniques and a body of work is now developing (Tager-Flusberg & Kasari, 2013). Furthermore, new research is investigating language longitudinally using younger diagnosed or high-likelihood siblings (with greater chance of later autism diagnosis due to having an autistic sibling), in order to capture early acquisition processes that may be less evident when examining school age children (Hudry et al., 2014; Leonard, Bedford, Pickles & Hill, 2015; Tager-Flusberg, 2010).

1.1.3. Causal explanations for language impairment in autism

Language impairment occurs so frequently alongside autism that early causal models suggested social communication symptoms were secondary to the central language difficulties (Rutter & Bartak, 1971). This model has since been refuted, not least by the empirical evidence that many autistic individuals have no current or historical structural language disorder. Additionally, autistic children display features not seen in DLD and unrelated to the severity of their language deficit (e.g. echolalia, pronoun reversal). Some postulate that the reverse relationship is true – atypical language development is a secondary effect of the primary autism features (Williams et al., 2008). This begs the question: why do some autistic individuals still manage to acquire typical structural language and others do not?

One somewhat controversial answer to this question is to suggest that DLD and autism share a high degree of co-morbidity. Approximately 7% of the population has significantly deficient language proficiency in the absence of behavioural, cognitive or neurological issues and is thus described as having DLD (Tomblin et al., 1997; Norbury et al., 2016). Hallmarks of this disorder are difficulties with morphosyntax and phonological processing, although other difficulties can also be present.

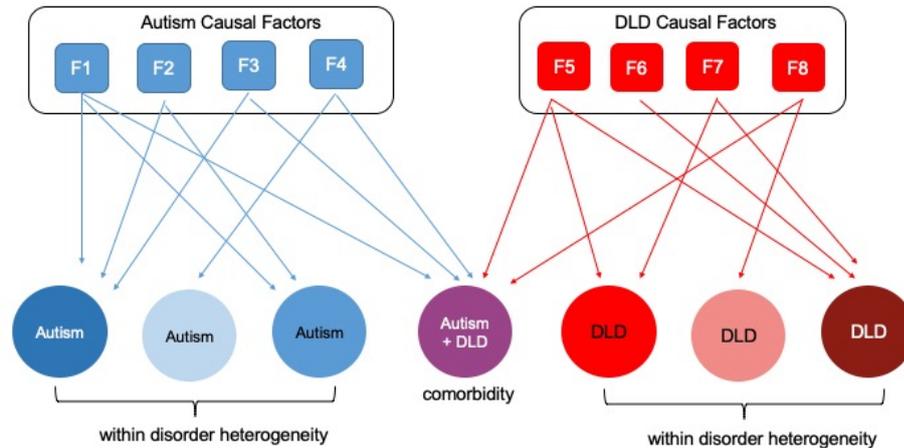


Figure 1. Comorbidity model where causal factors overlap in some individuals

Note: DLD: Developmental Language Disorder

Under this explanation, illustrated in Figure 1, some individuals would only display autistic or DLD symptoms and others would experience ‘multiple hits’ and display features of both autism and DLD. Evidence put forward (and disputed) for this theory has consisted of analysing similarities of those with ALL and those with DLD on behavioural, cognitive and neurobiological levels (e.g. Kjelgaard & Tager-Flusberg, 2001; Williams et al., 2008). Whilst I will briefly re-visit the evidence put forward in that debate, it is necessary to explore how we can more fully account for the full range of language outcomes in autism. If co-morbidity is part of the picture, I will endeavour to further specify it, in order to formulate testable hypotheses.

1.1.4. Comorbidity

Co-morbidity is the norm in developmental disorders, thought to occur because most disorders are multifactorial involving complex interactions between genetic, biological and environmental risk factors (Pennington, 2006; Pennington & Bishop, 2009). Much emerging evidence suggests that risk may be shared across disorders to some degree. It is widely accepted that disorders such as Attention

Deficit Hyperactivity Disorder (ADHD) frequently co-occur with autism and have a combined interactive impact on the individual, or that other groups of disorder tend to cluster in individuals (e.g. dyslexia, DLD and speech sound disorder; Pennington & Bishop, 2009; Ramus, Marshall, Rosen & Van Der Lely, 2013). However, for DLD and autism this appears to be a more problematic proposition, possibly because there is such strong bidirectional linkage between the dimensions of pragmatic communication skills and structural language skills.

Within-syndrome heterogeneity is also a common feature of developmental disorders and has important consequences for causal modelling; even aetiologically well-defined syndromes like Williams syndrome are still very heterogeneous in expression (Porter & Coltheart, 2005). Furthermore, many have posited that autism and DLD themselves are not discrete categories (Tomblin, 2011; Ronald et al., 2006; Brunsdon & Happé, 2014). Some studies characterise DLD as heterogeneous and comprising various subtypes (Conti-Ramsden, Crutchley & Botting, 1997, van Weerdenburg, Verhoeven & van Balkom, 2006). DLD may also not be very specific to language, commonly occurring with co-morbid motor impairments (Bishop, 2002; Hill, 2001). Lenroot and Yeung (2013) reviewed neuroimaging findings on heterogeneity within autism, acknowledging that diagnostic categories may be of limited use when examining neurobiological findings. As an alternative, Levy and Ebstein (2009) advocate cross-disorder studies, which focus on traits rather than potentially artificial diagnostic categories.

Given that there is unlikely to be a single cause for all the features of autism (Ronald et al., 2006, Fletcher-Watson & Happé, 2019) a fuller explanatory model would need to examine the multiple causal factors which combine and interact to result in a given individual's profile, regardless of the diagnostic category that best describes them. Whilst within-child cognitive risk and protective factors are crucial to consider in such a model, so is their social and linguistic environment, and their particular constellation of developmental experiences. I will look at evidence on a brain, cognitive and behavioural level under the assumption that genetic alterations affect neural connectivity and brain morphology which in turn cause atypical information processing, which influences behavioural phenotypes. However,

influence can also flow in the opposite direction: certain behavioural features could impair development of a cognitive skill, which could influence brain development as set out in Figure 2.

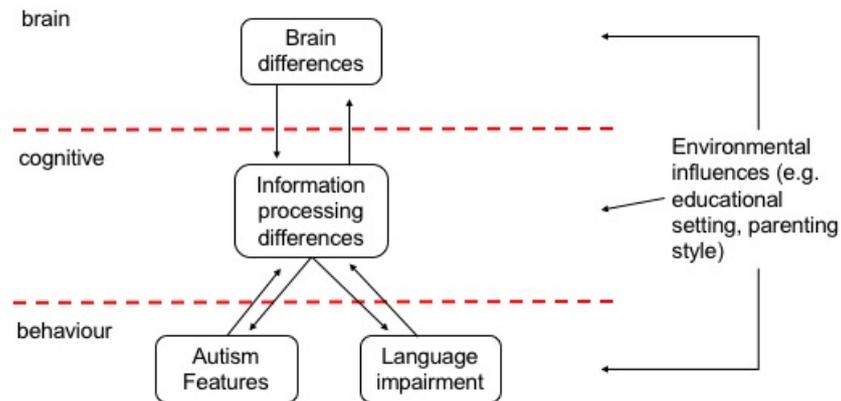


Figure 2. Bidirectional influences on behaviour, cognitive and brain level in autism and language impairment

For example, if a compensatory process is invoked, the neural networks for that process may be strengthened at the expense of those for the typical process. I also acknowledge that causal models are interactive, probabilistic rather than deterministic and subject to environmental influence at each level (Pennington, 2006; Pennington & Bishop, 2009). A possible multifactorial model taking these considerations into account is outlined in 1.4.

1.2. Language impairment as secondary effect of autism core features

Under one explanation, the core features of autism may derail or delay children's structural language development, i.e. language impairment is a secondary effect of social communication deficit and/or behavioural rigidity. Social cognition deficits in autism have been variously characterised as stemming from a weakness in social orienting (Mundy & Neal, 2001), social motivation (Dawson et al., 2004) or the development of joint attention (Sigman & Ruskin, 1999); a full exploration of these theories is beyond the scope of this Chapter, however I outline the evidence below that purports to link weak social cognition with impaired language development. The evidence linking non-social core autism features (repetitive and restricted behaviours, sensory sensitivities) to language impairment is also discussed.

1.2.1. Social orienting

Under this account, autistic children do not prioritise social information from an early age, and this contributes to atypical social development, with sequelae for language development.

A lack of exposure to social stimuli has been associated with "quasi-autistic symptoms" and impaired language development in early social deprivation (Rutter et al., 1999). Autistic infants may experience self-initiated social deprivation, which could result in impaired language development.

1.2.1.1. Social orienting to voices

In a recent review of possible mechanisms behind differences in lexical acquisition in autism, lack of attention to speech is highlighted (Arunachalam & Luyster, 2016). Klin (1991, 1992) reported that in a play-based experiment, autistic children aged 3-6 lacked the preference for child directed speech that was demonstrated by their

Typically Developing (TD) and developmentally delayed (non-autistic) peers. Dawson, Meltzoff, Osterling, Rinaldi and Brown (1998) reported reduced orientation to social stimuli (name being called, clapping) in autistic participants compared to TD peers and those with Down syndrome. These findings suggest a failure to orient to voices is worthy of further investigation, however the original sample sizes were small and heterogeneous. The non-speech stimuli in question was a background noise of multiple voices, whereas subsequent experiments have used rotated speech, since it controls for the acoustic properties of speech. Overall preference for non-speech was also found in 29 autistic 2-3 year olds (Kuhl Coffey-Corina, Padden & Dawson, 2005), whereas TD mental-age matched controls displayed no preference as a group. Results in the autistic group were heterogeneous, and those who preferred non-speech also demonstrated weaker speech perception in an event-related potential analysis. This is suggestive of a link between attention to speech and language skills, however without language-matched controls or follow-up measures of language outcomes, no firm conclusions can be drawn, and one cannot imply causality in either direction.

In an effort to explore links with longer-term language outcomes, Paul, Chawarska, Fowler, Cicchetti & Volkmar (2007) examined whether preference for infant-directed speech differed between autistic 2-3 year olds and several control groups. When compared with the autistic group, typical age-matched peers demonstrated a significantly greater preference for speech, however no significant differences were observed between younger language-matched controls or age-matched controls with other developmental delays, suggesting that preference for speech may be linked to developmental stage. Furthermore, in the autistic group, time spent attending to speech (but not overall attention) was positively related to receptive language, concurrently and one year later. Watson, Baranek, Roberts, David and Twyla (2011) examined behavioural and physiological reactions to speech vs. non-speech in 22 autistic preschoolers. Both measures were related to concurrent and later language skills. Only increased vagal activity during speech was a significant predictor of later communication skills after accounting for initial communication skills.

Little is known about speech preferences in younger autistic infants, since the earliest diagnoses are received towards the end of the second year. At 12 months, infants with a higher chance of developing autism (due to being sibling of an autistic child) do not have a clear preference for speech (vs. siblings of typically developing children), and this correlates with autism symptom severity at 18 months (Curtin & Vouloumanos, 2013). Across the sample, preference for speech also correlated with expressive language at 18 months. However, higher intragroup variability was observed in speech preference for the autism siblings group, likely reflecting heterogeneity of the sample: only 20% of siblings are expected to receive a subsequent autism diagnosis. These results would be strengthened by follow-up analyses including eventual diagnostic categories and language outcomes.

1.2.1.2. *Social orienting to faces*

As with social attention in the auditory modality, eye-tracking investigations of visual attention to faces have revealed group differences between TD and autistic groups across the lifespan and using a variety of stimuli (see Frazier et al. 2017 for a meta-analysis). However, evidence for the relationship between atypical looking patterns and extent of language impairment is still unclear.

Autistic participants allocate less attention to faces and more to bodies, objects and other background areas during online viewing paradigms or naturalistic social interactions (Ozonoff et al., 2010; Rice, Moriuchi, Klin & Jones, 2012; Klin, Jones, Schultz, Volkmar & Cohen, 2002, however cf. Guillon, Hadjikhani, Baduel & Rogé, 2014, for a critical review). Work is ongoing to determine the precise developmental timing of when attention to eyes begins to diminish, as subtle group differences may emerge in the first 6 months of life (Jones & Klin, 2013; however, cf. Young, Merin, Rogers & Ozonoff, 2009).

Hosozawa and Tanaka (2012) undertook a cross-disorder comparison of visual attention to social scenes of children aged 2-3 years with either autism, DLD or TD (n=66). TD and DLD groups both fixated faces at appropriate times, whereas the autistic group's visual attention was more likely to be elsewhere in the scene.

Compared to TD children, the DLD group allocated a greater proportion of face attention to mouths rather than eyes, whereas autistic children did not differ from TD (albeit with much less overall face attention). Multidimensional scaling revealed that TD and DLD overall gaze patterns were similar, but patterns of autistic individuals differed both from the comparison groups and each other. The autistic children in this study were matched with TD controls on developmental age, with no specific reporting of language measures.

Some studies have shown reduced attention to eyes and greater attention to mouths in autism, however controversies remain regarding the specificity of this finding to autism and its relationship with social and communicative competence. Jones, Carr and Klin (2008) observed significantly reduced attention to eyes and increased attention to mouths in 15 autistic two year olds, compared to typical controls and developmentally delayed controls (mental-age and language matched). Attention to eyes was correlated with social competence for autistic participants. Across all groups, mouth-looking was not correlated with verbal abilities. Opposing findings may be due to different patterns of looking at different stages of development (Nakano et al., 2010).

Young et al. (2009) examined whether gaze behaviour at six months predicted diagnostic outcomes at two years in a siblings group (n=55). Early eye contact was not related to diagnostic outcome, however greater fixation to mothers' mouths (vs. eyes) during a Still Face paradigm was positively related to expressive language growth. Only 3 participants in this sample went on to receive an autism diagnosis, so these conclusions lack specificity to autism.

Hanley et al. (2014) explored social gaze in autism, DLD and typical development (TD) in a naturalistic eye tracking study. The authors examined 'mouth bias' in both groups to evaluate whether it might be a 'social' or 'communicative' compensation strategy. Autistic participants were aged 7-13 and were matched on age and non-verbal ability with the TD group and on age and verbal ability with the DLD group. The autistic group displayed reduced attention to faces and eyes, whereas the DLD and TD groups did not differ significantly. No group differences were observed

regarding allocation of attention to mouths, however there was a large amount of individual variance in both autistic and DLD groups.

Norbury et al. (2009) investigated looking behaviour in a heterogeneous group of autistic adolescents, comprising an ALN and an ALI subgroup (matched on non-verbal ability and autism symptom severity), as well as TD controls. Video stimuli were designed to be maximally relevant to the participants' daily lives, and looking measures were related to social and communicative competence. The ALN group were slower to view, and spent less time fixating eye regions, whereas the ALI and TD groups did not differ significantly. Furthermore, looking patterns did not relate to social competence, whilst communicative competence was correlated with more time attending mouths and less time attending to eyes.

Autistic individuals may also display differences in timing of looking patterns and ability to transfer gaze in a dynamically evolving social interaction (Fletcher-Watson, Leekham, Benson, Frank & Findlay, 2009). Landry and Bryson (2004) compared autistic children's ability to disengage visual attention in a simple gaze-shifting paradigm, compared to TD controls and a group with Down Syndrome. They found evidence of "sticky attention" in the autistic participants, a domain-general difficulty which is hypothesised to contribute to difficulties processing complex social stimuli. Replicating these findings with DLD, ALI and ALN groups, comparing their attention shifting skills to gaze patterns and language abilities could deepen our understanding of the role of looking patterns in language development.

1.2.1.3. Specificity to autism?

Although this evidence suggests that impaired language may be associated with a reluctance to engage with social stimuli, it does not imply causality, and does not explain how some autistic individuals acquire and process language despite orienting less to social stimuli. The inconclusive findings regarding 'mouth bias' and whether it extends to those with DLD also does not imply this must be due to co-morbidity - perhaps DLD and ALI groups fail to orient to or benefit from social input

for different reasons. Studies investigating whether inability to process social stimuli reflects a top down or bottom up difficulty are inconclusive (Gervais et al., 2004; Whitehouse & Bishop, 2008a). Sensory overload, as frequently reported by autistic individuals (Grandin, 1996; Donna Williams, 1999; Bogdashina, 2011) may result in avoidance of social stimuli in some cases. Alternatively, the observed patterns may reflect an indifference, rather than aversion, to social stimuli (Moriuchi, Klin & Jones, 2017).

1.2.2. Social motivation

This account is similar to the social orienting account, however it postulates that atypical social development occurs because autistic children do not find social interaction inherently rewarding, and this impacts language development.

Multiple sources of evidence suggest that social motivation is a key ingredient to early speech perception and production development (Kuhl, 2007; Gros-Louis, 2014). TD infants experience perceptual narrowing during the first year, whereby they lose the ability to detect non-native phonemic contrasts by 12 months, signalling a specialisation in phonemes of their own language. Infants exposed to 12 sessions of Chinese via a live interaction demonstrated phonemic discrimination akin to infants raised in a monolingual Chinese household, whereas infants receiving the same input via video did not demonstrate any phonemic learning (Kuhl, Tsao & Liu, 2003). This suggests that learning to discriminate native speech sounds may be socially gated (Kuhl, 2007). Elsabbagh et al. (2013) demonstrated that in a 'high-contingency' group of TD infants (i.e. those receiving linguistic input from highly responsive caregivers), this specialisation to the native language occurred already by six months, as evidenced by inability to discriminate non-native phonemes in a looking paradigm. Little is known about how autistic infants acquire speech discrimination skills, however Kuhl et al. (2005) used an oddball paradigm to detect physiological evidence of phoneme discrimination in autistic preschoolers. Unlike a mental-age matched TD group, the autistic group did not demonstrate discrimination, suggesting weaker attunement to the native

language, and suggesting further evidence of the socially-gated learning hypothesis.

Interestingly, research has suggested that the process of acquisition of vocalisation skills may also be socially shaped. Gros-Louis (2014) reported that between eight and 14 months, the quality of maternal responses to vocalisations predicted the amount and quality of later infant vocalisations. If parent interaction is a key ingredient to phonemic learning, this may disadvantage autistic infants. Parents' contingent input relies on there being something to respond to, i.e. child intentional communication (gaze, gesture, object exploration), which may be less frequent in autism (Vallotton, 2009) resulting in an impaired social feedback loop (Warlaumont & Oller, 2014). Autistic infants may thus not benefit from or elicit contingent responsiveness from caregivers and thus take longer to attune to their native language.

There are however problems for the social motivation account: the evidence of masking behaviour in many cognitively and verbally able autistic individuals would suggest that they do have inherent social motivation (Fletcher-Watson & Happé, 2019; Jaswal & Akhtar, 2018). Additionally, no evidence has been found of atypical overt social skills before 12 months in high-likelihood infants (Fletcher-Watson & Happé, 2019).

1.2.3. Joint attention

Autistic infants may fail to synchronise with their carer, leading to poor dyadic and subsequently triadic relating. As a result, they experience fewer social interactions in which to make world to word mappings, hindering acquisition of grammatical and lexical knowledge. For both autistic and TD children, early joint attention abilities (initiating and responding to joint attention) predict later language competence (Bottema-Beutel, 2016, Charman, Baron-Cohen, Swettenham, Baird, Cox & Drew, 2003; Yoder, Watson & Lambert, 2015). This underscores the importance of social-cognitive skills such as understanding speaker intention in language acquisition

(Baldwin & Moses, 2001; Briganti & Cohen, 2011; Roseberry, Hirsh-Patek & Golinkoff, 2014). Luyster, Kadlec, Carter and Tager-Flusberg (2008) found that response to (and not initiation of) joint attention significantly predicted concurrent receptive language skills in a large sample (n=164) of autistic toddlers. Delincolas et al. (2007) also observed a similar differential relationship between response to joint attention and receptive and expressive language skills in a concurrent study of autistic 2-6 year olds (n=56). They hypothesised that initiating behaviours may be separable and relate more to social than linguistic development, a finding echoed by Pickard and Ingersoll (2014) in 53 autistic children aged 22-93 months. Bottema-Beutel's meta-regression (2016) confirmed the specific importance of response to joint attention for expressive and receptive language development, and concluded that this relation was stronger in autism than TD. The author theorised that this may be because a minimum level of joint attention responding is needed to support typical language, leading to a reduced impact of higher levels (expected to be found in TD) and a heightened impact of lower levels (expected to be found in autistic participants).

1.2.4. Non-social features of autism

Very few studies have explicitly examined the relationship between non-social core autism features and language impairment. Watson et al. (2011) examined associations between different sensory profiles and socio-communicative, language and adaptive measures in four years olds with either a diagnosis of autism or developmental delay (n=116). Across both groups, both sensory hypo-responsiveness and sensory seeking symptoms were significantly negatively correlated with language skills, whereas there was no significant association for sensory hyper-responsiveness. Baranek, Watson, Boyd, Poe, David & McGuire (2013) measured hypo-responsiveness to social and non-social stimuli in a cross-sectional study of TD, developmentally delayed and autistic children (n=178). For autistic participants, social and non-social hypo-responsiveness was significantly negatively correlated with expressive and receptive language. These findings suggest that sensory differences could lead autistic children to experience different

daily opportunities for world to word mapping which contribute to language learning. However, there needs to be more research in this area to build up a clearer picture of the precise role of sensory differences, repetitive behaviours, insistence on sameness and special interests in language learning. For example, it is possible that certain non-social autistic features could be a boon to language learning (e.g. a special interest in letters, words). Repetitive exposure to specific stimuli (such as a favourite book or video) illustrates a way in which autistic learners may shape their linguistic environment to facilitate learning.

1.2.5. Problems for core social account

Early interventions targeting social engagement via enhanced parental responsiveness and building joint attention skills have been developed for autistic children, however language outcomes (when measured) are often small or non-significant (Edmunds, Kover & Stone, 2019). Pickles et al. (2016) reported a long term follow up of the Preschool Autism Communication Trial (PACT), which aimed to enhance parent-child interaction in autistic under-fives (n=121). There was a statistically significant improvement in autism symptom severity but no significant treatment effect on language. Carter et al. (2011) randomized autistic toddlers (n=62) to a control group and an intervention focused on parent responsivity and child communication. There were no main intervention effects on language, however improvements were observed in a subgroup who started the trial with low object interest. Kasari, Gulsrud, Freeman, Paparella & Helleman (2012) conducted a five-year follow up of participants in a three-arm intervention study (joint attention intervention, symbolic play intervention and control group). Language outcomes were predicted by child age at intervention, initial joint attention and play skills. Being assigned to one of the treatment groups was an additional predictor even when these variables were controlled. Children's response rate to social-engagement interventions is variable and difficult to predict, suggesting other factors jointly influence structural language development.

This explanation also fails to account for those who develop appropriate structural language either on schedule or with some delay despite social cognition differences. Some evidence shows that young autistic children can use social information on a basic level for word learning, which is also problematic for a purely social explanation for delayed/impaired language acquisition in autism (Luyster & Lord, 2009). A population based study evaluated 9-14 year old participants for symptoms diagnostic of autism and language impairment, classifying children as either belonging to an ALN (n=31), ALI (n=41) or DLD group (n=25). They found that those qualifying as ALI did not have greater symptom severity than those in the ALN group, evidence which does not support a secondary-effect explanation (Loucas et al., 2008).

1.3. Language impairment as result of co-morbidity with DLD

In determining whether co-morbid DLD could explain language heterogeneity in autism, it is important to determine whether ALI and DLD represent deviation from typical development (TD) in two qualitatively different ways or the same way at the symptom (phenotype) level. This in turn would inform whether the explanation is one of shared aetiology or merely a product of phenomimicry (different causal pathways creating symptom profiles which resemble each other on surface level only, see Figure 3). In this section I will examine support for the co-morbidity hypothesis on a behavioural and brain level.

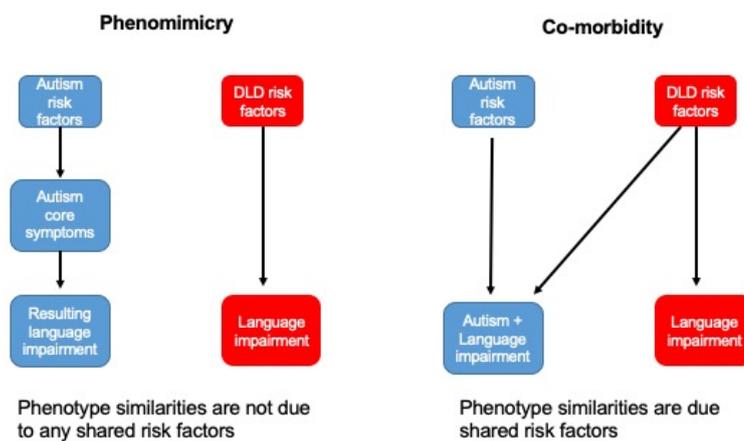


Figure 3. Phenomimicry and co-morbidity explanations of language impairment in autism

Note: DLD: Developmental Language Disorder

1.3.1. Behavioural evidence of shared pathways

1.3.1.1. Shared behavioural profile: blurred boundaries between diagnostic categories

One strand of evidence suggestive of overlapping features of ALI and DLD is provided by those who begin with a diagnosis of one and subsequently or concurrently also qualify for diagnosis of the other. Children with DLD are at higher risk of social difficulties (peer rejection, poor social competence), even if they fall below the threshold for an autism diagnosis (St Clair, Durkin, Conti-Ramsden & Pickles, 2011; Hart, Fujiki, Brinton & Hart, 2004; Loucas et al., 2008). It remains unclear if this is downstream consequence of language impairment (e.g. concentrating on the message and thus failing to integrate social information during interactions) or whether it reflects a separate underlying social cognition deficit which co-varies with structural language difficulties. It is important to be aware that diagnosis is an external construct based on subjective measurement and is not fixed over time. According to current diagnosis criteria, a diagnosis of DLD excludes one of autism. A diagnosis of autism can be specified to be with or without accompanying language impairment, but this is not considered a separate diagnosis rather an additional descriptor of the primary diagnosis (APA, 2013). Social cognition and consequently pragmatic skills are reported to be a strength in DLD, but a subgroup do have semantic-pragmatic difficulties³. Leyfer, Tager-Flusberg, Dowd, Tomblin & Folstein (2008) found 41% of children diagnosed with DLD had additional pragmatic difficulties. Conti-Ramsden, Simkin and Botting (2006) examined autism symptom severity in 76 14-year olds with a history of DLD and found 3.9% who would qualify for a diagnosis of autism and many others with sub-clinical traits. Conti-Ramsden, Botting, Simkin and Knox (2001) found that 5% of children (n=200) who met criteria for DLD at age seven, also scored highly on an autism checklist at age 11. A Danish study followed up 469 individuals diagnosed in childhood with DLD, and found a significantly higher rate of autism

³ Previously also known as Semantic-Pragmatic Disorder or Pragmatic Language Impairment, now referred to as Social Communication Disorder

diagnosis 35 years later (2.1%) compared with 2,345 typical controls (0.1%, Mouridsen & Hauschild, 2009). Mawhood, Howlin and Rutter (2000) reassessed children diagnosed with either DLD or autism at age 7-8 as adults and found that distinctions grew less clear between the groups as they aged, with only a quarter of adults with DLD presenting with typical social functioning (Howlin, Mawhood & Rutter, 2000). It is unclear whether this stems from a snowball effect of latent social cognition problems during childhood in DLD or is a downstream effect of their language impairment (see Figure 4). Another explanation could be different degrees of educational support and stability in the DLD group compared to the autistic group (Mawhood et al., 2000).

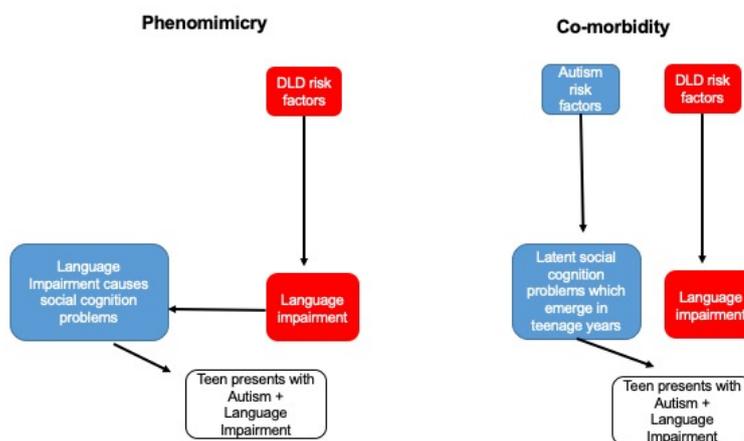


Figure 4. Phenomimicry and co-morbidity explanations for social cognition risk factors in DLD

Note: DLD: Developmental Language Disorder

Roy and Chiat (2014) also observed a changing profile over time, in their follow up study of participants diagnosed with language problems at age 2.5, many of them went on to receive diagnoses of autism or social and emotional behavioural disorders. Social communication abilities, like structural language, can be thought of as on a continuum. In clinical groups a poor score on one is moderately correlated with poor score on the other (Bishop & Norbury, 2002). Botting and Conti-Ramsden (2003) investigated commonly used psycholinguistic markers in clinical samples and noted they were able to distinguish the clinical sample from

TD but between diagnostic categories (autism, DLD and Social Communication Disorder) they had limited specificity.

1.3.1.2. Shared behavioural markers

Children with ALI's poor performance on measures typically used as clinical markers of DLD has been put forward as a further argument for co-morbidity (Kjelgard & Tager-Flusberg, 2001). These markers include: poor non-word repetition, morphosyntax errors and oromotor difficulties. I will examine the evidence of behavioural overlap for each in turn. A key methodological difficulty in evaluating this evidence, however, is that many studies compare performance of children with DLD to an autistic group with mixed or unspecified language profiles (ALN and ALI), whereas others are purely comparing ALI with DLD. In order to assess whether individuals with ALI share risk factors with those with DLD, it is crucial to exclude those with language within the normal range.

Non-word Repetition

Several studies report that participants with ALI and DLD share impairment in non-word repetition, with errors being similar in nature (phoneme deletion or substitution), though others argue for qualitative differences in error patterns (Tager-Flusberg, 2006; Lindgren, Folstein, Tomblin & Tager-Flusberg, 2009; Tager-Flusberg & Joseph, 2003; Roberts, Rice & Tager-Flusberg, 2004, Whitehouse et al., 2008). Whether non-word repetition can be characterised as a DLD-specific marker is debatable: (non-word repetition can detect DLD endophenotype even if resolved, Barry, Yasin, & Bishop, 2007; but is also present in Down syndrome, Eadie, Fey, Douglas & Parsons, 2002). Non-word repetition taps a number of underlying processes including speech perception, phonological memory, articulation and the construction of phonological representations (Coady & Evans, 2008), and could thus fail for many different reasons (regardless of a child's diagnostic status). Equally, qualitative differences observed between children with DLD and ALI may reflect additional cognitive or behavioural deficits rather than indicating a fundamentally different source of error. Some investigators

have found contrasting error patterns in non-word repetition. Whitehouse et al. (2008) compared children with DLD, ALI and ALN on non-word repetition. For stimuli with 2-3 syllables DLD and ALI performance was matched but on 4-5 syllables the children with DLD scored much lower. The authors suggested that core autism characteristics may be responsible for poor ALI group performance on the non-word task, rather than language impairments overlapping with DLD. Hill, van Santen, Gorman, Langhorst and Fombonne (2015) performed a similar group comparison and observed that children with DLD performed worse than children with ALI on the 2-3 syllable non-words. Given these equivocal findings and the small samples used in both studies, I do not think we can conclude that ALI and DLD do not share causal pathways. Difference in group performance does not demonstrate that the source of impairment is different but may be just reflecting a different profile of additional strengths/weakness due to non-shared cognitive factors, e.g. DLD clinical groups may perhaps be more affected by articulation impairments, poor literacy or poorer phonological storage capacity (Gathercole, 2006).

Morphosyntax

Morphosyntax is the combinatory rule system for a language, governing aspects such as verb and noun inflections and the appropriate use of function words. According to Rice, Warren and Betz (2005), in DLD, morphosyntax is out of sync for a longer period than other aspects of structural language (a delay within a delay). Tense marking is a syntactic feature commonly found in DLD (Bishop, 2014). Young TD children commonly fail to mark the third person in finite verbs (e.g. “he walk_ the dog” vs. “he walks the dog”, Rice, Wexler, & Hershberger, 2001) but in DLD tense marking is acquired later in development or remains absent. The specificity and universality of tense-marking errors to DLD is unclear (Roberts et al., 2004, Williams et al., 2008).

Some early work highlighted delayed or deviant syntax and morphology in autism (Bartolucci, Pierce & Streiner, 1980). Roberts et al. (2004) found similarities between ALI and DLD groups in tense errors in elicited sentences and Tager-

Flusberg found equivalent tense-specific difficulties in natural speech samples (2006). However, error analysis of Roberts et al.'s data revealed past tense errors were only made on irregular verbs by the ALI group rather than all verbs in the DLD group (Williams et al., 2008). Again due to differing cognitive profile of strengths/weaknesses, the ALI group could be a developmental stage ahead of the DLD group in morphosyntax development. This does not exclude the possibility of shared causal pathways. In natural speech samples, young autistic children exhibited a greater degree of *developmental scatter*, that is, they produced grammatical structures that are less predictable based on previous productions (Eigsti, Bennetto, & Dadlani, 2007). Eigsti et al. (2007) observed delayed syntactic development in a detailed analysis of grammatical forms used by a sample of verbal, cognitively able five year olds, compared with TD controls and those with other developmental delays. Presentation of specific grammatical error patterns in either disorder may vary depending on the developmental time point examined. Matching groups on non-verbal IQ or vocabulary size does not mean they will be at the same developmental point with grammar due to their other cognitive 'hits' or compensatory processes. In a study of verbal older children with a history of language delay, Eigsti and Bennetto (2009) discovered subtle but significant differences between TD and autistic groups' performance in an offline meta-cognitive grammaticality judgement task. More research is needed on early emerging morphosyntactic profiles in both populations to further establish the similarities and differences. Cross-disorder comparisons need to also examine receptive grammar skills (e.g. Fortunato-Tavares et al., 2015).

Oromotor difficulties

Kjelgaard and Tager-Flusberg (2001) examined group differences in oromotor skills and found limited support for shared features between DLD and ALI. The ALI group performed worse than ALN on phonology, but still within the normal range (i.e. better than the DLD group). Proponents of the view that language impairment in autism results from core social features rather than co-morbidity with DLD, cite relative strengths in phonology as evidence for distinct aetiologies (Williams et al.,

2008). In this section I will examine the evidence for and against this claim and attempt to reconcile the findings.

Early studies refer to delayed phonology in autism, e.g. Bartolucci, Pierce, Streiner & Eppel (1976). More recently, Rapin and Dunn (2003) found that 63% of autistic preschoolers had mixed receptive/expressive disorder that included phonology and grammar. Rapin, Dunn, Allen, Stevens and Fein (2009) re-examined these participants at 7-9yrs and found 11% had typical language, 73% higher order language problems only, and the rest were mostly globally impaired, with a few who had typical comprehension and impaired articulation. “The neurologists detected oromotor deficits (oromotor dyspraxia/dysarthria, a motor—not linguistic—deficit) ... in Clusters 1 and 2” (p.77). Further suggestion of impaired phonological decoding behind language impairment in both disorders is “the few non-verbal children with autism or DLD in Allen’s nursery who learned to speak words after they had learned to read them” (p.77).

In a recent assessment of a community sample of autistic 4-6 year olds without intellectual disability, almost 60% were found to have moderate to severe language difficulties, with 21% being identified as having speech-related difficulties (Kjellmer, Fernell, Gillberg & Norrelgen, 2018). Children with either a DLD or autism diagnosis are both frequently noted to have delayed first words (Trauner, Wulfeck, Tallal & Hesselink, 1995, Howlin, 2003). The extent that this reflects articulatory output processes rather than a lack of word knowledge in autism has not been fully explored in the literature.

Autistic children’s pre-speech vocalisations have atypical acoustic features (Oller et al., 2010). Young autistic children make fewer speech-like vocalisations relative to typically developing peers (Warlaumont & Oller, 2014; Plumb & Wetherby, 2013). Investigations of infant siblings of autistic children (who have a higher chance of obtaining an autism diagnosis), also indicate early differences in vocalisation rate and quality (Paul, Fuerst, Ramsay, Chawarska & Klin, 2011; Chenausky, Nelson & Tager-Flusberg, 2017; Heymann et al., 2018; Patten, Belardi, Baranek, Watson, Labban & Oller, 2014).

A growing literature also suggests that early vocal development difficulties may strongly impact spoken language development in autism. A recent meta-analysis concluded that pre-verbal vocalisations are correlated with concurrent and later expressive language in young autistic children (weighted effect size of $r=.50$, McDaniel, Ambrose & Yoder, 2018). Smith, Mirenda and Zaidman-Zait (2007) found verbal imitation ability (scored simply as present or absent) significantly predicted later language milestones.

A more specific oral motor dysfunction could contribute to speech delays in some autistic children (Adams, 1998). Belmonte et al. (2013) described a subset of autistic children whose receptive language outpaced their expressive skills, and these same children also had marked initial and ongoing oro-motor difficulties. Tierney, Mayes, Lohs, Black, Gisin and Veglia (2015) observed high co-morbidity of autism and apraxia in a clinical sample (of 11 autistic individuals, seven also met criteria for apraxia of speech). Chenausky et al. (2019) used a cross-sectional approach to group autistic participants with poor expressive language skills ($n=54$) into several groups depending on the presence or absence of characteristics associated with childhood apraxia of speech. In participants with no suspicion of speech-motor difficulties, receptive language predicted concurrent expressive language (as measured by number of different words spoken). In both the group suspected of apraxia and the group with insufficient speech to determine apraxia status, speech skills were the only significant predictor. Other authors have demonstrated difficulties with motor praxis more generally in autistic compared with TD populations (Dziuk, Larson, Apostu, Mahone, Denckla & Mostofsky, 2007, MacNeil & Mostofsky, 2012), with motor skill in infants being predictive of later autism symptom severity in a high-likelihood cohort (Leonard & Hill, 2014).

Finally, in older children who have acquired spoken language, sub-clinical phonological problems have been found. Both Cleland, Gibbon Peppé, O'Hare and Rutherford (2010), and Shriberg et al. (2001) found sub-clinical articulation difficulties in ALN groups.

Lack of comparable oromotor findings between DLD and ALI in comparison studies could be due to sampling bias - minimally verbal individuals, whose phonology and oromotor skills may be severely impacted, are often excluded from studies as they perform at floor on language or IQ tests or may not be able to follow or comply with task demands. Meanwhile, the least affected children with DLD who may have relatively intact phonology, tend to be excluded by dint of not being part of a clinical case-load (Tomblin, Records, Buckwalter, Zhang & O'Brien, 1997; Tomblin, 2011; Zhang & Tomblin, 2000).

Williams et al. (2008) points to the rapid resolution of any articulatory/phonological problems in autism versus the slower improvement in phonological skills observed in DLD to support claims that phonological deficits do not share a common causal mechanism across disorders, however detailed longitudinal investigations of phonological development in both groups are lacking and such assumptions may be due to the sampling bias discussed above. Analyses of phonology in autism have often excluded consideration of articulation/phonology in autistic children who remain minimally verbal. Equally, at a population level, relatively few children with DLD actually have articulation difficulties. Shriberg, Tomblin & McSweeney, (1999) examined six year olds in a population level sample, and found of those meeting DLD criteria only 5-8% had a persisting speech disorder.

Is this also an example of developmental trajectories diverging because different strengths compensate? Critics of the co-morbidity explanation posit that it cannot be that autistic children who have an additional 'impairment', can overtake children with DLD in grammatical and articulatory development, since the latter only have one 'pure' impairment and thus more compensatory skills:

"If linguistic deficits in each disorder have the same underlying basis then it is not obvious why the group with more severe and widespread initial deficits (i.e., the autism group in Bartak, Rutter & Cox, 1975, study) should show greater compensatory flexibility than the [DLD] group, whose early social, communicative and linguistic abilities were all superior to those of the [ALI] group" (Williams et al., 2008, p.949)

However, children with ALI may have developed atypical compensatory skills which they bring to bear in language learning process, and children with DLD may have additional non-shared cognitive ‘hits’ that further constrain linguistic development (e.g. phonological storage, poor literacy development). Children with an autism or DLD diagnosis may receive quite different educational and therapeutic input, which might contribute to compensatory flexibility (Mawhood et al., 2000). Both groups present along a continuum and group comparisons may belie the individual factors impacting language development, some of which may be shared across multiple disorders (e.g. attention, motor, executive functioning).

1.3.1.3. Other behavioural similarities

Other comparisons of DLD and ALI on a behavioural level hint at similarities in semantic processing and learning biases (e.g. McGregor et al., 2011; Riches, Loucas, Baird, Charman & Simonoff, 2012). However, Loucas, Riches, Baird and Pickles (2013) tested spoken word recognition and found participants with ALI to be differentially weaker, which could be due to non-shared cognitive strengths/weaknesses. Haebig, Kaushanskaya and Ellis Weismer (2015) examined the role of semantics in lexical processing using an ALI, DLD and TD comparison, and found that updating and shifting abilities predicted lexical processing in all participants. This reveals possible shared causal mechanisms. Similarly, eye-tracking data has been used to examine online processing differences to reveal whether similarities between DLD and ALI groups are due to common processing limitations or are net result of different dysfunctions (see review by Norbury, 2014). Other research has investigated the late development or absence of a shape bias in autism and what this might say about word learning mechanisms (Tek, Jaffery, Fein & Naigles, 2008; Field, Allen & Lewis, 2016), it has also found to be absent in late talkers and those with DLD (Arunachalam & Luyster, 2016). In behavioural tests of prosody comprehension and production both groups display weaknesses (Marshall, Harcourt-Brown, Ramus & van der Lely, 2009; Peppé, Cleland, Gibbon, O’Hare & Martínez Castilla, 2011). Other co-morbidities common to both disorders include ADHD, anxiety, motor impairments and

dyslexia. All of these are likely to further shape language learning across development.

1.3.1.4. Notable differences between language development in DLD and autism

There are certainly some aspects of structural language, which appear to be selectively impaired in autism and not DLD, and one could suggest that these have the most direct link with pragmatic weaknesses. Deixis (use of certain referential terms like pronouns) depends on understanding the interlocutor's perspective but may also depend upon processing demands (Brehme, 2014). Echolalia (which could also reflect an inability to understand speech roles) is also commonly observed in autism (Sterponi & Shankey, 2013). These two difficulties arise briefly in TD (Bloom, Rocissano & Hood, 1976) but are not reported as persisting in DLD (Adamo, 1996, cf. van Santen, Sproat & Hill, 2013). Again, not all autistic children go through a phase of echolalia or impaired deixis, indicating that a combination of individual factors influence this aspect of language too. However, these two aspects can arise and persist in autism in the absence of other ALI-like structural language weaknesses (van Santen, Sproat & Hill, 2013). Little is known about these specific phenomena during early language development across the spectrum.

Another notable difference in language trajectories between groups is the much higher degree of language skill regression reported in autism (Barger, Campbell & Mcdonough, 2013; Goin-Kochel, Esler, Kanne & Hus, 2014; Thurm, Manwaring, Luckenbaugh, Lord & Swedo, 2014). One of the few studies to examine language regression in DLD reported only 1% of participants were affected (Pickles et al., 2009). Regression was historically believed to only affect 25% of autistic individuals, however recent prospective approaches are revealing higher prevalence, thought to be due to biases in retrospective studies, which focussed on overt expressive behaviours (Ozonoff & Iosif, 2019). A sustained plateau rather than loss of language skills has also been revealed for some participants (Shumway et al., 2011). This aspect of autistic language trajectories remains

poorly understood, and any causal model with overlapping aetiologies needs to account for this fairly clear cut difference in groups.

Finally, in terms of common co-morbidities, epilepsy is far more prevalent in autism than in DLD or TD. The relationship between epilepsy and language development remains under-researched and poorly understood (Ballaban-Gil & Tuchman, 2000).

1.3.1.5. Summary

In sum, there is evidence for behavioural similarities between children with DLD and those with ALI: the higher incidence of autistic traits in DLD, particularly emerging in late childhood; shared difficulties in non-word repetition, morphosyntax and in some cases speech production, particularly at an early stage of language learning.

There is also evidence that DLD and ALI linguistic profiles diverge over time (Geurts & Embrechts, 2008) and comprise specific non-shared features, notably regression, echolalia, deixis and the greater tendency to have no spoken language at all.

If we assume that linguistic profiles are shaped by a range of child-specific cognitive risk factors and environmental inputs, an explanatory model where children with ALI and DLD share some causal risk factors but not others, could explain both the similarities and differences. This underscores the importance of investigating early language trajectories in autism and DLD in order to determine whether this divergence reflects different underlying causal mechanisms, or shared mechanisms which combine with non-shared features to produce dissimilar trajectories. It is also important to examine specific skill deficits in the context of overall child profile and not in isolation.

A methodological difficulty thus far is that some group design studies seek to identify whether there are shared risk factors common to all autistic children (regardless of language status) and those with DLD, whereas other researchers

have been asking a different question: do autistic individuals who also have language deficits share risk factors with those of DLD. It is difficult to weigh the evidence when it is not explicit which of the two questions is being asked.

1.3.2. Neurobiological evidence of shared pathways

Given the behavioural similarities, researchers have examined possible similarities on a neurobiological basis (genes, brain morphology, functional imaging). There is some evidence for shared genetic risk factors, however neuroimaging data suffers from small and poorly matched samples (where age, task, performance and language profile may not be well-equated) and is therefore much harder to interpret.

1.3.2.1. Genetic evidence

Both DLD and autism are highly heritable, indicating they are each at least partly genetically influenced. No classic Mendelian pattern of heritability exists, thus they are likely to be due to polygenic and multifactorial risk factors, as for many other common human conditions (Bishop, 2010, Betancur, 2011). Results from family studies are mixed on whether relatives of probands with autism and DLD have different linguistic and cognitive profiles or not. Two studies found language difficulties in relatives of autistic individuals (Bailey, Palferman, Heavey & Le Couteur, 1998; Piven et al., 1997). Tomblin, Hafeman and O'Brien (2003) found increased occurrence of autism in siblings of probands with DLD. Gamliel, Yirmiya, Jaffe, Manor and Sigman (2009) found heightened DLD risk in siblings of autistic individuals. Ruser et al. (2007) found a higher rate of communication impairments in parents of probands with autism and DLD (vs. Downs Syndrome controls). However, a family study comparing ALI and DLD groups found relatives had different profiles:

“Given that ALI relatives performed better than [DLD] relatives across the language measures, the hypothesis that ALI and [DLD] families share similar genetic loading for language is not strongly supported.”
(Lindgren et al., 2009, p.22)

Bishop (2010) examined the ‘correlated additive risks’ model of heredity using simulation techniques. This model supposes independent causal pathways for DLD and autism, with a correlation between the risk factors to explain the higher

than expected co-occurrence of both diagnoses. Prevalence rates for probands and their relatives were simulated using hypothetical gene models with different degrees of pleiotropy, to reflect different levels of correlation between risk factors. The simulation predicted a lower than observed prevalence rate of ALI and failed to replicate observed family study data (e.g. Lindgren et al., 2009). A second simulation employed a gene x gene interaction (epistasis) and this generated data more consistent with family study data and prevalence rates. In epistasis certain genes have an increased impact when they are found in the presence of specific other genes. Reduced similarity between probands and first degree relatives is a hallmark of epistasis. Similarly, Pickles et al. (1995) posits that based on monozygotic twin co-morbidity rate of 80%, the dizygotic twin and siblings rate should be 50% not 10%, which supports the existence of non-additive gene interactions.

Some specific genes for language impairment (occurring in both individuals with ALI and DLD) have been identified, giving further weight to the co-morbidity hypothesis. Variants of the Contactin-Associated Protein-Like 2 gene (CNTNAP2) have been more broadly linked with language impairment across disorders. Vernes et al. (2008) found performance on behavioural tests for markers of DLD correlates negatively with polymorphisms on the CNTNAP2 locus in children with DLD. Alarcon et al. (2008) found a link between the same CNTNAP2 polymorphisms and age of first word in autistic probands. CNTNAP2 encodes neuronal transmembrane proteins involved in cell adhesion and has shown enhanced expression in language-related brain circuits (Abrahams et al., 2007). CNTNAP2 is a downstream target of Forkhead box protein P2 (FOXP2), which has been implicated in a rare, inherited, monogenic language disorder in the KE family (Fisher & Scharff, 2009). Additionally, Whitehouse, Bishop, Ang, Pennell and Fisher (2011) revealed that the same polymorphisms in CNTNAP2 are associated with a measure of language development at age two in an epidemiological study of over 1,000 typically developing children. There is much genetic evidence for shared risk factors for language problems across other neurological disorders. Eicher and Gruen (2015) support the idea of multiple generalist genes, which are implicated in many neurological developmental disorders. They found a

relationship between risk genes previously implicated in DLD and dyslexia, and receptive language skills in a large autistic sample (n=1,990).

Summary: genetic evidence points to overlapping genetic risk factors, and even the family study data can be explained by epistasis. This does not negate the possibility (likelihood) of genetic factors specific to either disorder that may also impinge on language learning.

1.3.2.2. Brain structure evidence

Synthesis of studies exploring brain structure correlates of language impairment in autism are hampered by issues of small sample size, recruitment bias and the evolving definitions of both autism and DLD. Studies with autistic participants often do not report language abilities or relate them to neuroimaging findings, and DLD studies which do this rarely have autistic controls to provide a comparison. Ages of participants also vary widely, making results difficult to interpret.

In typical development we observe a Left > Right asymmetry in frontal language areas, which is reduced or reversed in ALI and DLD but not ALN groups (De Fossé et al., 2004; Herbert et al., 2005; Floris et al., 2016), which could signal a shared neurobiological pathway for language impairment in both disorders. However, these structural differences are not always matched by atypical activations in functional imaging studies (see below).

Planum Temporale reversed asymmetry compared to typical development has been found in autistic (Rojas, Bawn, Benkers, Reite & Rogers, 2002) and DLD groups (de Guibert et al., 2011), however relatively few studies have compared the degree of atypical lateralisation with language skills or development history (cf. Gage et al., 2009).

There have been structural differences found in the cerebellum in participants with ALI which are similar to those seen in DLD (Hodge et al., 2010). These are

abnormalities in neurodevelopment of fronto-corticocerebellar circuits that manage motor control and the processing of language, cognition, working memory, and attention. D'Mello, Moore, Crocetti, Mostofsky and Stoodley (2016) looked at autistic individuals with versus without early language delay, as well as TD controls. They analysed the grey matter volume of the whole brain and found differences centred on the cerebellum. As age of first word or phrase went up, cerebellular grey matter decreased. Right cerebellum grey matter volume was reduced for both autistic groups. Left was only reduced for the autistic individuals with early language delay. There is still much to be learned about the role of the cerebellum in autism (Fatemi et al., 2012; Magal, 2013, Stoodley, Sillitoe & Colledge, 2014). The cerebellum is thought to play a role in prosody and procedural learning, both of which may be impaired in individuals with both DLD and ALI:

“We suggest that dyspraxia in autism involves cerebellar mechanisms of movement control and the integration of these mechanisms with cortical networks implicated in praxis.” (Miller, Chukoskie, Zinni, Townsend & Trauner, 2014, p.95)

In a small study (n=5), minimally verbal autistic children were compared to TD controls and neuroimaging revealed atypical lateralisation of the white tract arcuate fasciculus (Wan, Marchina, Norton & Schlaug, 2012).

In summary: there are replicated findings of reduced or reversed frontal language area Left > Right asymmetry, as well as cerebellular abnormalities being associated with language difficulties across disorders, however evidence is limited by methodological issues outlined above.

1.3.2.3. Functional imaging evidence

The fMRI literature is plagued by similar issues to the brain morphology one. Direct fMRI comparisons between those with DLD and ALI are lacking. The tasks employed during these studies are also diverse and sometimes result in a wide disparity between task success within and between groups, so it is not clear if the

same cognitive process is being measured in each participant, e.g. if a task is not even attempted then the targeted cognitive process is not measured, and no conclusion can be drawn about that process being impaired or not.

Methodological difficulties notwithstanding, right hemisphere hyper-activation of language areas both during passive and active language tasks is frequently found in autistic participants of all language abilities. Activation of these areas seems to get more bilateral with age, whereas in TD population the reverse is true, left lateralisation increases (Flagg, Cardy, Roberts & Roberts, 2005; Seery, Vogel-Farley, Tager-Flusberg & Nelson, 2013). Herringshaw, Ammons, DeRamus and Kana (2016) performed a meta-analysis of fMRI data on all types of language processing task in autism and concluded that in lots of areas autistic and TD performance overlap. Key areas of difference are: 1) increased activation in right hemisphere language areas, Superior Temporal Gyrus and Inferior Frontal Gyrus (particularly when performance is deviant); 2) increased recruitment of left lingual gyrus (particularly when performance is typical); and 3) hypo-activation of Middle Temporal Gyrus bilaterally. To the extent that activation for language processing is more right biased, this suggests commonalities with studies on participants with DLD (de Guibert et al., 2011).

However, there are many contradictory findings. Redcay and Courchesne (2008) found that early Right Hemisphere recruitment was predictive of a better language outcome in their study of autistic toddlers. Atypical activations are also observed despite age appropriate performance on language tasks (decreased frontal along with increased posterior language activation, Just, Cherkassky, Keller, & Minshew, 2004).

In one of the few direct comparisons between participants with ALI and DLD, Whitehouse and Bishop (2008b) measured cerebral dominance using functional transcranial doppler (fTCD) ultrasonography, assessing blood flow through the middle cerebral arteries in adults (with either ALI, DLD, TD and a group with a history of DLD). Here, the DLD group presented a greater right than left hemisphere activation during a word generation task while the reverse was true for

all other groups, despite no differences in task performance – indicating potential neurobiological divergence between the two language impaired groups. However, this study had a small sample size given there were four groups (n=42) and attempts to replicate the finding have not been successful (Bishop et al., 2014; Wilson & Bishop, 2018).

Summary: adequately powered functional imaging evidence directly comparing participants with ALI and DLD is lacking. Longitudinal studies which measure language abilities in a multi-dimensional way and include the full spectrum of language impairment in autism, as well as controls with DLD and TD, are necessary to further elucidate links between linguistic performance and a common activation pattern across or within disorders.

1.4. Multifactorial explanation for language impairment in autism and DLD

It is likely that the factors contributing to both autism symptomatology and language impairment are multifactorial and interactive.

Influential causal factors may include enhanced abilities as well as deficits, which would either give rise to compensatory processes or compounding effects respectively. The resulting outcome for any given individual will depend on the interaction between and severity of a host of positive and negative factors, resulting in within-disorder heterogeneity. Against this backdrop, there may be many routes to impaired language in autism, as illustrated in Figure 5.

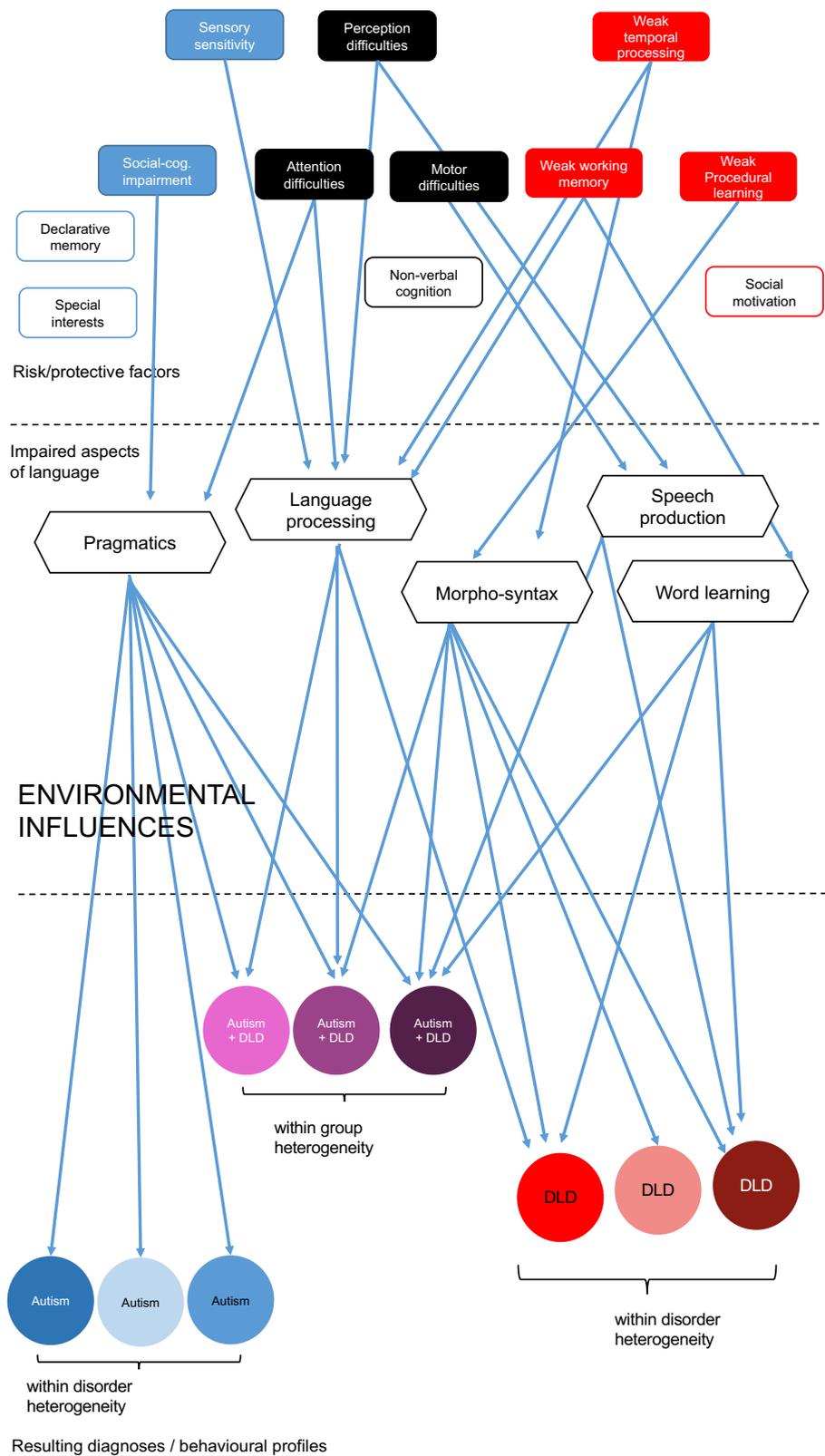


Figure 5. Multi-factor interactive model of shared and non-shared factors

Note: DLD: Developmental Language Disorder.

A crucial question for research is to establish which abilities (cognitive, perceptual, motor, etc.) are necessary for autistic individuals to eventually achieve typical language (as measured by standardised tests of grammar and vocabulary, whilst acknowledging that there may still be subtle differences in language abilities between ALN and TD language). If language is impaired via multiple routes in autism, this reduces the fruitfulness of group based comparisons. A more nuanced approach may be to look at intermediate level endophenotypes that act as risk or protective factors: procedural learning, memory, attentional or perceptual capabilities that are either atypical or out of sync with developmental stage.

Below I will discuss potential key abilities and outline their evidence base pertaining to language development in autism, and where relevant, DLD. However, it is first important to consider how these within-child abilities interact with the external environment, particularly the linguistic input that the child is exposed to. As Figure 6 illustrates, in order for caregiver input to translate into language acquisition, a number of things need to happen:

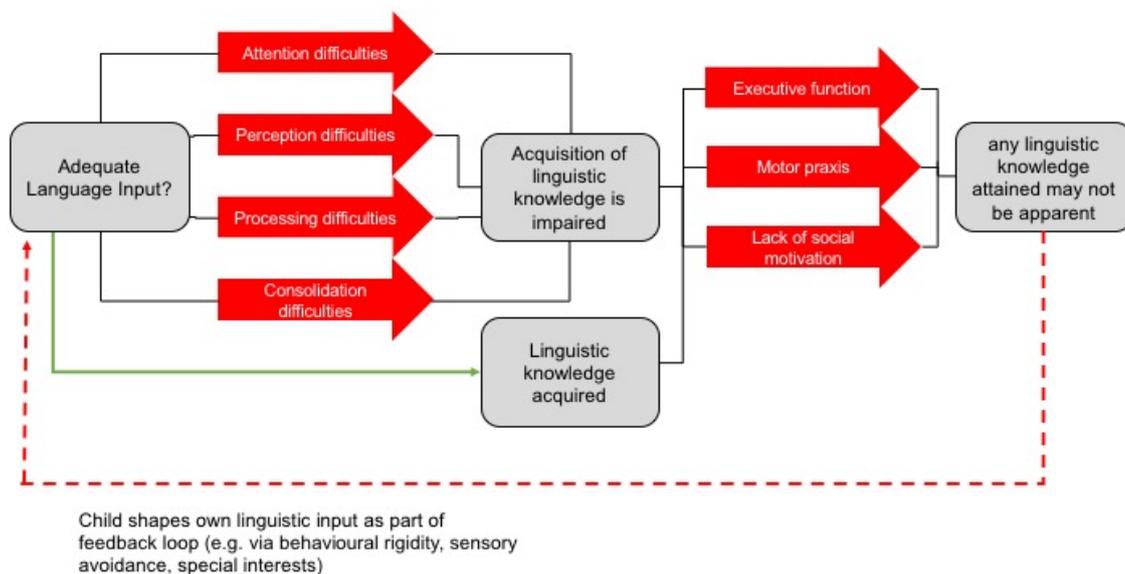


Figure 6. Opportunities for disruption in language learning

The learner must be exposed to the input. Exposure to caregiver speech is assumed to play a key role in fostering language acquisition. The impact of quality and quantity of infant-directed speech on language has been identified in the TD literature (Hoff, 2003; Rowe, 2012). A key ingredient is assumed to be the contingent nature of parent input (i.e. infant-directed speech which relates to the child's immediate focus of attention, Bornstein & Tamis-LeMonda, 1997). Autistic children may limit their own exposure to social interaction, thus reducing quantity and type of input they are exposed to. Here any additional motor, executive function or attentional global deficits will affect the amount of embodied experiences the child has and thus their chance to accumulate exposure to linguistic or other regularities in the social world (see also de Barbaro, Johnson, Forster & Deák, 2013, who describes how sensorimotor advances in infants may drive the 12-month revolution). The importance of contingent responsiveness in the input received by children with expressive language delays has been documented (e.g. Girolametto, Weitzman, Wiigs & Pearce, 1999; Hudson, Levickis, Down, Nicholls & Wake, 2014; Yoder & Warren, 1999). McDuffie & Yoder (2010) found that parent follow-ins and responses to communication acts predict unique variance in expressive language growth in a group of autistic under-fives with expressive language delay. This suggests that enhanced input may counteract the possible effects of language learning deficits in some children. One could posit that in this aspect children with DLD, who are relatively less impaired in social cognition or cognitive flexibility, may be exposed to quantitatively and qualitatively more linguistic input in the early years.

The learner must attend to the input. The role of attention in language learning is not clearly delineated – on the one hand, some implicit learning can take place without attention (e.g. Saffran, Sengas & Trueswell, 2001), but it remains to be determined how much complex language learning can occur without some engagement with the input. Difficulties with inhibition, an aspect of executive functioning, may relate to verbal skills in DLD and autism (Bishop & Norbury, 2005). See also '**Attention**' and '**Executive function**' below.

Input must be accurately perceived by learner. Autistic learners could have an additional hit to the perceptual system, whereby complex, social, or multimodal stimuli are particularly difficult to perceive. This could take the form of a top down deficit or a bottom up impairment but its result would be to reduce the amount of total input that can be fed into the language learning system, even if it was 'experienced' and 'attended to'. In this, some evidence suggests similar or overlapping perceptual deficits with DLD (e.g. Tallal & Piercy, 1973). See also '**Perception**' below.

Input must be processed efficiently by the learner. Speed of processing has been implicated in some but not all cases of DLD (Miller, Kail, Leonard & Tomblin, 2001). The sparse data on speed of processing in autism reveal a relative strength (Scheuffgen, Happé, Anderson & Frith, 2000; Wallace, Anderson & Happé, 2009). Inspection Time refers to the exposure duration needed to reliably identify a simple stimulus. Inspection Time speeds in autistic participants with intellectual impairment are discrepantly fast versus matched non-autistic peers with intellectual impairment, and in intellectually able groups the negative correlation between Inspection Time and measured IQ, which is robustly observed in TD, breaks down. Data comparing speed of processing in children with ALI and DLD is lacking in order to explore whether this relationship holds in language impaired groups.

Input must be subject to implicit or explicit learning mechanisms in order to translate into linguistic knowledge. See '**Procedural learning**' below.

The learner must consolidate what is learned. Consolidation of learned material appears to be an important component of learning, i.e. the processes that go on after training to stabilize the fragile initial learning trace and place it in memory. Some findings suggest children with DLD are particularly weak at consolidating implicitly trained (non-linguistic) sequences (Desmottes, Meulemans & Maillart, 2017), however replications using linguistic stimuli are lacking. Henderson, Powell, Gaskell and Norbury (2014) and Fletcher, Knowland, Walker, Gaskell, Norbury and Henderson (2019) compared children from ALN and TD groups in word learning

tasks, measuring word learning performance before and after sleep consolidation. Autistic participants did not show typical effects of competition with similar phonological forms post-consolidation (e.g. biscial vs. biscuit), suggesting that the new words were not integrated into their lexicon in the same way as for TD children (Henderson et al., 2014). In Fletcher et al.'s study (2019), participants in the ALN group showed reduced consolidation of unique features of new vocabulary compared to TD, implying they gain a less secure semantic representation of words learnt. Some postulate that our ability to consolidate new words is closely related to the level of our existing vocabulary (James, Gaskell, Weighall & Henderson, 2017) therefore future studies need to be closely matched on language ability. To date, consolidation studies involving participants with ALI and DLD are lacking.

Finally, in order to transfer the language knowledge acquired from any learning into functional language use, **the learner requires fundamental motor planning and execution abilities** for expressive language (manual, written or spoken language). See oromotor section above and '**Speech**' below. Expressive language skills of the child will impact the type and amount of input the learner receives from the environment (back to the beginning of the loop illustrated in Figure 6).

A difficulty with any single one or combination of the above stages could have the effect of slowing the acquisition of linguistic knowledge and language capabilities in autism or DLD.

Whilst it is common for autistic children to remain minimally verbal (c. 14-29%), the majority of those with DLD do acquire language skills. Non-autistic individuals who remain minimally verbal are likely to receive a different diagnosis to DLD (e.g. Verbal Auditory Agnosia, Verbal Dyspraxia). These disorders are rare and participants would usually be excluded from DLD studies. It could be hypothesised that the extreme halting of language development at an initial stage resulting in minimal verbal language is more prevalent in autism due to different non-shared risk factors.

1.4.1. Protective factors

Equally, for both autism and DLD, positive factors could have a compensatory action in spite of any difficulties experienced. Abilities to consider include the following:

Declarative memory: i.e. the ability to learn associations between two stimuli in a more explicit fashion, which could support explicit language learning (Ullman & Pullman, 2015). Vocabulary is often (particularly early in development) a strength in autism even in the face of language delays. Perhaps autistic children begin by treating fragments of language as vocabulary items, learning via echolalia and then gradually parse using either implicit learning or explicit hypothesis testing? In other words, they may bootstrap grammatical knowledge from lexical knowledge. Individuals with ALI may have more pervasive declarative memory impairments than individuals with ALN (Boucher, Bigham, Mayes & Muskett, 2008; Boucher, Mayes & Bigham, 2012).

Verbal short term memory: if a child can retain a chunk of input from the speech stream for long enough to form sound-meaning associations, this could support language learning. According to group comparisons verbal short term memory has been reported to be a relative strength in autism, which is not so in DLD (phonological storage deficit), however, this needs to be further understood as a contributing factor to ALI. The trajectory of memory abilities in ALI may be atypical, as Hill et al. (2015) found weaker verbal memory in 5-8yr olds with ALI and DLD vs. ALN. Evans, Selinger and Pollak (2011)'s concurrent study of visual and auditory working memory and event related potential correlates thereof, found deficits in both in DLD suggesting a domain general deficit. Enhanced visual functioning in autism (Samson, Mottron, Soulières & Zeffiro, 2012) may support language learning when visual supports are in place.

Special interests and repetitive behaviours, insistence on sameness. Whilst these core features of autism can present a myriad of behavioural and functional challenges, their role in language learning is potentially positive. To the extent a

child repetitively engages with verbal stimuli (e.g. many hours of viewing one episode or being read one book), this could facilitate language learning by increasing exposure to a smaller amount of input. Many autistic children are extremely interested in letters or phonics at an early age, which may lead to acquisition of phonological representations via a declarative memory route. Unusual intensity of focus on special interests could itself be a by-product of an imbalance in memory strengths and weaknesses (Romero-Munguia, 2008).

Executive function: Both autistic and DLD populations are known to have broad deficits with executive function, which is the term for a range of oversight functions such as working memory, planning, inhibition and attention. However, within group differences are great (Henry, Messer & Nash, 2012; Vugs, Hendriks, Cuperus & Verhoeven, 2013; Pellicano, 2010; Gooch, Hulme, Nash & Snowling, 2014). Executive function is likely to influence multiple aspects of the language learning process as it governs the deliberate control of goal orientated actions (e.g. listening, processing information, producing gestures and words), which can support language learning. However, the relationship could flow in the opposite direction (language facilitates executive function via self-talk and verbally mediated strategies). Gooch, Thompson, Nash, Snowling and Hulme (2016) found executive function was correlated with language skill in children with TD and DLD (n=243) but cross-lagged correlations did not demonstrate causal pathways in either direction. Johnson (2012) argues that executive function may play a compensatory role in neurodevelopmental disorders in that reduced executive function alone may not result in ADHD or autism but it could have developmental consequences when combined with additional risk factors in a multi-hit model, which would explain why deficits are not universal.

Having outlined the potential for both deficits and protective factors to influence the language learning process, I will now explore several of the key areas of potential deficit in further detail.

1.4.2. Procedural learning

Before discussing the potential relevance of procedural learning to language in autism I will outline the theory and key findings as they relate to DLD. Ullman and Pierpoint (2005) proposed that the underlying deficit for those with DLD was to be found in the procedural memory system, which is concerned with slow and implicit acquisition and retention of skills and habits via repetition, and is thought to be required for language skill acquisition. Procedural learning is needed to learn 'compiled' or automatic skills, the acquisition of which free up cognitive resources for other operations. According to the declarative-procedural model of typical language acquisition (Ullman, 2004), declarative memory supports lexical knowledge whilst procedural memory undergirds phonology, morphology, and syntax. It has however been acknowledged that in real language learning one cannot fully dissociate the two systems, for example sensitivity to word boundaries (acquired implicitly) can aid word learning (an explicit process). Likewise, some degree of phonological representation (acquired implicitly) is needed in order to encode vocabulary (an explicit process) (Hsu & Bishop, 2014). Initial evidence for a procedural learning deficit in children with DLD was their reduced ability to acquire structural language rules as expected, yet their preserved ability to acquire adequate vocabulary. Additionally, difficulties with domain general procedural learning were observed in the DLD population.

Subsequent empirical evidence has been established by comparing children with DLD with TD age- and/or language-matched controls in various experiments targeting procedural learning, across different domains (e.g. speech stream segmentation (Evans, Saffran & Robe-Torres, 2009), artificial grammar learning (Hsu, Tomblin & Christiansen, 2014), serial reaction tasks (Lum, Conti-Ramsden, Morgan & Ullman, 2014)). In each case learning is incremental and evidenced by tiny improvements in accuracy or reaction time to novel stimuli where the implicit rules would enhance performance. Some additional tasks do not depend on sequential learning (e.g. weather prediction, rotary pursuit), and findings have been more mixed (Mayor-Dubois, Zesiger, Van der Linden & Roulet-Perez, 2012; Adi-Japha, Strulovich-Schwartz & Julius, 2011; Kemeny & Lukacs, 2010; Hsu &

Bishop, 2014), leading some to question whether the implicit deficit is sequence specific.

Although children with DLD have been shown to learn more slowly in implicit learning tasks, they do learn, depending on length of exposure (Lum et al., 2014; Evans et al., 2009), and complexity of implicit rule (Gabriel, Maillart, Stefaniak, Lejeune, Desmottes & Meulemans, 2013). This suggests either implicit learning does occur but more slowly, or other compensatory systems are used, given enough input. The finding that they learn at the same speed as grammar matched controls might suggest that they have an immature procedural learning system (Hsu & Bishop, 2014). Indeed, the difference between children with TD and DLD in procedural learning ability appears to decrease with age (Lum et al., 2014). One difficulty for these experiments is to ensure that they are not confounded by differences in other component skills required for the tasks such as attention, phonological perception or motor planning (particularly as few studies screen for or exclude co-morbid ADHD or dyspraxia despite their high rate of co-morbidity). West, Vadillo, Shanks and Hulme (2017) recently highlighted the poor reliability of implicit learning tasks used to measure procedural learning in a large study of 7-8 year olds (n=101).

Nevertheless, performance on implicit learning tasks has been found to account for a significant amount of individual differences in language comprehension in TD adults (Misyak & Christiansen, 2012). Siegelman and Frost (2015) found implicit learning to be also independent of general cognitive abilities such as intelligence or working memory. Grammar abilities in TD children correlated with their implicit learning performance (Evans et al., 2009). This relationship does not appear to hold so robustly for children with DLD (Desmottes et al., 2016; Lum, Conti-Ramsden, Page & Ullman, 2012, cf. Gabriel et al., 2013), leading to the suggestion that morphosyntax may be acquired via other mechanisms in this group (i.e. more explicit declarative learning). Lum et al. (2012) found positive correlations between grammatical abilities and procedural memory in a TD group but in the DLD group it was correlated with a declarative memory measure. The procedural deficit hypothesis has also been implicated in Dyslexia (Nicolson & Fawcett, 2007).

However, West, Clayton, Shanks and Hulme (2019) found no significant group differences in implicit learning task performance between typical children and children with dyslexia, when matched on reading age. Furthermore, implicit learning performance was not related to reading ability in either group. The authors postulate that differences in task performance observed in the implicit learning literature may be due to impairments in motor learning.

Could procedural learning be compromised in a similar way in children with ALI? As discussed, it is unlikely that any one process is compromised in all individuals. Experiments testing procedural learning in autism have so far lacked appropriate comparison between language phenotypes and have largely included only those with unimpaired language skills (e.g. Mayo & Eigsti, 2012; see Obeid, Brooks, Powers, Gillespie-Lynch & Lum, 2016 for a review and cf. Mostofsky, Goldberg, Landa & Denckla, 2000). Gordon and Stark (2007) documented an implicit learning deficit in a highly heterogeneous autistic and language impaired group, although longer exposure did lead to some learning. Brown, Aczel, Jiménez, Kaufman and Grant (2010) found no group differences on a range of implicit learning paradigms, however, groups were matched on age and verbal IQ (thus perhaps not recruiting from the ALI population). It is possible that other null result findings were based on tasks that were too deterministic (so declarative learning could be used) or poor group matching.

A procedural learning deficit in ALI and DLD is relevant to the idea of there being a critical period for native language acquisition (Johnson & Newport, 1989). Eigsti and Bennetto (2009) hypothesised that deviant grammatical development might result from the fact that for the child with markedly delayed language onset, the process of language acquisition “*with a cognitive[ly] and socially ‘older’ brain*” (p.1001) might resemble effortful second language learning. Given West et al.’s findings (2017), ensuring the reliability of implicit learning tasks as well as the recruitment of appropriately matched comparison groups to test this hypothesis is critical.

1.4.3. Attention

Attentional engagement is an important multi-sensory mechanism used to aid perception of a complex sensory world, by homing in on a specific object for further processing. 'Sticky attention' shifting, i.e. a long latency to disengage attention from one item to the next, is an early sign of autism, observed at 12 months (Sacrey, Bryson & Zwaigenbaum, 2013). The "Sluggish Attentional Shifting" hypothesis has also been investigated in dyslexia (Hari & Renvall, 2001) and DLD (Dispaldro, Leonard, Corradi, Ruffino, Bronte & Facoetti, 2013).

The importance of the link between language impairment and visual attention is also evident from eye-tracking studies (Norbury, 2014). Autistic children may experience additional attentional difficulties other than slow shifting: tendency to focus on local details rather than global gestalt, reduced attention to social stimuli. These differences in attention may impact language acquisition in that they reduce opportunities for mapping events and objects in the environment to lexical forms.

1.4.4. Perception

Impaired temporal processing, impacting the perception of brief and/or multimodal stimuli has been implicated in DLD (Tallal & Piercy, 1973). Kaganovich, Shumakerm Leonard, Gustafson and Macias (2014) investigated TD children, TD adults, and children with a history of DLD (resolved), and their ability to detect whether a visual and auditory stimuli were presented simultaneously or in rapid succession. TD adults outperformed children (in reaction time, accuracy), and children with history of DLD performed significantly worse than TD children. In this study age of participant and ADHD co-morbidity were additional predictors of performance, suggesting ability to integrate multimodal stimuli matures with age and is susceptible to attentional impairments. A correlation was found between language ability and task performance in the group with a history of DLD and this was replicated with a larger sample (Kaganovich, 2017).

Pons, Andreu, Sanz-Torrent, Buil-Legaz, and Lewkowicz (2013) investigated sensitivity to asynchronous speech by asking which video children chose when one had synchronous audio and one had asynchronous audio. No groups could detect which video was asynchronous at 366ms offset but at 600ms the TD group could and the DLD group could not. Righi et al. (2018) demonstrated that autistic participants were less sensitive than language matched controls in detecting audio-visual asynchrony, and these differences correlated with scores on a standardized language test.

Children with DLD also experience less fused percepts during McGurk effect tests (Norrix, Plante, Vance, and Boliek, 2007). Ability to integrate visual and auditory cues is potentially important in early language learning, in order to learn pairings between sounds and the motor process required to produce them. It also plays a role in facilitating comprehension in sub-optimal conditions, such as speech in noise. Weikum et al. (2007) showed that six month olds can tell native from non-native language based on visual cues alone. Altvater-Mackensen and Grossmann (2014) found audiovisual perception abilities in TD six month olds correlated with later vocabulary size.

Both children with ALI and DLD may struggle with integrating visual and auditory information. The superior temporal sulcus plays role in social (visuo-motor) and auditory perception (Redcay, 2008), and is one of the key shared neuroanatomical areas where subjects with language impairment differ from peers from TD or ALN groups (atypical lateralisation).

Production and comprehension of aurally conveyed emotion via speech prosody has been shown to be impaired in DLD and autism. Interestingly, emotion perception problems have been demonstrated to be multimodal in autism (e.g. facial expressions, voices) whereas in DLD they seem to be limited to speech modality (Taylor, Maybery & Whitehouse, 2012).

An unexpected but consistent finding in the literature is that among autistic participants, there are a proportion of individuals with enhanced auditory

perception, as evidenced by smaller discrimination thresholds for changes in pitch of simple tones (Heaton, Hudry, Ludlow & Hill, 2008). Uncertainty remains over whether the enhanced perception extends to speech or only simple stimuli. Jones et al. (2009) and Samson, Mottron, Jemel, Belin and Ciocca (2006) found evidence for simple stimuli only. Jarvinen-Pasley, Wallace, Ramus, Happe, & Heaton (2008) compared performance of a heterogeneous group of autistic children and language-matched controls on an open-ended task where sentences could be matched with either a depiction of their perceptual (pitch-contour) or linguistic content. Autistic children were better at recalling the pitch contour of a sentence than the controls, however these findings need replication.

How these perception abilities arise or develop over time remains a mystery. Could the enhanced perception of minute variation in speech sounds delay but not prevent the acquisition of basic building blocks of language like the perceptual magnets of phoneme categories (Kuhl et al., 2008)? It may impede the perceptual grouping of speech stimuli and extraction of invariance (despite unimpaired procedural learning skills) if each token of, say, the word “bus” was encoded in such detail to include the speaker’s vocal characteristics (pitch, timbre), sub-phonemic variations due to geographic or sociolinguistic factors. It could ‘flip’ into an asset if and when basic phonology and grammar has been acquired, enabling individuals to retain a more accurate representation of heard speech in short term verbal memory, and thus aid word learning and verbal processing.

According to Markram’s Intense World Theory “*hyper-perception, hyper-attention and hyper-memory*” (Markram, Rinaldi & Markram, 2007, p.77) experienced in autism, render the social world unbearable and lead to sensory shut down. Heightened vigilance to unexpected events enables pattern learning, however if child is on permanent state of vigilance, this may impede detection of rules and extraction of invariance (i.e. unimpaired procedural learning abilities but perceptual features of the environment prevent procedural learning from taking place). Several authors (e.g. Gerrard & Rugg, 2009) have proposed a neuroconstructivist model whereby peripheral sensory abnormalities disrupt compilation of complex

skills; impact on synaptogenesis, synaptic pruning and myelination; and subsequently manifest themselves as autistic behaviours.

1.4.5. Speech

The acquisition of basic speech production skills (i.e. infant babble) was once thought to be merely the unfolding of a pre-programmed biological process, beyond control of environment, however we now know that input influences babble hugely and parent responsiveness is one of the key factors (Gros-Louis, 2014). Studies have also found babble to predict later expressive language in typical development (McGillion et al., 2017). However, interactive vocal learning requires a child to not only have social motivation but also the perceptual, attentional and procedural learning abilities to learn the motor programmes for speech.

Difficulties with motor abilities have already been implicated in language development for both DLD and autism (Bhat, Galloway & Landa, 2012; Hill, 2001; LeBarton & Landa, 2019; Mody et al., 2017). In **section 1.3.1.2.3** above I reviewed the evidence for speech-motor difficulties in autism from early studies (Bartolucci et al., 1976), investigations of apraxic features (Chenausky et al., 2019; Tierney et al., 2015), oromotor skills (Belmonte et al., 2013), and atypical early vocalisations (Chenausky et al., 2017, Patten et al., 2014; Paul et al., 2011; Plumb & Wetherby, 2013; Warlaumont & Oller, 2014). Significant relationships between expressive language development and verbal imitation skills (V. Smith et al., 2007), consonant inventories (Yoder et al., 2015), and vocalisations (McDaniel et al., 2018) have been observed.

A lack of speech output may have subsequent cascading effects on quality of future input received, embodied experiences, chances to make world to word mappings and experience interaction. Speech acquisition could be a 'pivotal skill', which once attained can catapult learning to the next level. Nation and Hulme (2011) examined the impact of learning to read on language processing systems in children with reading disorder. A similar relationship could exist for acquiring speech, once it is acquired, other processing systems become more efficient.

1.5. Conclusion

I have reviewed what is currently known about the heterogeneity in language development in autism. Structural language impairment can have adverse consequences for autistic children's well-being, interacts with core autism features and may compound overall developmental difficulties (as many support strategies are verbally mediated). We thus need to take language differences seriously, by developing a fuller understanding of how such heterogeneity arises and what can be done to alleviate language impairment in autism. Truly understanding language development in autism involves understanding not only how it differs from TD, but how it compares with other developmental conditions, where co-morbidities are present. Although structural and pragmatic language difficulties are correlated in clinical groups, this does not mean a single shared causal pathway nor that one symptom is a secondary effect of the other. Structural language problems are likely to arise from multiple, interacting factors (including protective strengths and environmental influences), some of which will overlap with effects due to other co-morbid developmental conditions. These may be perceptual, mnemonic, motor or attentional in nature or pertain to executive functioning.

We need to understand how multiple risks (some common across all developmental disorders) interact with protective factors and environmental influences to shape language differences. Such models predict some similarities across groups, but considerable individual qualitative differences determined by unique combinations of risk and protective factors. Given the unique autism profile, one would expect to see differences compared to individuals with DLD even if many risk factors are shared.

Further research is needed to specify the multiple causal mechanisms and their interaction. More subtle and specific analysis tools will be important in this endeavour. The field is already moving away from reliance on standard test scores, toward more naturalistic testing or techniques with no explicit response required, which may tease apart subtle differences and further elucidate reasons for individual variation. Equally there is a need for a cross-disorder comparison

looking at key traits and avoiding sampling bias via arbitrary cut-offs or exclusions. Due to the multiple bidirectional relationships inherent in the language learning process, it may be fruitful to examine non-linguistic skills in order to identify possible endophenotypes (e.g. Lee, Mueller & Tomblin, 2015), to try to avoid confounding effects of accumulated linguistic knowledge (which may have been acquired via atypical routes). For example, longitudinal measurement of phonological abilities in autism could further clarify their contribution to expressive language outcomes. A longitudinal investigation of declarative and implicit learning in young autistic children would provide much needed evidence towards the relevance of the procedural deficit hypothesis in ALI. In a similar vein, testing participants on execution of complex non-speech oral motor skills may reveal a different pattern of deficits compared to simply looking at sub-clinical articulation differences or basic oral motor movements (where compiled skills may be strong, particularly in older groups, despite an underlying weakness). Finally, increased scrutiny of individuals who have achieved late language learning (post age five) may be informative (e.g. Pickett, Pullara, O'Grady & Gordon, 2009).

Under a multifactorial model, changing core autism features (ie. social engagement and cognitive flexibility) may not always yield improvement on language structure and function – we may need to develop supplemental strategies to directly target language forms and functions more specifically. A full exploration of what these may be is beyond the scope of this review, however, there could be a stronger focus on motor production of speech and sign as a pivotal skill to be explicitly taught. Consideration could be given to how we can make language input more appealing or digestible to this group in order to support language development (e.g. modify perceptual characteristics of speech input, use of visual support/technology, modification of the type or amount of input). Finally, there is a need for robustly designed well-powered intervention studies investigating how to build language skills in those who remain minimally verbal by school age.

I have reviewed explanatory models for language impairment in autism, and highlighted a multitude of potential causal factors which may influence language development and result in a diverse range of language outcomes. Better

understanding these mechanisms is of particular importance to those with extremely limited expressive language and their families, as they have the most severe barriers to communication. I have therefore chosen to focus on those who remain minimally verbal in this thesis, and to focus on one specific factor: speech skills. In the oro-motor section of 1.3.1.2, I described the mounting evidence linking early speech skills with expressive language in autism, even when other putative predictors are controlled for. I outlined the studies investigating oral-motor difficulties and their influence on language development for some autistic children. In section 1.4.5 I considered how a lack of speech output may have cascading effects on overall language development.

The extant literature lacks unified definitions relating to preverbal speech skills, so I define speech skills in this thesis as the ability to make purposeful communicative vocalisations containing identifiable speech sounds or phonemes (consonants and vowels) within syllables (e.g. /d/ and /a/ in the syllable /da/ used in the context of request for more juice). Thus a preverbal child can have a range of speech skills, despite not producing any words. I conceptualise speech skills as a continuous variable, describing the repertoire of speech sounds a child produces. There is no agreed way to measure speech skills in this population. Speech skills are usually observed by transcribing a natural interaction between the adult and the child and analysing which individual sounds are used communicatively in syllabic vocalisations, or they are tested more formally by asking the child to imitate an adult model of a specific sound. I recognise that both approaches have strengths and weaknesses.

In summary, I posit that speech skills could be compromised in some (but not all) autistic children and this could exert an important influence on their expressive language trajectories. I explore the hypothesis that for some minimally verbal autistic children, an additional co-morbid speech-motor difficulty plays a significant role in their expressive language development.

My two main objectives are:

- 1) To extend work testing the hypothesis that early speech skills predict later expressive language in minimally verbal autistic children (Chapter 3)
- 2) To test the hypothesis that providing minimally verbal autistic children with a novel speech skills intervention could improve their speech skills and in turn positively influence their broader expressive language skills (Chapters 4 and 5)

Prior to developing the intervention required for (2) and in order to situate this study in the literature, I conducted a review of interventions targeting expressive language or speech skills in minimally verbal autistic children (Chapter 2).

2. Spoken language interventions in autism

2.1. Introduction

2.1.1. Rationale for the review

This thesis examines the hypothesis that speech skills are an important and potentially malleable influence on expressive language development for autistic children who remain minimally verbal. I investigate whether speech production skills can be improved in this population via daily parent-mediated practice using a motivating smartphone application (Chapters 4 and 5). This chapter aims to situate this novel intervention in the literature by exploring the following questions:

1. How has **spoken language** been targeted in previous autism interventions, particularly for **minimally verbal participants**?
2. Which other therapies have been evaluated for **minimally verbal autistic participants** and what outcome measures were targeted by these interventions?
3. Given the expected dearth of intervention studies which directly target spoken language for minimally verbal autistic individuals, how is **spoken language targeted in other relevant populations** (e.g. cerebral palsy, childhood apraxia of speech, Down syndrome, aphasia)?

Investigations of how to enhance expressive verbal language for autistic participants have emanated from somewhat disparate disciplines, including behavioural analysis, developmental science, special education and speech-language therapy. I searched the literature for systematic reviews and meta-analyses, and subsequently for additional relevant studies that were not represented in the reviews. In this chapter I aim to delineate the interventions that have been investigated thus far, as well as to evaluate the quality of this evidence. It is therefore useful to outline which features contribute to a high quality intervention study, why there so are few high quality intervention studies on

communication skills for minimally verbal autistic children, and how this can be addressed in future.

Randomized Control Trials (RCTs) have long been considered to provide 'gold standard' evidence of intervention success, so long as threats to internal validity are well controlled, e.g. by eliminating sources of bias in the recruitment, allocation and attrition of participants or the assessment and reporting of outcome measures. Performance bias is particularly difficult to overcome in language intervention studies, where parents and/or therapists are usually explicitly aware of group allocation (Brignell et al., 2018). A recent Cochrane review of communication interventions for minimally verbal autistic children (Brignell et al., 2018) concluded that although there have been many intervention studies in this field, there were only two studies worthy of inclusion, and thus the evidence quality is very low.

Another study design worth considering is the single case series, whereby each participant acts as their own control and is continuously assessed during a baseline and an intervention period. Important features include: systematic introduction of the intervention, strong inter-assessor agreement on outcome variables, sufficient data points (>5 in each phase) and opportunity to demonstrate repeated treatment effects (either within participant in an ABAB design, or between participants in a case series). What Works Clearing House suggests that a single case series demonstrating 3 or more treatment effects represents strong evidence, and if non-effects are also present this is downgraded to moderate evidence (Kratochwill et al., 2010). Recent innovations have included the introduction of randomisation to support more robust statistical analysis, and meta-analytic techniques to combine effects across multiple studies or multiple individuals within the same study (Rvachew & Matthews, 2017). Single case series have been used when the target population may be hard to recruit or heterogeneous, or if individual targets may vary by participant (i.e. one may struggle with the sound 'p', another may struggle with 'ee').

I will conclude that there is no consensus as to what sort of intervention would yield the best expressive language outcomes for minimally verbal autistic children, as

there have not been enough high quality studies. I will argue that focussing solely on RCTs may not be the best way forward for such a hard to reach and heterogeneous group, where single case series may be a viable and informative alternative (Kourea & Lo, 2016).

2.2. Spoken language in autism interventions

2.2.1. Background

Spoken language is a broad term which may include interventions seeking to increase the amount or quality of expressive language in minimally verbal or verbal participants, or enhance the intelligibility of verbal participants whose speech is unclear. A narrower search term would be expressive phonology or speech skills, which is closer to the outcome measure targeted by my intervention study, however I have included all spoken language interventions within my scope. In reviewing relevant studies, I was particularly interested in what age and stage of language development the participants were at, and how many of them were minimally verbal, as well as the proposed mechanism through which the intervention is theorised to work.

The autism intervention literature is characterised by a long history of studies of the impact of behavioural interventions. These interventions seek to build social competence, covering a wide range of cognitive, motor, play, imitative and language skills. They are often referred to as comprehensive or multi-faceted as they have multiple broad ranging goals, when compared to 'targeted' interventions focussed on individual aspects of language (Hampton & Kaiser, 2016). Multi-faceted interventions may be delivered intensively by therapists, or in a parent-mediated 'low-intensity' model. They are often administered from a relatively young age, hence the name Early Intensive Behavioural Intervention. They tend to be based on learning theory from the science of Applied Behaviour Analysis (ABA), but under this umbrella there is substantial variety in methods, which range on a continuum from highly structured and contrived, to naturalistic and play-based. For example, Discrete Trial Training, which was first described by Wolf et al. (1964), is considered highly structured or 'contrived'. It comprises adult-directed, massed trial training, and is based on a fixed curriculum. Participants are motivated via reward systems, external to the task in hand. Naturalistic Applied Behaviour Analysis approaches, in contrast, use naturally occurring reinforcement in a child-led, individualised approach, variously described as Natural Environment Teaching

(Sundberg & Partington, 1998), Pivotal Response Training (PRT, Koegel, Koegel, Harrower & Carter, 1999) or Enhanced Milieu Teaching (EMT, Hancock & Kaiser, 2002). Another example of comprehensive intervention is Social Communication, Emotional Regulation Transactional Support (SCERTS, Prizant et al., 2006). An Applied Behaviour Analysis variant with an emerging evidence base is the Early Start Denver Model (see Baril & Humphreys, 2017 for a review), which seeks to overlay a developmental-pragmatics understanding to existing behavioural methods, i.e. children struggle with language because they lack pre-linguistic skills such as joint attention, social engagement, imitative ability, symbolic understanding, and these should be targeted in Early Intensive Behavioural Intervention.

It is outside the scope of this review to evaluate the different Early Intensive Behavioural Intervention programmes and their features. However, in order to understand which interventions are comparable to the intervention in this thesis, in specifically and directly targeting expressive language alone, it must be borne in mind that the majority of interventions available to autistic preschoolers and subject to empirical testing so far, are multi-faceted. These interventions do seek to improve multiple aspects of language but this is not usually their sole objective, which stands to reason since autistic children have pervasive developmental difficulties across many domains. Consequently, outcome measures in trials will include autism symptom severity, adaptive and social variables as well as language outcomes (Green et al., 2010, Carter et al., 2011, Kasari et al., 2014b). Even among behavioural interventions which identify as targeted towards language only, they may target multiple aspects of language, such as pre-linguistic communication skills, receptive skills, pragmatic skills, semantics and use of alternative communication systems, in contrast to the intervention outlined in Chapters 4 and 5. This makes cross-study comparisons difficult as each intervention targets a different constellation of outcome measures.

2.2.2. Systematic Reviews and Meta-analyses

The most relevant systematic review is Brignell et al.'s (2018) Cochrane review of communication interventions in minimally verbal autistic children, which identified only two intervention studies of sufficient methodological rigour. These were a school-based trial of Picture Exchange Communication System (PECs) in children aged 4-11 (n=84, Howlin, Gordon, Pasco, Wade & Charman, 2007) and a parent-mediated focused play therapy for 32-82 month-olds (n=70, Siller, Hutman & Sigman, 2013). The review only included RCTs in which there was a control group receiving treatment as usual or no treatment; studies in which two interventions were explicitly compared were excluded from review. This intended to control for the impact of spontaneous maturation (or regression to the mean) on treatment effects, which further limited the pool of eligible studies. The authors also downgraded the quality of evidence due to small sample size, wide confidence intervals and only partial blinding. Strikingly, neither Howlin et al. (2007) or Siller et al. (2013) reported a significant treatment effect on expressive language. Howlin et al. (2007) reported that PECs training resulted in greater interactions and PECs use, but had no impact on spoken language, and gains in PECs use were not maintained at follow up. Siller et al. (2013) reported a treatment effect of improved parent behaviour, but there was only a conditional effect on expressive language outcomes (see below).

Hampton & Kaiser (2016) carried out a systematic review of intervention effects on spoken-language outcomes for autistic children, and made the distinction between targeted and comprehensive/multi-faceted interventions. They limited the search to early (<8 years) and behavioural interventions, and only included high quality studies up to 2014 (i.e. RCT or quasi-experimental designs). These criteria mean that it is likely that some targeted interventions were omitted from the search, as these are usually smaller in scale and evaluated via case series. The authors concluded that there was some evidence that early behavioural intervention modestly improved spoken language outcomes ($g=.26$), especially when parents are trained in therapy delivery alongside practitioners ($g=.42$), however it was unclear whether these small effects translated to a meaningful functional

improvement in skills. The 26 studies included 1,738 participants with an average age of 3.3 years, so one can infer that participants were thus mainly at the early stages of language learning. However, this cannot be assumed: Hardan et al. (2015) evaluated a parent Pivotal Response Training intervention via an RCT (control condition was psycho-education about autism with no specific Pivotal Response Training or language component) for 53 participants aged 2-6 years. Following a 12-week course, Pivotal Response Training group children made significant improvements based on a natural language sample pre- and post-intervention. However, participants would not qualify as minimally verbal - a prerequisite to study entry was "child being able to vocalise with intent when prompted by a clinician during the screening visit", and their mean expressive vocabulary was over 130 words at intake.

Kane, Connell and Pellicchia (2010) undertook a review of the behavioural literature to determine spoken language effects in autism intervention, however they included only single case series in order to synthesise the data quantitatively with a Percentage of Non-overlapping Data (PND) metric. The 22 studies included were all of a targeted nature, with the dependent variable being one specific language function, for example naming, requesting, participating in conversational exchanges, use of prepositions or adjectives (n=65, ages 2-12 with mean age of 6.2 years). The meta-analysis resulted in a mean Percentage of Non-overlapping Data of 65% and 83% for contrived and naturalistic approaches respectively, meaning that the majority of data points in the intervention phase were higher than the highest score in baseline. However, inspection of the studies concerned reveals that only one (Esch, Carr & Michael, 2005), related to strengthening the vocal repertoire of children with minimal speech, whereas the other studies all involved children with pre-intervention verbal skills. Esch and colleagues (2005) used a stimulus-stimulus pairing technique (a variant of Discrete Trial Training, whereby the target vocal sound is presented simultaneously with a desired object or activity, with no contingency placed on the child to vocalise). Investigations of this method to date have had mixed results (see below).

Mulhern et al. (2017) performed a relevant systematic review of procedures for the induction of speech among persons with developmental disabilities. Despite having a wider scope than just autistic participants, 72.5% of the 1,213 participants in 50 out of the 78 studies had a primary diagnosis of autism. The authors evaluated mainly case series, using the Percentage of Non-overlapping Data method and concluded that established behavioural techniques (e.g. differential reinforcement, prompt-delay) were supported by the evidence. Of the non-behavioural or mixed studies reviewed, several contained elements of music or rhythm (e.g. Sandiford, Mainess & Daher, 2013, Lim & Draper, 2011), others included computer assisted practice with visual feedback (Bernard-Opitz, Sriram & Sapuan, 1999) but overall the evidence base for non-behavioural methods was judged to be insufficient.

Lane, Lieberman-Betz and Gast (2016) examined naturalistic language interventions targeting spontaneous spoken language in autism and their efficacy. A total of 24 studies included 45 participants (age range: 2-9 years) but only limited information was available regarding language stage of the participants. The What Works Clearinghouse guidelines (Kratochwill et al., 2010) were used to evaluate methodological rigour in each study design. It was deemed to be adequate for half the studies, and of those, interventions were determined to have increased spoken language, but no overall effect size was generated due to the varied design of the studies included.

Finally, Goldstein (2002) reviewed the evidence for interventions seeking to increase communicative repertoires (which includes but is not limited to spoken language) in autistic children. This review was across all levels of language competency, e.g. some conversational interventions were included. Amongst the conclusions were that the total communication approach (combining sign language and vocalisations) was more successful than prompting speech alone, in particular when verbal imitation is weak, a sentiment echoed in the Alternative and Augmentive Communication (AAC) literature. Goldstein also critiqued the current state of research, suggesting that interventions are not well described, designs not robust, there are too few studies and no replications.

Summary: A rigorous comparison of high quality studies (Brignell et al., 2018) revealed only two eligible studies, neither of which demonstrated a clear improvement in spoken language outcomes. Both studies did however show improvement in the environmental supports thought to be the agents of change (use of PECs in classroom and use of parent responsive behaviours at home). Additional narrative reviews reveal a wide range of behavioural and non-behavioural approaches to improving language forms and functions in autism. Studies are diverse in the language ability profile of participants, the specific outcome measures used and the analysis methods employed, this lack of uniformity further hinders synthesis of the evidence base.

2.2.3. Individual studies

2.2.3.1. Articulation training for autistic participants with some verbal skills

Koegel, Camarata, Koegel, Ben-Tall and Smith (1998) compared structured (Discrete Trial Training based) articulation training with a naturalistic approach for five school aged children in a counterbalanced reversal design. Children in this study had some speech but had reduced intelligibility. They were assigned previously unmastered target sounds such as /j/ and /th/. Therapy was received for two x 45 minutes per week over a period of 54-95 sessions. Both methods increased speech accuracy during trials but generalization to conversational language samples was only observed for phonemes taught using the naturalistic procedures.

2.2.3.2. Interventions for minimally verbal autistic participants

The intervention studies reviewed in this section are the most relevant to the current study, since they all aim to improve an aspect of spoken expressive language for minimally verbal autistic participants. I have therefore summarised the studies in Table 1.

Evaluating Specific Behavioural techniques

Koegel, O'Dell and Dunlap (1988) described a case series where four non-verbal school-aged autistic children underwent two types of therapy aimed at enhancing spoken output, in a repeated reversal design. In one intervention, all speech attempts were rewarded and in the other therapists sought to shape speech output by only rewarding the 'target' production (in an adaptive design based on previous successful productions). Speech productions were rated according to analysis of their distinctive features, to give a measure of accuracy. Not only production accuracy but also affect during sessions were higher in the 'reinforce all attempts' condition, highlighting the importance of motivation in speech acquisition.

Lepper, Pettursdottir and Esch (2013) compared two Discrete Trial Training approaches to increasing vocalisations in three autistic boys aged 2-4 years and with no echoic abilities. An alternating design was employed for approximately 50 sessions. They found that both Stimulus-Stimulus Pairing (SSP, described above) and Operant Discrimination Training (ODT) were effective in increasing the rate of target vocalisations compared to the control condition, but ODT was preferred by participants. In ODT, the reinforcement was conditioned on a set behaviour by the child (arm raising, which was prompted until mastered). A further study by Lepper and Petursdottir (2017) compared two specific variants of Stimulus-Stimulus Pairing, Response Independent Pairing (RIP) and Response Contingent Pairing (RCP) on three boys, aged 4-6 years. RCP was shown to be more reliable in evoking target vocalisations.

Esch, Carr and Grow (2009) described a case series of three boys aged 2;0 to 5;7, using Stimulus-Stimulus Pairing procedures over 40-60 sessions. Vocalisation of the target syllable or word did increase, non-target words did not.

A specific weakness of this type of training study is that it is unclear whether improvements during the sessions would result in generalisation of the sounds being learnt to use in other contexts or to combination with other sounds into words. This design could be improved by including a naturalistic generalisation probe, for

example opportunities to use newly acquired speech sounds in different environments, during different activities and/or with different people. Given the large amount of therapy time required to evidence the acquisition of one or two target sounds in these studies, it is worth considering whether prioritisation of speech sound learning in minimally verbal participants is appropriate. On the other hand, little information is given in these studies on the other communication targets and functional skills are being worked on in conjunction with speech sounds.

Evaluating Naturalistic vs. Structured Applied Behaviour Analysis

Paul, Campbell, Gilbert and Tsiouri (2013) directly compared two interventions, in order to determine whether Rapid Motor Imitation Antecedent (which is a variant of Discrete Trial Training) or Enhanced Milieu Training (EMT, a naturalistic approach), was more effective in inducing spoken language in minimally verbal autistic preschoolers (n=22, ages 2-6). The main outcome measures were reported expressive vocabulary, observed different words spoken and total number of word tokens. The rationale for Rapid Motor Imitation Antecedent is its reported efficacy in teaching vocal imitation, considered to be a necessary precursor to benefit from Discrete Trial Training. Both intervention arms were delivered by speech therapists over 12 weeks in 36 45-minute sessions, and also had a parent responsiveness training component. Both groups improved on spoken language measures, and joint attention moderated the response to both treatments. Children with better receptive skills fared better with Enhanced Milieu Training and those with lower receptive skills responded better to Discrete Trial Training. About 50% of children across both interventions succeeded in transitioning to the second stage of language learning (putting words together), as per Tager-Flusberg et al. (2009). This study placed emphasis on the importance of pre-intervention motor imitation skills as a precursor to vocal imitation. A subset of children who did not have enough motor imitation skills to participate in Rapid Motor Imitation Antecedent received motor skills training beforehand, and those who succeeded on this were randomly assigned to both interventions. Those who did not were assigned to Enhanced Milieu Training, which renders the study a quasi-experimental design.

Evaluating parent responsiveness training

Siller et al. (2013) used an RCT to evaluate the efficacy of a focussed playtime intervention designed to boost parent responsiveness (12 weekly sessions of 90 minutes), compared with treatment as usual. Participants were 70 minimally verbal 2-6 year olds. A treatment effect was observed on parent behaviour, but only a conditional effect on expressive language outcomes: those with baseline language scores <12 months benefitted most from the intervention (n=24). This highlights the need to consider not only what works best but also for whom.

Evaluating the role of AAC in evoking spoken language

PECS

The Picture Exchange Communication System (PECS, Bondy & Frost 2001) is an augmentive communication system, frequently implemented with minimally verbal autistic preschoolers. Some studies have observed an increase in spoken language following PECS interventions which is maintained and generalised post-intervention (e.g. Charlop-Christy, Carpenter, Le, LeBlano & Kellet, 2002). In contrast, Howlin, Gordon, Pasco, Wade and Charman (2007) conducted an RCT of teacher PECs training and concluded that despite significantly greater initiations and PECs use in the intervention group, no increase in spoken language was observed. Furthermore, gains were not maintained once the intervention ceased.

Schreibman and Stahmer (2014) carried out an RCT focusing on verbal communication using minimally verbal participants (n=39, age<4 years), comparing two alternative treatments: Pivotal Response Training and PECs. They found that the two groups' spoken language did not significantly differ post-intervention. Children in both groups increased their spoken vocabularies by an average of over 80 words, although variance was high, indicating heterogeneous responses. Seventy-eight per cent of children exited the programme (averaging 247 hours over 23 weeks) with more than 10 functional words. Parent-report highlighted that PECs was more difficult to implement than Pivotal Response Training. The authors attribute this to the need to create and utilise specific

additional materials for PECs (laminated pictures) whilst Pivotal Response Training employed items already available in the home setting.

Yoder and Stone (2006) carried out a lower intensity version of this experiment, also targeting spoken language in those who had <20 different words pre-treatment. Participants (n=36) were aged 18-60 months and randomly received either Responsive Education and Pre-linguistic Milieu Teaching (RPMT, a naturalistic behavioural intervention) or PECs for 24 hours over six months, with tests before, immediately after and six months after the intervention. The outcome measure, non-imitative spoken communication acts, was deemed to improve more in the PECS group. Children rating highly on an initial object exploration measure experienced heightened success in the PECs group, this is theorised to be because the PECs intervention did not teach play skills, so interest in a variety of objects is a useful pre-requisite, whereas on the RPMT intervention play would be naturally targeted.

Speech Generating Devices

Gevarter et al. (2016) actively targeted vocal word approximations whilst using a Speech Generating Device (SGD) combined with behavioural methods (e.g. reinforcement delay, vocal prompt). In a multiple baseline case series, three out of four minimally verbal children (ages: 4;0 to 7;9) with prior SGD experience, were taught to vocalise to request a preferred item alongside activation of an SGD. Importantly, these vocal approximations were generalised to situations where the SGD was not present. The authors discuss parallels and advantages over PECS (e.g. consistent vocal model automatically provided, child can use vocalisations to request more effectively with a wide range of communicative partners by supplementing them with SGD). Amongst inclusion criteria, however, was a score greater than 0 on an echoic skills battery, and targeted word approximations contained at least one sound that had been judged to be within the child's repertoire, indicating that results need may not be generalizable to those with more impaired speech imitation skills.

Mirenda (2008) provides a useful discussion of the merits of SGD use in what is variously known as Natural Aided Language intervention (NAL, Cafiero, 2001); Aided Language Modelling (ALM, Drager et al., 2006); and System for Augmenting Language (SAL, Ronski et al. 2010). In these interventions, adults frequently and naturalistically combine vocal models, pointing and SGD activation when commenting on items in the child's environment. Mirenda (2008) supports the rationale for input-only SGD interventions, by arguing that typical children are provided daily with thousands of exemplars of verbal language prior to speaking their first word, whereas the number of exemplars of SGD use experienced by the average user prior to being expected to use this system is much smaller. Ronski et al. (2010) compared this approach to an analogue spoken language intervention as well as a more 'traditional' output-only SGD intervention in an RCT (n=62, children with developmental delays and <10 words, aged 2-3⁴). Both the contingent and non-contingent SGD approaches resulted in greater expressive vocabularies and spoken output, suggesting that SGD use may help rather than hinder vocal development.

Multimodal speech training

Wan, Chenausky and colleagues (Wan et al., 2011; Chenausky, Norton, Tager-Flusberg & Schlaug, 2016, 2018) have approached the question of how to induce speech in minimally verbal autistic children from a different angle, trialling a novel intervention called Auditory-Motor Mapping Training (AMMT) in a proof of concept study, a matched group study and an ongoing RCT. The two main components of Auditory-Motor Mapping Training are intonation (singing) and rhythmic tapping and the aim is to enable sound-motor mapping, in a modified version of Melodic Intonation Therapy, used in aphasia. The proposed mechanism is more directly linked to speech production abilities and thus more comparable with the intervention study in this thesis (Chapters 4 and 5).

⁴ not all participants were autistic

Auditory-Motor Mapping Training was compared with traditional Speech Repetition Therapy (SRT) in a matched group study (n=14), and children were tested on trained and untrained bisyllabic words at various time points (Chenausky et al., 2016). Both therapies were implemented by speech therapists, in relatively intensive fashion (45 minutes per day for five days a week for six weeks). Results indicated that Auditory-Motor Mapping Training participants achieved a higher rate of change in % syllables approximated (29% improvement for Auditory-Motor Mapping Training vs. 3.6% for Speech Repetition Therapy). A key limitation of this study when generalising to all minimally verbal autistic children is the high number of screened participants (30) who were excluded from the study because "...they could not participate in table-top activities for at least 15 minutes, were unable to imitate any speech sounds, [or] were completely non-vocal."

Chenausky et al. (2018) describes regression analysis performed on a larger pool of participants who had all received 25 hours of speech-targeting therapy (either Auditory-Motor Mapping Training or Speech Repetition Therapy), in order to establish predictors of change scores. The participants were all minimally verbal (<20 intelligible words and no productive syntax) and ages ranged from 3;5 to 10;8. Phonetic inventory predicted over 30% of the variance in change score on a '% syllables approximately correct' outcome measure in the full sample of 38 participants, meaning that those participants with some initial speech production skills benefited more from the intervention. When phonemic inventory was combined with ADOS score (a measure of autism symptom severity) in a smaller subset of these children with complete cases, they predicted 73% of variance. Importantly, phonetic inventory in this study was measured by a repetition task rather than taken from a natural language sample as in previous works (Yoder et al., 2015) and the longitudinal study described in Chapter 3.

Finally, Chenausky, Norton and Schlaug (2017) explored treatment effects (when contrasted with Speech Repetition Therapy) of Auditory-Motor Mapping Training in a matched pair of participants with greater verbal skills (at 'word combination' stage). They noted a greater effect size in outcome measure (% syllables

approximately correct), highlighting that it might be an efficient intervention for those with some verbal skills as well.

Sandiford et al. (2013) compared Melodic Based Communication Therapy (MBCT) to traditional speech-language therapy in a small pilot RCT (n=12, ages 5-7). Participants were minimally verbal and underwent four 45-minute sessions per week for five weeks, targeting imitation of 25 words. Melodic Based Communication Therapy differs from Auditory-Motor Mapping Training in that each word has its own melody, melodies are played from a pre-recorded device by the therapist and individualised reinforcement is used for each child. Results showed Melodic Based Communication Therapy was at least as effective as traditional therapy in increasing imitative and non-imitative speech productions of the 25 target stimuli.

Another form of multimodal speech training is described in an exploratory study by Vernay, Harma, Marrone and Roussey (2017). Four participants, of whom two were minimally verbal and two were reported to use no words, took part in 10 sessions of self-paced segmentation of common multisyllabic words on a touchscreen. The rationale was that visually emphasising syllables in written words would serve as an enhanced prompt to produce them. This pilot study showed no overall learning effects, and although three out of four participants seemed to make more correct productions following the self-paced segmentation phase, findings are hard to interpret due to confounding order effects and lack of baseline, follow-up or generalisation testing.

Sensory-motor speech training

Rogers et al. (2006) used a single case design to compare a multi-faceted intervention (Early Start Denver Model), with a tactile-motor, non-autism specific, speech focussed intervention (Prompts for Restructuring Oral Muscular Phonetic Targets, PROMPT). Ten participants aged 20–65 months with fewer than five words at intake were allocated to matched pairs and randomly assigned to interventions. The study used an A-B-A design with a three-month follow-up. Participants in both interventions received 12 one-hour weekly sessions of therapy and daily 1-hour home follow-up delivered by parents. Both intervention groups demonstrated progress on the outcome measure (total words/hour). Eight out of 10 children were able to use more than five novel words by the end of the intervention. Nine out of the 10 showed an improvement in reported vocabulary score, some remarkably so (from 20 to 190+ words spoken). Notably, children in both interventions were less verbal in the generalisation probe than intervention probes, highlighting the importance of ensuring generalisation. Children ranged in their past experience of intervention and in terms of pre-requisite skills, autism symptom severity and cognitive skills. A pair of children who were older with limited progress on extensive past therapy made good gains, as did a pair who were younger and had not previously received much therapy. It is difficult to infer anything given the sample size and lack of no treatment control, but it is an encouraging pilot study.

Sweeney and Lebersfeld (2016) undertook a single case series employing Prompts for Restructuring Oral Muscular Phonetic Targets (PROMPT) alongside strategies based on motor learning theory (e.g. block practice drills, repetition). Participants were aged 4-25 (n=10), had a diagnosis of autism alongside symptoms of childhood apraxia of speech, and spoke one word or less. After 30 sessions over a period of 6-7 weeks, nine participants had learnt to say at least one word (mean=2.8). Further studies are necessary to assess replicability of these findings, as well as functional impact for the participants involved.

Spoken Language Intervention pre-diagnosis?

Another small case series looked at a parent-mediated Pivotal Response Training intervention on three preverbal toddlers with social communication delays and signs of autism, but who had no formal diagnosis due to their age, using a non-concurrent multiple baseline design (Bradshaw, Koegel & Koegel, 2017). Intervention was 12 1-hour weekly parent-coaching sessions. The dependent variable was child verbal expressive language from weekly 10-minute parent-child interaction, operationalised as syllabic word approximations that were classified as either 'initiated', 'responses' or 'prompted'. The intervention was successful in modifying parent behaviour and this was maintained at follow-up. All three children demonstrated improvements in frequency of word approximations. One difficulty with trialling interventions for very young children is that it has not yet been established that they have a persistent problem with expressive language, so they may be a different population to those who are minimally verbal at age five. Larger scale and longitudinal studies would be needed to establish if such early intervention had a lasting effect on spoken language, and a control group would be needed as many late-talkers resolve their expressive delays spontaneously.

In summary, there have been a range of innovative approaches to induce speech in minimally verbal autistic children, however high quality evidence is still lacking. Case series have demonstrated gains in accuracy of specifically targeted sounds or number of words produced, however this is sometimes the result of hours of trials and evidence of generalisation to other settings or communicative partners is limited. Larger RCTs have shown some promising results but treatment response is highly variable. Further research is needed to determine which pre-treatment characteristics are associated with intervention outcomes.

Table 1: Summary of individual intervention studies aiming to improve spoken expressive language in minimally verbal autistic children

Study	Intervention approach	Participants	Outcome measure(s)	Study design	Effect direction
Bradshaw et al., 2017	Pivotal Response Training, parent-mediated	N=3 Age <2yrs**	Words spoken	Case Series	Increase in outcome variable during intervention
Chenausky et al., 2016	Auditory Motor Mapping Training vs traditional speech therapy, clinician led	N=14 Age 5-9yrs	Accuracy on trained and untrained words	Matched group study	Outcome increase was higher in Auditory Motor Mapping Training condition
Esch et al., 2005	Stimulus Stimulus Pairing (a form of Discrete Trial Training) clinician led	N=3 Age 6-8yrs	Accurate production of target sounds/syllables	Case series	Limited improvement
Esch et al., 2009	Stimulus Stimulus Pairing combined with programmed reinforcement, clinician led	N=3 Age 2-6yrs	Accurate production of target and non-target sounds/syllables	Case series, alternating design	Positive effect on target sounds/syllables and no effect on non-target sounds/syllables
Gevarter et al., 2016	SGD combined with behavioural methods, clinician led	N=4 Age 4-8yrs	Vocal word approximations	Case series	Three out of four children learnt to request preferred item with vocal approximation and generalised this skill to situations without device
Howlin et al., 2007	PECs training in school setting vs. no treatment	N=84 Age 4-6yrs	Communicative initiations PECs use Speech	RCT	Communicative initiations and PECs use increased but changes not maintained at follow up. Speech: no effect
Koegel et al., 1988	Behavioural vocal imitation training, comparison of 'reinforce all attempts' vs. 'reinforce target production' clinician led	N=4 Age=3-11yrs	Accurate production of target sounds/syllables Affect	Case series, repeated alternating phase design	Positive effect for accuracy and affect in reinforce all attempts condition 1 child made no progress
Lepper et al., 2013	Operant Discrimination Training vs. Stimulus Stimulus Pairing	N=3 Age 2-4yrs	Number of times target sounds/syllables produced	Case series, alternating design	Positive effect on outcome measure in both conditions, Operant

	(both forms of Discrete Trial Training) clinician led				Discrimination Training was preferred
Lepper et al., 2017	Response Independent Pairing vs. Response Contingent Pairing (both forms of Stimulus Stimulus Pairing) clinician led	N=3 Age 4-6yrs	Number of times target sounds/syllables produced	Case series, alternating design	Response Contingent Pairing was more effective
Paul et al., 2013	Rapid Motor Imitation Antecedent (Discrete Trial Training variant) vs. Enhanced Milieu Training (naturalistic behavioural intervention), clinician led	N=22 Age 2-6yrs	Expressive vocabulary (parent report) Number of spoken words (observation) Language age (Mullen)	Quasi-experimental (group study but some participants not fully randomized)	Positive effect on outcome measures in both groups, mediated by initial joint attention level. No difference between conditions. 50% transitioned from minimally verbal status
Rogers et al., 2006	Early Start Denver Model vs PROMPT, clinician led with parent training	N=10 Age 1-5yrs	Words spoken	Matched group study	Both groups made comparable progress Eight out of 10 children used 5 or more words by the end of the study
Romski et al., 2010	SGD output only vs. SGD input vs. traditional speech therapy, clinician led with parent training	N=62* Age 2-3yrs	Spontaneous augmented and spoken target words	RCT	Positive effect of both SGD approaches vs. traditional speech therapy
Schreibman & Stahmer, 2014	Pivotal Response Training vs. PECS, clinician led with parent training	N=39 Age<4yrs	Expressive vocabulary (parent report) Expressive language age (Mullen) Adaptive communication	RCT	Positive effect on outcome measures in both groups, no difference between conditions. 78% exited intervention with >10 words
Siller et al., 2013	Focused Playtime parent coaching in home setting vs. no treatment	N=70 Age 2-6yrs	Parent responsiveness Expressive language age equivalent (Mullen)	RCT	Parent responsiveness increased. Expressive language age: no main effect. Conditional positive effect for those with lowest

					baseline (language age<12 months)
Sweeney & Lebersfeld, 2016	PROMPT combined with motor learning strategies, clinician led	N=10 Age 4-25yrs	Words spoken	Case series	9/10 made progress of at least 1 word
Yoder & Stone, 2006	Responsive Education and Pre-linguistic Milieu Teaching vs. PECS, clinician led with parent training	N=36 Age 1-5yrs	Non-imitative spoken communication acts	RCT	More positive effect on outcome measure in PECS condition
Sandiford et al., 2013	Melodic Based Communication Therapy vs. traditional speech therapy, clinician led	N=12 Age 5-7yrs	Imitative and non-imitative production of target stimuli	RCT	Both groups made similar levels of progress
Vernay et al., 2017	Self-paced segmentation of visual stimuli using touchscreen, clinician led	N=4 Age 8-12yrs	Imitative production of target stimuli	Case Series	No learning effect demonstrated

Note: Mullen=Mullen Scale of Early Development (Mullen, 1995); PECS= Picture Exchange Communication System; PROMPT=Prompts for Restructuring Oral Muscular Phonetic Targets; RCT=Randomized Controlled Trial; SGD=Speech Generating Device; * not all participants were autistic; ** participants only had working diagnosis of autism due to their age

2.3. Other interventions for minimally verbal participants not focussed on spoken language

A recurring theme is the emphasis on broader communication goals, targeting expressive language via other means, such as PECs, sign, or Speech Generating Devices. Much research on this topic has utilised single case designs (e.g. Olive, Lang & Davis, 2008; see Ganz et al., 2012 for a review). However, there have been recent large RCTs with adaptive components embedding AAC use in a multi-faceted behavioural programme (Almirall et al., 2016, Kasari et al., 2014a). Kasari et al. (2014a) found that when a developmental/behavioural intervention focussed on social engagement was combined with Speech Generating Device training, overall spontaneous communicative gestures increased relative to the behavioural intervention alone (effect size .62), and as a by-product, spoken language also increased significantly. Distefano et al., (2016) discusses how interaction quality in this study (number and length of child-therapist interactions) may drive the increase in spoken language experienced by those in the intervention group receiving behaviour and Speech Generating Device input.

In a similar vein, pre-linguistic skills such as joint attention and play are frequently targeted in this population. Goods, Ishijima, Chang & Kasari (2013) trained these skills in an RCT for 3-5 year-old minimally verbal autistic children. For the intervention group, one hour per week of their 30-hour Applied Behaviour Analysis programme was replaced with Joint Attention Symbolic Play Engagement and Regulation (JASPER) for 12 weeks, following a 12-week baseline period. The control group continued with Applied Behaviour Analysis as usual. Significant differences were observed in post-treatment play diversity, proportion of time spent engaged (in any activity) and initiation of spontaneous gestures. Imitation is another pre-linguistic skill that interventions seek to improve. Ingersoll and Schreibman (2006) evaluated the effects of parent delivered Reciprocal Imitation Training on five young autistic children (ages 31-42 months), two of whom were minimally verbal. These two participants improved versus baseline at object imitation, which was maintained at one-month follow-up. Language measures were also taken during treatment, post-treatment and at follow-up. One child made no

language progress, but increased her rate of imitative language, albeit with no generalisation. The other child improved his imitative language with limited generalisation, and had more spontaneous language post-treatment. Compared to the language gains made by the three verbal children in the study, progress for minimally verbal children was more attenuated.

In a rare attempt to specifically target listening comprehension in minimally verbal autistic students, Mucchetti (2013) described case studies of four children (6-8 years old, <20 words), who took part in an adapted shared book reading intervention, where enhanced multimodal shared book reading techniques (e.g. visual supports, three-dimensional objects, simplified text) were contrasted with usual shared book reading methods in an alternating single case design. Significantly higher student engagement and story comprehension were observed relative to the usual shared book reading methods used at baseline (Percentage of Non-overlapping Data $\geq 80\%$).

Another common intervention target is the reduction of challenging behaviours, which frequently affect minimally verbal autistic participants (Tager-Flusberg et al., 2016). Carr and Durand (1985) theorised that challenging behaviours (e.g. biting, bolting, self-harm) could serve a communicative function and therefore teaching specific communication skills to replace them could lead to a reduction in challenging behaviours. Functional Communication Therapy is considered to have a solid evidence base, with over 90 peer reviewed studies suggesting that it is a valuable tool to decrease challenging behaviours in autism (Gerow et al., 2018; Tiger, Hanley & Bruzek, 2008; Wong et al., 2014). The teaching of functional communication employs many of the methods discussed above but intervention outcome measures concern the impact on the rate or type of challenging behaviours.

2.4. Speech interventions for other minimally verbal populations

In light of the limited studies targeting speech in intervention studies in autism, it is potentially useful to consider the intervention literature aimed at increasing the variability, frequency or quality of speech sounds produced for other clinical populations with similarly severe spoken language limitations (Lindblad, 2012). Are there any conclusions to draw about whether these could be adapted or extended to minimally verbal autistic preschoolers? Several specific therapies discussed in this review are also provided to other populations with limited spoken language. Prompts for Restructuring Oral Muscular Phonetic Targets (PROMPT) has been evaluated for participants with cerebral palsy (Ward, Leitão & Strauss, 2014) and aphasia (Richardson, 2012). Similarly, AAC has been used for a range of communication targets with participants with Down syndrome (e.g. Dada & Alant, 2009, Lanter, Russell, Kuriakose & Blevins, 2016). These populations are likely to have a different profile of social skills underpinning their language learning to autistic participants, as well as a variable range of cognitive abilities. Although a full review is beyond the scope of this thesis, reported trials of PROMPT and AAC for other groups were positive, however studies have been small, with poor experimental control and heterogeneous participants.

Another population of interest are those with speech production impairments in the absence of any other pervasive developmental difficulties (usually known as childhood apraxia or developmental verbal dyspraxia). These children are quite different from minimally verbal autistic children in that they tend to have average range cognitive, social, and attentional skills. Interventions described below are focused on increasing intelligibility in already verbal children rather than evoking first words or sounds. Numerous reviews have synthesised the existing evidence base (Baker & McLeod, 2011; Morgan & Vogel, 2008, Murray, McCabe & Ballard, 2014, Sugden, Baker, Munro & Williams, 2016). Morgan, Murray and Liégeois (2018) completed a recent Cochrane Review in this field, and concluded that only one well-controlled study was identified (Murray, McCabe & Ballard, 2012, 2015). This RCT (n=26, ages 4-12) contrasted Rapid Syllable Transition (ReST)

Treatment with the Nuffield Dyspraxia, Third Edition (ND3). Both interventions demonstrated a significant treatment effect (the outcome measure was accurate productions on a 292-item probe derived from ND3 assessment protocol) and generalisation, but gains were maintained more effectively by the ReST intervention group four months later.

Additionally, there have been several attempts to incorporate technology into interventions for speech sound disorders. A novel short-term intensive intervention for childhood apraxia of speech using ultrasound biofeedback is described in a case series with mixed success, both in terms of retention and generalisation of the problematic speech sounds (n=3, ages 10-14 years) (Preston, Leece & Maas, 2016). McLeod et al. (2017) reported no significant difference in outcomes following a recent RCT comparing treatment as usual with a computer-assisted phonology practice programme, for children with identified speech sound disorders (n=123, ages 4-5 years, outcome measure was % of consonants accurately produced).

Finally, Martin et al. (2015) reported results from a two-year intensive, school-based, structured, multimodal, phonetic intervention designed to improve articulation in 12 participants with developmental verbal dyspraxia. Participants did not have a co-morbid autism diagnosis but several had ADHD. Standard (age-adjusted scores) on an articulation test significantly increased, as did secondary outcome measures such as Mean Length of Utterance and reported resilience, however with the absence of a control group it is difficult to draw any firm conclusions.

2.5. Conclusion

In conclusion, so far there is no consensus on how best to induce spoken language in minimally verbal autistic children. No systematic review to date has covered the exact remit of this question, although several related reviews are informative. Understandably, multi-faceted interventions, many based on the principles of Applied Behaviour Analysis, have dominated the field to date and evidence is robust for their efficacy in improving measures relating to social engagement (e.g. time spent engaged, number of child initiations, parent responsive behaviours), with more variable impacts on language, play skills, adaptive functioning or autism symptom severity (Green et al., 2010, Carter et al., 2011, Kasari et al., 2014b).

Language skills, in their broadest sense, are central to these types of intervention but not the only skill that they target, and language skills are not always designated as a primary outcome measure. There has rightly been an evidence-driven shift towards naturalistic, and developmentally appropriate methods and targets in such programmes, along with the importance of skill generalisation. Although some young preverbal children make excellent spoken language progress on such programmes, there is an increased focus on understanding which child pre-treatment characteristics mediate intervention response and define slow or poor responders. Paul et al. (2013) analysed pre-treatment moderators of intervention success and identified joint attention as a significant predictor of positive outcome. Chenausky et al. (2018) compared pre-treatment factors with success on a speech-focussed intervention and found that only autism symptom severity and initial speech skills were predictive. Further studies are needed to explain these contrasting results, but effects are likely to be impacted by intervention format, goals and sample characteristics.

Behavioural approaches form the backbone of most autism language interventions, however, a range of innovative strategies based on motor and sensory considerations have been trialled to more directly target the goal of inducing speech in minimally verbal autistic children. To date conclusions would be premature due to small, heterogeneous samples which are poorly suited to

RCTs. Importantly, many studies aiming to improve spoken language in those who are minimally verbal still exclude those with the greatest impairments from taking part, thus making it difficult to generalise their conclusions. A focus on building basic echoic skills in such severely impaired children is needed in order to enable them to take part in more advanced interventions.

More interventions are being designed for those with persistent minimal verbal language at school age, some targeting spoken language and others targeting language pre-requisites (e.g. communication via other means, imitation, joint attention, symbolic understanding and comprehension). One reason why the evidence base remains weak is that it is difficult to include minimally verbal participants in large-scale RCTs, which use robust standardised measures. Their presentation lends itself better to observed measures, taken across multiple natural interactions, where behaviours are evaluated and counted, by multiple raters, which is more labour intensive and thus difficult to do on a large scale. An additional difficulty is that recruiting minimally verbal children into studies is also challenging. Work is ongoing to remedy this imbalance in the literature and apply more vigorous statistical methods (Tager-Flusberg & Kasari, 2013).

Finally, evidence emanating from intervention studies for children with other clinical conditions impacting speech production may be useful to consider, and there is some overlap (e.g. PROMPT, AAC use). Evidence based interventions devised for childhood apraxia of speech are also limited in quantity and quality. Future research should investigate if aspects of accepted practice for these disorders could be adapted for minimally verbal autistic children and in particular whether novel multimodal approaches in this field are a viable addition to the therapeutic arsenal.

A case series design was chosen to evaluate the intervention in this thesis (Chapters 4 and 5), due to a number of reasons. Firstly, although RCTs do provide the most robust test of intervention efficacy (Ebbels, 2017), they are not always feasible. Minimally verbal autistic preschoolers are difficult to recruit and evaluate in large numbers. I visited participants in their homes, as a clinic visit may have

been too onerous given the children's profiles. Their heterogeneity also meant that speech targets may need to be individualised, and although children could have been allocated into groups, the degree of individual variation would mean it may not be possible to balance the groups on every variable of interest. It is also not prudent to launch an RCT until there is some indication of feasibility of an intervention and that it is likely to have a positive impact (see Figure 7, adapted from Craig et al., 2006)

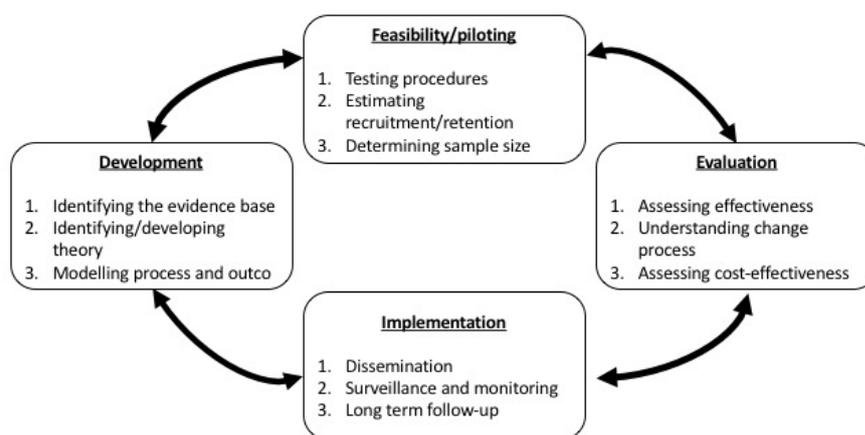


Figure 7. Key elements of the intervention development and evaluation process (adapted from Craig et al., 2006, p.8).

Additionally, numerous enhancements to case study design have been recently proposed, in order to reduce bias and increase confidence in any treatment effects. Among these include the use of randomisation and meta-analytic approaches to increase scale (Rvachew & Matthews, 2017).

In a randomized single case design, randomization is introduced by randomly selecting the week that intervention will commence, such that probes occur continuously during A weeks (baseline) and B weeks (intervention). The mean difference score of A versus B weeks can be computed and compared to all other possible permutations via resampling, to determine a distribution of all possible mean difference scores. The obtained score can then be compared to this distribution to derive a p-value, indicating the likelihood of this mean difference

score occurring by chance. When multiple participants undergo this process, the resulting p-values can be evaluated via meta-analysis to determine the likelihood of them occurring by chance. Further experimental controls will be provided by blind second-coding of probe data and the inclusion of control probe items. This procedure is outlined further in Chapter 5.

3. Study 1: Longitudinal Study

Does phonetic repertoire in minimally verbal autistic preschoolers predict the severity of later expressive language impairment?

3.1. Introduction

In Chapter 1 I reviewed empirically and theoretically motivated factors, thought to drive language variation in autism. Identifying early risk and protective factors for expressive language is important for identifying theoretically sound intervention targets and understanding individual differences in language outcome.

Several prospective and retrospective studies have evaluated the contribution of empirically tested and theoretically motivated predictive variables to early expressive language growth in autistic children. Expressive language is either a continuous outcome variable (e.g. vocabulary size) or a categorical one (e.g. acquisition of phrase speech). One type of prospective study tracks infants from an early age, who have higher than usual chance of receiving an autism diagnosis, due to having an autistic sibling. This method has the advantage of exploring prodromal development, however it can result in cohorts with very diverse diagnostic profiles and expressive language skills. Another prospective approach is to establish a more homogenous cohort of young children, who meet autism diagnosis and minimal language criteria (e.g. Yoder et al., 2015). If one is particularly interested in what drives and sustains expressive language difficulties for certain children, it is important to establish a relevant and homogenous cohort, recognizing that predictors may vary in influence according to a child's age and stage of development.

As I concluded in Chapter 1, it is likely that a variety of early observable factors, both child-related and environmental, feed into language abilities in autism in an interactive fashion during the course of development. Positive correlations have

been found between later expressive language and earlier attention to speech, joint attention, receptive language, communicative intent, imitative and non-imitative motor skills, play skills, speech production abilities as well as various features of the input (Yoder et al. 2015).

A common problem in longitudinal studies of language is that many of the putative predictors are highly inter-correlated, making it difficult to isolate causal mechanisms. In a bid to address this, Yoder et al. (2015) undertook a 16-month longitudinal study to isolate value-added predictors of expressive language growth in minimally verbal autistic preschoolers (mean age 2;11, n=87). The approach tested nine predictors, identified from the literature as well as two background variables (autism symptom severity and cognitive impairment). Value-added means that the correlation between predictors is taken into account during model selection. Predictors retained in the model were parental responsiveness, child response to joint attention, child communicative intent, and consonant inventory.

The rationale for a causal role for the first three of these predictors has been extensively explored and due to their perceived 'malleability', they have been included as core developmental targets in early interventions (e.g. Green et al., 2010, Carter et al., 2011, Kasari et al., 2014b).

Less is known about the role or malleability of consonant inventory, which indexes speech production ability. As discussed in Chapter 1, a growing literature also suggests that early vocal development difficulties may strongly impact spoken language development in autism. Young autistic children make fewer speech-like vocalisations relative to typically developing peers (Warlaumont & Oller, 2014; Plumb & Wetherby, 2013). Investigations of infant siblings of autistic children (who have a higher chance of obtaining an autism diagnosis), also indicate early differences in vocalisation rate and quality (Paul, Fuerst, Ramsay, Chawarska & Klin, 2011; Chenausky, Nelson & Tager-Flusberg, 2017; Patten, Belardi, Baranek, Watson, Labban & Oller, 2014). A recent meta-analysis concluded that pre-verbal vocalisations are correlated with concurrent and later expressive language in young autistic children (weighted effect size of $r=.50$, McDaniel et al., 2018).

The reasons behind limited vocal development in some autistic individuals are yet to be fully elucidated. The speech attunement theory (Shriberg, Paul, Black & van Santen, 2011; Paul et al., 2013) suggests that it is the failure to attend to others' verbal output ("tune in"), combined with limited motivation to interact and thus practice their own speech production ("tune up"), that results in some autistic children's poor expressive language development. This view links expressive language development to core autism features rather than a speech-specific difficulty. Empirical evidence for this theory comprises studies which show that although vocal development is often delayed in autism, phonetic development is in line with overall language development and does not follow an atypical trajectory (e.g. Shriberg, Paul, Black & van Santen, 2011). However, Shriberg et al. (2011) selected a sample that would not include the most severely speech impaired children (fluent language production and mental-age above 4), making it difficult to generalize these findings across the autism spectrum.

Another hypothesis is that reduced consonant inventory reflects the presence of a speech-motor co-morbidity, which would constitute an additional barrier to developing expressive verbal language. Motor and imitation problems have been observed to occur early in autism (e.g. Zwaigenbaum, Bryson & Garon, 2013). Early motor skills and later communication abilities have been linked in prospective (Bhat et al., 2012; LeBarton & Landa, 2019) and retrospective studies (Mody et al., 2017). Gernsbacher, Sauer, Geye, Schweigert and Hill Goldsmith (2008) found a significant relationship between infant and toddler oral and manual imitation skills and later language outcome in autism. Stone, Ousley and Littleford (1997) found that not only were autistic toddlers more impaired in the ability to imitate body movements than developmentally matched clinical controls, but this skill predicted speech development 14 months later. Pecukonis, Plesa Skwerer, Eggleston, Meyer and Tager-Flusberg (2019) found manual imitation skills predicted concurrent expressive language in minimally verbal autistic children and adolescents (n=37), whilst play and joint attention skills were not significant predictors.

A more specific oral motor dysfunction could contribute to speech delays in some autistic children (Adams, 1998). Belmonte et al. (2013) described a subset of autistic children whose receptive language outpaced their expressive skills, and these same children also had marked initial and ongoing oro-motor difficulties. Tierney, Mayes, Lohs, Black, Gisin and Veglia (2015) observed high co-morbidity of autism and apraxia in a clinical sample (of 11 autistic individuals, seven also met criteria for apraxia of speech). Smith, Miranda and Zaidman-Zait (2007) found verbal imitation ability (scored simply as present or absent) significantly predicted later language milestones. Chenausky et al. (2019) identified a subgroup within a minimally verbal or low verbal sample (n=54), for whom childhood apraxia was either suspected or could not be ruled out. Expressive language for these participants was predicted by speech skills alone, whereas for the participants with no identifiable speech-motor difficulties, only receptive language was predictive.

Further study is warranted to investigate the role of early speech production abilities in expressive language development. Speech production abilities are typically indexed by a consonant inventory taken from a communication sample, but this method may not be appropriate across the spectrum of verbal ability. When a skill is emerging, it is advantageous to incorporate various sources of reporting (observation, parent report, experimental measures, Broome, McCabe, Docking & Doble, 2017). If a child does not enjoy interacting with experimenter, the consonant inventory may underestimate the child's true competencies. Consonant inventories from brief samples may also be unreliable (Van Severen, Van Den Berg, Molemans & Gillis, 2012). Thus, a parent reported measure of communicative sound production may be helpful. Given previous findings regarding predictive value of presence/absence of verbal imitation, a measure that includes elicited sounds could also facilitate a fuller picture of a child's speech skills. Combining these approaches in a composite would aim to reduce error by measuring speech skills from multiple angles.

Auditory processing and speech perception difficulties may also be atypical in autism and could be another source of variance in language outcomes (Boucher, 2012; Haesen, Boets & Wagemans, 2011; Kujala, Lepistö & Näätänen, 2013;

O'Connor, 2012). This hypothesis is difficult to test in young minimally verbal autistic children, however several studies have done so using event-related potentials mismatch paradigms. Key, Yoder & Stone (2016) compared event-related potentials of age-matched autistic (n=24) and typically developing children (n=18). They found reduced consonant differentiation in the autistic group, which was correlated with degree of discrepancy between verbal and non-verbal skills. Matsuzaki et al. (2019) used an oddball paradigm with vowel stimuli to examine mismatch fields in 84 typically developing and autistic children, some of whom (n=9) were minimally verbal. Degree of delayed auditory discrimination correlated with language skills.

The current study aims to apply Yoder et al.'s (2015) findings to an independent sample over a 12-month period, and to further explore the possible link between speech production abilities and later language development in a group of minimally verbal autistic preschoolers. Specifically, I compare the predictive power of a multi-faceted speech skills composite and a novel alphabet knowledge measure, with that of consonant inventory alone. I use the value-added predictors identified by Yoder et al. (2015) as a starting point, rather than seeking to re-evaluate their value-added nature.

3.2. Methods

A longitudinal correlational design was used to evaluate early predictors of later expressive language growth in a group of minimally verbal autistic preschoolers. The experiment design, hypotheses and analysis plan were pre-registered prior to data collection on <https://osf.io/x2wgc>. The pre-registered protocol was followed except where specified below.

3.2.1. Participants

Recruitment took place over a 7-month period. Twelve children were recruited via social media, referrals from independent professionals, specialist nurseries and units. A further 20 participants were recruited via the ASD-UK research database, an agency who help recruit autistic participants for research projects in the UK (<http://www.asd-uk.org>).

Ethical approval was obtained from the UCL Research Ethics Committee (Project ID 9733/001) and informed written consent was sought from parents on behalf of each participant.

The flow chart in Figure 8 demonstrates how the sample of 27 participants was reached from initial enquiries from 52 families.

Participants were aged 2-5 years at intake, had a confirmed diagnosis of autism and presented at Time 1 as minimally verbal, defined here as fewer than 24 spoken words as reported by parents. Four participants displayed significantly more words and phrases at Time 1 (both observed and by parent report) and were thus excluded from the main analysis. A further participant was excluded from analysis due to providing dependent variable data for only one time-point. The following exclusions were also applied: epilepsy; known neurological, genetic, visual or hearing problems; English as an Additional Language.

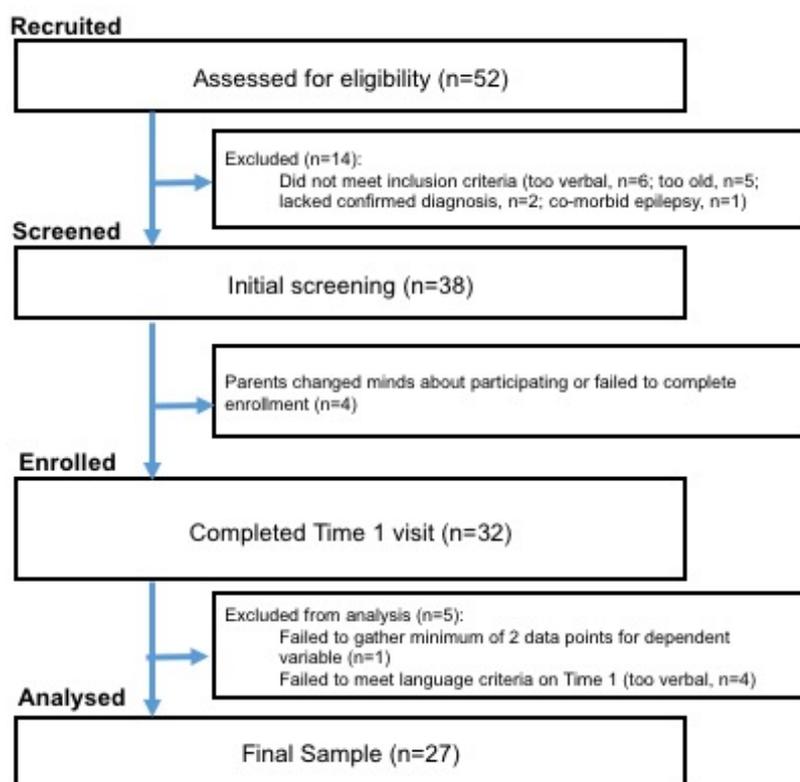


Figure 8. Recruitment Flow Chart Study 1

My original protocol stated that I would include participants with fewer than 20 spoken words by parent report, which is in line with Kasari et al. (2013) and Yoder et al. (2015). However, this criterion was expanded to 24 words in order to include three participants with 21, 22 and 23 reported words respectively, in order to maximize sample size. Each of these ‘borderline’ children also only uttered up to five different words during the 20-min Communication and Symbolic Behavior Scales (CSBS) language sample, which provided an additional check on expressive language status and is consistent with participant language use in Yoder et al. (2015). These ‘borderline’ children would still qualify as having a small repertoire of words and phrases (Kasari et al., 2013) and meet the definition of preverbal language stage (Tager-Flusberg et al., 2009).

At Time 1 the final sample thus comprised 27 children (male: 21, female: 6), who were aged between 35 and 62 months (mean=50, sd=7.6). This is approximately 15 months older than the Yoder et al. (2015) sample, who were aged 20 to 47

months (mean=35, SD=7). This is an unintended consequence of the difficulties recruiting this sample in the UK context: the original protocol targeted 40 participants, aged 24 to 48 months.

Parents reported 24 participants to be White, one to be Black, one to be Asian and one to be Mixed Race. The formal education levels of the primary caregivers were distributed as follows: 11 completed high school, eight completed university education and eight completed post-graduate studies or equivalent. Additional descriptive information on participants is provided in Table 3.

3.2.2. Variables

Variables were divided into background variables, predictor variables and a dependent variable, as shown in Table 2. The background variables merely serve to characterize the sample and were not entered into the statistical model. Further description of data transformation criteria is set out in '*anticipated data transformations*' below.

Table 2: Variables measured in Study 1

	Time	Measure	Procedure	Transformation
Back-ground Variables	1	Autism Symptom Severity	CARS (Schopler, Reichler & Renner, 1988) raw score	N/A
	2	NVIQ	Visual Reception and Fine Motor subtests of Mullen Scales of Early Learning (Mullen, 1995) transformed into Developmental Quotient (developmental age/chronological age)	N/A
	1	Receptive Language	Oxford CDI words understood (Hamilton, Plunkett & Schafer, 2000) raw score	N/A
Predictor variables	1	Intentional communication	Number of communicative acts across all pragmatic functions during communication temptations sub-section of CSBS (Wetherby & Prizant, 2002) converted to a rate due to differing sample lengths (Cohen & Cohen, 1984)	N/A
		Response to Joint attention	6 presses modified from ESCS (Mundy, Delgado, Block, Venezia, Hogan & Seibert, 2003) proportion correct	Square root
		Parent responsiveness	Parent input derived from recorded naturalistic interaction at Time 1 (coded for % of contingent linguistic responses following child lead)	N/A
		Consonant inventory	CSBS Scale 11 (Wetherby & Prizant, 2002) raw score	Square root
		Phonetic repertoire	Composite comprising Elicited Phonemes, Reported Phonemes and Observed Phoneme inventory	Square root
		Alphabet and phonics score	Percentage of correct trials	Square root
Dependent Variable	1,2,3 & 4	Expressive language	Oxford CDI words spoken (Hamilton, Plunkett & Schafer, 2000) raw score	Log10

Note: CARS: Childhood Autism Rating Scale; CDI: Communicative Development Inventory; CSBS: Communication and symbolic behavior scales; ESCS: Early Social Communication Scales; Time 1, 2, 3 and 4 separated by 4 months (mean =4.1; sd=0.4)

3.2.2.1. Alternative Phoneme Measures

Phonetic Repertoire

At Time 1, three additional measures of speech sound repertoire were taken, in order to compare their combined predictive power versus consonant inventory alone. These comprised Observed Phoneme Inventory (derived from CSBS language sample, Appendix A), parent reported core phonemes used communicatively (derived from Reported Phonemes questionnaire in Appendix B) and Elicited phonemes, which used a procedure adapted from Kaufman Speech Praxis Test (Kaufman, 1995) to determine participants' existing echoic repertoires with single phonemes, e.g., /m/).

Alphabet and Phonics Knowledge

To accurately measure speech perception skills in this cohort would require a laboratory visit and ample testing time. Instead the child's ability to receptively identify different speech sounds using a letter/sound recognition paradigm was measured at Time 1, in order to determine whether an ability to link sounds with letter mappings may act as protective factor for expressive language growth. The child was asked to give the experimenter one of three letter cards upon hearing either a corresponding phonics sound or a letter name as part of a counterbalanced pre-determined sequence (see Appendix C). Scores were translated into a percentage of trials completed. This variable was added to the test battery as an exploratory measure, despite confounds with prior print exposure and global attention skills.

3.2.3. Procedure

Data was collected in children's homes in four sessions separated by four months each as summarized in Figure 9. A £5 gift voucher was provided to each child following each visit. Predictor and background variables were taken at Time 1, apart from the non-verbal IQ measure which was carried out on Time 2 to

accommodate the limited concentration span of participants. Additional demographic information was gathered at Time 1 (see Appendix D for Family Background Questionnaire). At each visit, the dependent variable, Oxford CDI words spoken, was completed by parents. This is a UK adaptation of the MacArthur-Bates CDI measure (Fenson, Marchman, Thal, Dale, Reznick & Bates, 2007). Additionally, at each time-point parents completed a Therapy questionnaire detailing the type and amount of weekly therapy received by the child in the preceding 4 months (see Appendix E). Testing sessions were video and audio recorded for later coding and transcription.

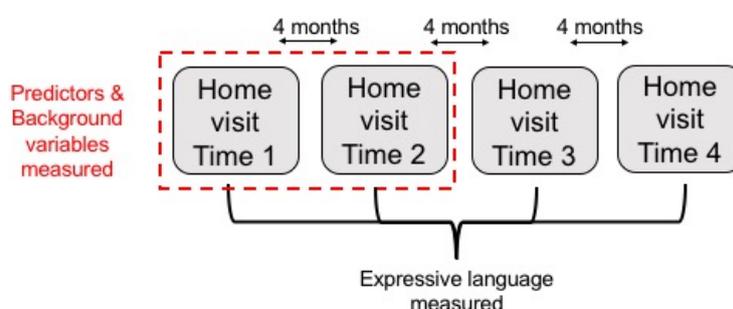


Figure 9. Data Collection for Study 1

3.2.3.1. Video coding

Parent Child interaction

Parents (mothers: n=26, fathers: n=1) were given a set of developmentally appropriate toys and asked to interact as they normally would with their child for 15 minutes. The coding manual was obtained from Paul Yoder and closely followed the procedures described in Yoder et al. (2015). The following adaptations were made to the current study: communication behavior was coding using ELAN (ELAN, 2018) rather than ProCoder software, and the selection of toys used was different (see Appendix F for list of items and coding manual). As per Yoder et al. (2015), the video was divided into 5 second intervals, which were classified as

codeable or non-codeable, depending on whether both participants and their actions were visible. Each codeable interval was examined for evidence of a child lead, and if so, the referent of that lead. Child leads were attentional (e.g. looking at a referent) or physical (e.g. manipulating a referent). Each interval containing an identifiable child lead was then coded for parent response (either linguistic, physical or both). Finally, the percentage of child leads that resulted in a parental linguistic response was computed. Mean sample lengths was 15.0 minutes (sd=1.3). A random sample of 22% of all coded sessions from media files were analysed by a second coder, blind to specific research question. The intra-class correlation coefficient was .98.

CSBS

The communication temptations section of the CSBS was administered at Time 1 according to the manual. Each communication behaviour displayed by the child was coded according to its function (initiating or responding to behavioural regulation, joint attention or social interaction) and the communicative means (with or without gesture, vocalisation or words). This was also coded in ELAN and subsequent information to be extracted was as follows: number of intentional communication attempts in each category (in order to compute total communicative acts), and phoneme and consonant inventory (phonemes and consonants were only counted if they occurred as part of a deliberate communication act and were part of a syllable). Mean sample length was 24.0 minutes (sd=7.2). Correlation between the sample length and communicative acts was .32 ($p=.10$). The total communicative acts measure was conservatively converted to a rate in order to avoid bias caused by variation in sample duration (to avoid conflating shorter samples with fewer communicative behaviours, if the cause of shorter samples was behavioural or attentional) as per Cohen and Cohen (1984).

A random sample of 10-22% of all coded sessions from media files were analysed by a second coder, blind to specific research question. The variables tested for reliability were those entered into the statistical model. Inter-observer agreement

was .86 (rate of communicative acts, communicative intent); .95 (number of consonants, consonant inventory) and .99 (number of phonemes, phoneme inventory). All inter-observer agreement statistics were computed using the intra-class correlation ICC() command in the psych R package (Revelle, 2018). For additional information, I calculated an agreement matrix to determine what percentage of the time raters agreed on individual phoneme and consonant judgements (rather than overall number of phonemes in the repertoire). This was a mean of 84% for consonant inventory (sd=17%, range= 55-100%) and 80% for phoneme inventory (sd=15%, range=57-100%).

3.2.4. Data Analysis

3.2.4.1. Exclusion Criteria

The participant exclusion criteria from the pre-registration was followed, resulting in the removal of one participant who had only provided dependent variable data on one time-point.

I did not plan to exclude outliers, in order to reflect the heterogeneity in expressive language development, however four data points represented significant outliers (due to two participants making very large language gains by Time 3 which were maintained at Time 4). These data were adjusted to the time-point mean + 3 standard deviations, in order to avoid any undue influence on the analysis (Kutner, Nachtsheim & Neter, 2004; Field, 2013).

3.2.4.2. Missing Data

Very few data were missing, only one predictor data point (Reported Phonemes for one participant, >4%) and one dependent variable measure (Time 1 CDI value for one participant, >4%). The missing data were multiply imputed following Enders (2010). Measures requiring transformation were transformed before imputation (von Hippel, 2009). Forty imputed data sets were used in order to minimize bias in parameter estimates (Graham, 2009). After imputed data sets were created,

imputed scores were deleted for the one missing dependent variable data point, since not doing so may bias regression estimates (von Hippel, 2007).

3.2.4.3. *Anticipated Data Transformations*

The analysis measures used assume multivariate normality. Multivariate normality is more likely when univariate distributions do not grossly depart from descriptors of the normal distribution (Tabachnick & Fidell, 2001). All variables were transformed if they had univariate skewness $>|.8|$ or kurtosis $>|3.0|$. Transformations were selected in accordance with the principles in Tabachnick and Fidell (2001). Transformations that were applied are listed in Table 2.

3.2.4.4. *Linear Mixed Models*

All data analysis was conducted using linear mixed effects models, fit in R (R Core Team, 2017) with the `lmer()` function of the `lme4` library (Bates, Maechler, Bolker & Walker, 2015). In line with recommendations in Barr, Levy, Scheepers and Tily (2013) my analysis assumed a maximal model with random intercepts and slopes. Time was centered at Time 4, meaning that the intercept corresponded with expressive language outcome at the end of 12 months. This was deemed more meaningful than centering at Time 1 when expressive language fell within a tight range (0 to 23 words), and is in line with the approach taken by Yoder et al. (2015). Time was entered into the model as a nominal value (i.e. a number between 1 and 4) rather than on a continuous basis, given the adherence to a regular time interval between assessments, which also mirrors Yoder et al. (2015).

Model comparisons were made using the deviance statistic, or change in the -2 log likelihood, when comparing nested models. A significant change is one with a Chi squared p-value of less than 0.05. Non-nested models were compared using Bayesian information criterion (BIC).

3.2.5. Planned Confirmatory Analyses

The following specific hypotheses were tested:

Hypothesis 1: All value-added predictors identified in Yoder et al. (2015) will be significant positive predictors of expressive language in this sample (parental responsiveness, child response to joint attention, child communicative intent and consonant inventory).

Hypothesis 2a: Phonetic repertoire will provide a better model fit in predicting expressive language compared to consonant inventory.

Hypothesis 2b: Alphabet and phonics knowledge will provide a better model fit in predicting expressive language compared to consonant inventory.

3.3. Results

3.3.1. Preliminary Results

3.3.1.1. Expressive Language Growth

Descriptive measures for dependent, independent and background variables are described in Table 3.

Table 3: Descriptive Measures in Study 1

Measure	n	Mean	sd	Min	Max
Age at Time 1 (months)	27	49.6	7.6	35.4	61.8
Autism Symptom Severity at Time 1 (raw score)	27	41.3	5.6	28.5	52.5
NVIQ at Time 2 (DQ)	27	.42	.17	.13	.77
Receptive Language Time 1 (words)	26	150	111	0	342
Expressive Language Time 1 (words)	26	4.5	7.4	0.0	23.0
Expressive Language Time 2 (words)	27	13.3	16.2	0.0	48.0
Expressive Language Time 3 (words)	27	41.4	84.4	0.0	323.0
Expressive Language Time 4 (words)	27	48.7	93.8	0.0	356.0
Communicative intent (total comm. Acts)	27	23.6	12.2	7.0	45.0
Response to Joint attention (% correct)	27	25%	36%	0%	100%
Parent child interaction (% leads)	27	53%	16%	18%	84%
Consonant inventory (raw score)	27	4.3	4.1	0.0	14.0
Phonetic repertoire (raw score)	26	12.8	10.9	0.0	40.0
Alphabet score (% correct)	27	22%	40%	0%	100%
Weekly SLT therapy Time 1 (hours)	26	0.91	2.06	0	10
Total weekly therapy Time 1 (hours)	26	4.22	5.33	0	22

Note: SD: standard deviation; NVIQ: non-verbal intelligence quotient; DQ: developmental quotient (developmental age/chronological age); SLT: speech and language therapy

All participants commenced the study at a mean age of 4;2 years with extremely limited expressive language. Over the 12-month period of the study, individual expressive language growth was highly variable, as illustrated in Figures 10 and 11. Using the threshold indicated in the original sample selection criteria (< 24

words by parent report), 65% of the sample remained minimally verbal at Time 4. Furthermore, 27% of all participants were at floor on this measure at Time 4, reportedly using no words at all.

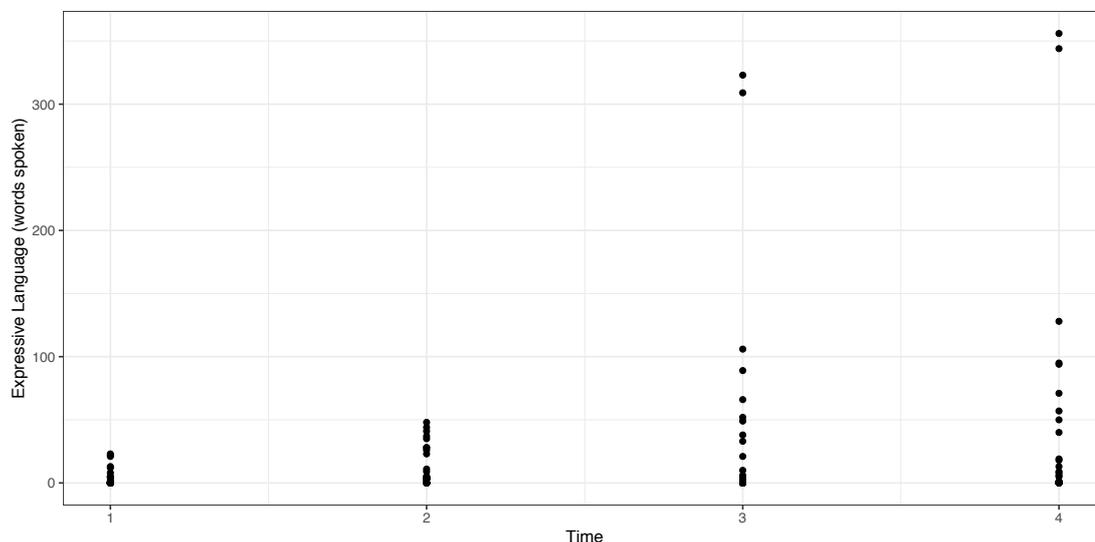


Figure 10. Expressive language (parent report of number of words child reliably produces) at each of the four time-points (each separated by four months).

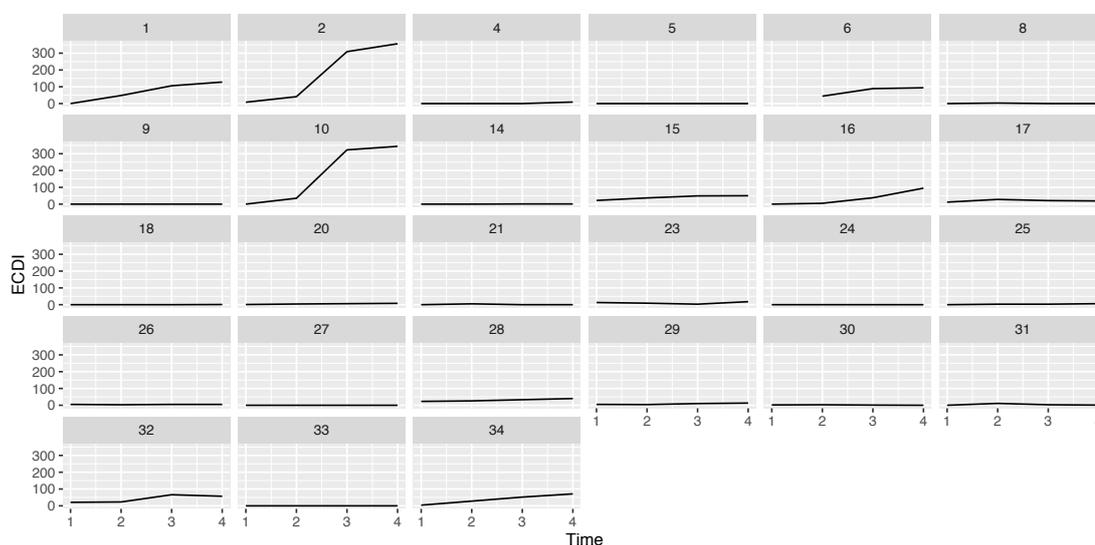


Figure 11. Individual expressive language trajectories (parent report of number of words child reliably produces) at each of the four time-points (each separated by four months).

The average gain in expressive vocabulary was 43 words (sd = 95); however, this figure is biased by the presence of two participants whose Time 3 and Time 4 scores were significant outliers. These two participants both gained over 340 words during the 12-month period. The mean gain excluding these outliers is 17 words (sd = 33).

There was high stability in expressive language, evidenced by high correlations between expressive language scores as measured at each time-point, as illustrated in Table 4. Despite equally spaced time-points, the degree of correlation was much higher between later time-points than it was between Time 1 and Time 2.

Table 4: Expressive Language Correlations

ECDI	Time 2	Time 3	Time 4
Time 1	.62***	.56**	.52**
Time 2		.90***	.83***
Time 3			.95***

Note: ** p<.01; *** p<.001; ECDI: Expressive Communicative Development Inventory (words spoken by parent report)

3.3.2. Putative Predictors

The predictor and background variables are summarised in Table 3 and their correlations are presented in Table 5. Background variables including autism symptom severity, NVIQ and Time 1 receptive language, and predictor variables consonant inventory, phonetic repertoire and alphabet score all correlated with Time 4 expressive language level. Conversely, communicative intent, parent responsiveness and response to joint attention were not significantly correlated with Time 4 expressive language. Expressive language change over 12 months (i.e. Time 4 minus Time 1 expressive language) was correlated with autism symptom severity, NVIQ, phonetic repertoire and alphabet score.

Table 5: Correlations

	2	3	4	5	6	7	8	9	10	11
Background variables										
1. Autism Symptom Severity	-0.66***	-0.41*	-0.55**	-0.40*	-0.35	-0.49*	-0.56**	-0.40*	-0.53**	-0.46*
2. NVIQ		0.50**	0.55**	0.39*	0.42*	0.46*	0.51**	0.51**	0.54**	0.54**
3. Receptive Language			0.19	0.40*	0.40*	0.20	0.29	0.41*	0.40*	0.28
Predictor variables										
4. Response to Joint Attention				0.60***	0.37	0.32	0.32	0.44*	0.19	0.12
5. Communicative Intent					0.32	0.59**	0.49*	0.19	0.26	0.03
6. Parent Responsiveness						0.27	0.31	0.03	0.21	-0.07
7. Consonant Inventory							0.87***	0.10	0.56**	0.24
8. Phonetic Repertoire								0.21	0.71***	0.39+
9. Alphabet Score									0.44*	0.57**
Dependent variable										
10. Expressive Language Outcome										0.77***
11. Expressive Language Change										

Note: +p<.051 * p<.05; ** p<.01; *** p<.001

One composite variable was planned (Phonetic Repertoire), so the intercorrelations among component measures of this construct were verified. Observed Phoneme Inventory, Reported phonemes, and Elicited phonemes were all measured at Time 1. Elicited phonemes correlated significantly with Observed Phoneme Inventory ($r=.45$, $p<.05$) but not Reported phonemes ($r=.28$, n.s). Likewise, Reported phonemes correlated significantly with Observed Phoneme Inventory ($r=.46$, $p<.05$). The resulting Phonetic Repertoire measure correlated with each component $r>.60$ and also significantly with consonant inventory ($r=.87$, all $ps <.01$).

3.3.3. Confirmatory analyses

Below are the steps taken to evaluate the pre-registered hypotheses, beginning with an unconditional growth model containing random effects of individual

differences between participants on the intercept and the slope (i.e. the linear effect of time) and a fixed effect of time, then adding in the previously identified value-added predictors and finally comparing the predictive model this generates with one using alternative predictors. Coefficients for each model are set out in Table 6.

3.3.3.1. Selection of unconditional models for language growth – Model 1

A model with Time centered at Time 4, containing fixed and random intercepts and slopes was the best fit to the data, with an adjusted R squared of .07.

3.3.3.2. Conditional model using Yoder et al. (2015) predictors – Model 2

Of the four original predictors, only consonant inventory had a significant zero-order correlation with expressive language change or outcome, therefore the other three predictors were not entered into the model.

A fixed effect of consonant inventory significantly improved model fit versus Model 1 (Chi sq= 12.19, df=1, $p<.001$). The increase in adjusted R squared was .23. Adding further interactions with Time did not significantly improve the model fit, so this was deemed the best model using the original predictors, and thus the one used to compare against novel predictors to address Hypotheses 2a and 2b. A significant fixed effect of consonant inventory in this model means that Time 1 consonant inventory significantly predicts intercepts in the growth curve at Time 4, i.e. the expressive language outcome at Time 4. The significant effect of Time indicates that expressive language scores did increase with Time. The lack of interaction between Time and consonant inventory in this model indicates that consonant inventory does not significantly predict the rate of individual growth (slope), over and above its effect on the Time 4 outcome (intercept).

3.3.3.3. Testing novel predictors - Model 3

The second objective was to test the suitability of two alternative predictors to be used in the model in place of consonant inventory. Model fit was compared using Bayes Information Criterion (BIC) since the models were not nested (i.e. one model did not contain all the parameters of the other model).

Replacing consonant inventory with Phonetic Repertoire, resulted in a decrease in BIC (148 vs. 139), and therefore indicated an increase in model fit. Adjusted R squared for this model was .45, an increase of .16.

The same process was used to test Alphabet score at Time 1 as a predictor. Taking Model 2 and replacing consonant inventory with Alphabet score resulted in a higher BIC (160), indicating a worse model fit. Therefore, no model containing Alphabet score was included in analysis.

Table 6: Model Summary

Predictors	Model 1		Model 2		Model 3	
	coefficient	SE	coefficient	SE	coefficient	SE
Intercept	0.991***	0.132	0.450*	0.188	0.109	0.187
Time	0.158***	0.032	0.159**	0.044	0.150**	0.044
Consonant Inventory			0.290***	0.072		
Phonetic Repertoire					0.267***	0.043
Adjusted R squared		.07		.29		.45
Change in Adjusted R squared		.07		.23		.16
Bayes Information Criterion		163		148		139

Note: * $p < .05$; ** $p < .01$; *** $p < .001$; SE: Standard Error.

3.4. Discussion

3.4.1. Expressive Language Growth

A significant proportion of participants (65%) remained minimally verbal after 12 months, at mean age 5;2. This figure is somewhat greater than the 40% reported to remain minimally verbal in Yoder et al. (2015), however the time periods are also not directly comparable (16 vs. 12 months). Few similar longitudinal studies are directly comparable due to differences in design, definition of minimally verbal, or sample characteristics (e.g. Anderson et al., 2007; Norrelgen et al., 2014; Bal, 2016).

Children made a mean gain of 45 words during this study, which is lower than the 75 words (sd=95) after 16 months reported in Yoder et al. (2015). When two significant outliers are excluded the comparison figure shrinks to 17 words (sd=33), which suggests that on average children on this study are not making progress at the same rate as the children observed in Yoder et al. (2015). A potential explanation is that participants in the current study were recruited at age 3 to 5 (mean 4;2) rather than age 1;8 to 3;11 (mean 2;11) in Yoder's study. It is possible that some of the younger children in Yoder et al. (2015) were less severely impaired and did not have such persistent expressive language impairment, but were experiencing a transient delay in language development which partially resolved during the study period. This suggests that my sample may include children with more severe difficulties and greater difficulty acquiring expressive language. The children in Yoder et al. (2015) share a similar range and mean to my cohort for developmental ratio. Like this cohort, they have a highly variable receptive language score at the start and end of the study, and by design they start the study with a similarly limited expressive vocabulary. Another possibility may be differences in intervention receipt; however, Yoder et al. (2015) do not report information about the types or duration of interventions children received and this study is not designed to evaluate the impact of intervention on expressive language outcome. Instead, the current study has focused on value-added predictors identified by Yoder and colleagues.

Expressive language measures were quite stable within each participant over time, and particularly between adjacent time-points at Time 2, 3 and 4 (all $r \geq .90$), correlation was only $r = .62$ between Time 1 and 2. This could be an artefact of the very low variability in the initial Time 1 expressive language level, or could reflect a decrease in measurement error over time in parental judgements of language skills. Language was stable across this period as those children with larger vocabularies at Time 1 tended to have the largest vocabularies at Time 4. This stability does not imply no change in language, 67% children showed some improvement in language scores, with an average increase of 17 words, excluding two outliers. Bornstein et al. (2018) used the Avon Longitudinal Study of Parents and Children dataset to evaluate stability of language over 13 time-points and 15 years. They found that core language was stable from an early age in both typical ($n = 4,111$) and atypical groups, including autism ($n = 89$). Average stability across all time-points was .65 for the autistic children and .56 in the typical group.

3.4.2. Confirmatory Analyses

This study did not find a meaningful relationship between three of the original putative predictors and expressive language. It is difficult to draw a firm conclusion from null results, and given the small sample this may be due to lower statistical power. Although the sample is smaller than the 87 participants used by Yoder et al. (2015), the current study also examines far fewer variables (Yoder et al. tested nine putative predictors and two background variables). However, all but phonetic repertoire were measured via single variables, whereas Yoder et al. (2015) used aggregate measures, which are known to enhance stability and validity. The sample size reflects the difficulty of recruiting this hard to reach population as well as financial and practical constraints on data collection in a repeated measures design. It is also possible, as mentioned above, that the sample might be qualitatively different to the Yoder et al. (2015) sample due to the older age at which participants were recruited, which could result in different predictive relationships: this cohort may have had more severe speech-motor deficits that are distinct from the social variables that associate with language development.

Finally, the high number of participants who continue to be at floor on the dependent variable may attenuate correlations with putative predictors. A key focus of future work should be ensuring that conclusions from younger and broader samples can be generalised to those with the most complex communication difficulties (e.g. Pecukonis et al., 2019).

In contrast, the significant correlation between early consonant inventory and expressive language growth seen in Yoder et al. (2015) was replicated in this sample. This adds to prior emerging evidence that speech production abilities are related to expressive language development in autistic preschoolers (McDaniel et al., 2018) and that speech production is worthy of further consideration when devising interventions (e.g. Chenausky et al., 2018). It is also noteworthy that there was no interaction between Time and consonant inventory in this sample, so consonant inventory and Time were both significant predictors of participants' expressive language outcome at Time 4, but consonant inventory did not predict slope of their growth curve to get to that outcome. Given that all children started the study with expressive language scores in a tight range at Time 1 (0 to 23 words), their growth rate is likely to be highly correlated with their Time 4 outcome, which may explain this lack of effect.

In this sample, replacing consonant inventory with a composite of three phonetic measures (phoneme inventory, elicited phonemes and reported phonemes) resulted in a better model fit and explained more variance. This supports the idea that for minimally verbal autistic children, a broader measure of speech skills, incorporating information from multiple sources, may be more nuanced and thus a better predictor of the same underlying construct, a sentiment echoed more generally in Tager-Flusberg et al. (2009).

The composite measure, phonetic repertoire, comprised three speech measures which were only moderately correlated with each other (r values ranging from .28 to .46), yet this measure proved to be a stronger predictor of expressive language than consonant inventory. Reasons for the low correlation may include measurement error. Some parents reported they found it difficult to evaluate the

communicativeness of their child's babble and identify the specific sounds within it, as required by the Reported Phonemes measure. Equally, children's engagement during the Elicited Phonemes task varied considerably, which could have understated some children's actual skills. On the other hand, the measures may be expected to truly vary as they measure different skills. Those needed for Elicited Phonemes (to attend to, process, and copy a specific sound, with no intrinsic motivation and with an unfamiliar interlocutor) compare with those for Reported Phonemes, where motivation may be present in the natural home environment (e.g. to obtain a desired item) and the interlocutor is familiar. Furthermore, no specific speech sound may be necessary (a gesture and a vocalisation may suffice to convey information) and performance pressure is reduced. Phonemic repertoire may have been a more informative predictor because different facets of speech skills were combined.

Conversely, Alphabet Knowledge did not appear to have a consistent relationship with expressive language in this sample, nor did it correlate with other phonetic repertoire measures. This novel measure was not continuously distributed across the sample. Given attentional difficulties, future work may employ parent questionnaires as a more effective and accurate way of tapping alphabet and phonics knowledge and relating them to language development.

The weaker correlation in this sample between socio-communicative measures (communicative intent, parent responsiveness and response to joint attention) and expressive (verbal) language, supports the idea that some minimally verbal autistic children could have an additional disorder of speech-motor development. If this were the case, stronger socio-communicative skills would not act as protective factors for expressive language to the same extent that they do in younger and thus more diverse minimally verbal autistic cohorts. To illustrate this point, a few children in this sample were frequent and productive users of alternative forms of communication (Makaton; speech generating application), despite their lack of verbal output. This is a further indicator of a specific additional difficulty with speech production rather than motivation or symbolic understanding. Belmonte et al. (2013) identified a motor-impaired subgroup comprising one third of participants

(cohort aged 22 to 65 months at intake, n=31). These children had weaker oral-motor skills and a disparity between their receptive and expressive language level, reinforcing the conclusion that motor difficulties contributed to their lack of speech progress.

Potentially relevant predictors not evaluated in the current study are non-verbal cognition and autism symptom severity, since they were not deemed to be value-added predictors in Yoder et al.'s (2015) findings. Previous cohort studies have identified associations between these variables and later language (Wodka, Mathy & Kalb, 2013; Thurm et al., 2007; Anderson et al., 2007; Thurm, Manwaring, Swineford & Farmer, 2015). NVIQ and symptom severity do associate with language outcomes in the current study such that those with more severe and pervasive development deficits have more limited consonant inventories and make more limited progress. However, it may be more useful to identify specific predictors, which are more narrowly defined and suitable as potential intervention targets, rather than confirm the pervasive association between later language and global measures of non-verbal cognition or symptom severity. Bal et al. (2019) recently investigated the role of early predictors in two independent cohorts of language-delayed autistic preschoolers (n=267) and identified fine-motor skills as a strong predictor of later expressive language. Their study highlighted the importance of looking at specific skill domains rather than broader indices of developmental level.

3.4.3. Limitations

This study has several limitations. The sample size is relatively small, which impacts statistical power. Secondly, for financial and logistical reasons, no formal independent diagnostic verification process took place (e.g. ADOS assessment). However, each family reported that autism had been diagnosed by a qualified health professional, and children scored a mean of 41.3 on the Childhood Autism Rating Scale (CARS) autism symptom severity assessment (only one child scored less than the 30 cut-off score). Thirdly, the study design involved a series of home visits. The data generated in such contexts is more vulnerable to measurement

error and confounding factors, due to poorer control of the testing environment, e.g. presence of pets, siblings, television screens and other distractions. However, home visits are preferred by families of children with complex needs and facilitate their participation, thus creating a broader representation of families within the study. Therefore, greater ecological validity was judged to be worth the trade-off with experimental precision. Finally, in order to limit testing time, single estimates were used for most predictor measures and for the dependent variable. Composite scores would have created more robust estimates and been preferable, however this is unlikely to substantively change the outcomes of this study.

3.4.4. Conclusion

These results underscore the striking variation in expressive language development during a 12-month period for a cohort with fairly homogenous starting vocabularies (0 to 23 words), with some remaining at zero words and others in excess of 340. They also further highlight the independent contribution of speech production abilities to expressive language development in minimally verbal autistic children.

The current findings strongly suggest that speech production may reflect an additional deficit for minimally verbal autistic children, rather than assuming that severe expressive language deficits are a consequence of core autism features. If we aim to help those autistic children most at risk of persistent expressive language difficulties, we need to understand the drivers of language growth more precisely and ensure that our conclusions are based on research evidence that includes this population, so that findings can be generalised and additional barriers to communication identified and addressed.

Future work could incorporate longitudinal measures of phonetic repertoire in order to build a more informed picture of what predicts phonetic abilities in this population (e.g. Woynaroski, Watson, Gardner, Newsom, Keceli-Kaysili & Yoder, 2016). Both segmental (i.e. phonetic) and supra-segmental aspects of preverbal vocalisations (e.g. prosody, utterance length, “speechiness”) warrant further examination. The

use of automated analysis of day-long recordings as a potential method for future research would also make studies of this nature more feasible (Woynaroski et al., 2017, Swanson et al., 2018). Finally, ways in which speech production could be supported in this group should be developed and evaluated (e.g. Chenausky et al., 2018).

4. Study 2: Intervention Feasibility

Feasibility of an app-based parent-mediated speech production intervention for minimally verbal autistic children: insights from a case series

4.1. Introduction

As explained in Chapter 1, multiple risk factors may interact and combine to impact language acquisition in autism and the effects of remaining minimally verbal can be widespread. Identifying barriers to spoken language development and tailoring interventions accordingly is therefore an important clinical and research aim. Longitudinal studies have shown a host of variables to predict expressive language in young preverbal autistic cohorts (e.g. parent responsiveness, child joint attention skills and communicative intent) and these findings have informed intervention design (e.g. Green et al., 2010, Carter, et al., 2011, Kasari et al., 2014b). These studies have shown that parent and child interactive behavior may be malleable, but downstream effects of enhanced joint engagement on child language measures are not always apparent. However, mounting evidence points to additional speech-motor barriers to language development in some autistic children (Adams, 1998; Gernsbacher, et al., 2008) which could explain different predictive patterns when older or more impaired cohorts are examined (Chenausky et al., 2019; Pecukonis et al., 2019; Saul & Norbury, 2020). In particular, speech production skills (as indexed by consonant inventory) may be an important predictor of expressive language growth in minimally verbal autistic children (Yoder et al., 2015; Saul & Norbury, 2020). However, there is not yet an established intervention tailored to developing speech production skills in this population.

The evidence base for interventions focusing on speech skills for minimally verbal autistic children is sparse. As described in Chapter 2, Brignell and colleagues (2018) completed a systematic review of communication interventions for minimally verbal autistic children, which only identified two high quality studies.

Only one of these targeted spoken language, and this was via a parent-mediated focused play therapy for 32-82 month-olds (Siller et al., 2013). This intervention focussed on improving engagement and broad communication goals rather than speech skills. Language interventions that have been rigorously tested tend to be multi-faceted, targeting a broad range of language and cognitive skills, and with a diverse young cohort. Targeted interventions, aiming to improve a single skill (e.g. speech production), have a lower quality evidence base, constituting small group studies or case series, which have limited power or experimental control. The participants in these studies tend to be older, globally more impaired and when progress is made it is often in a single skill that may not be generalised, following a large number of intervention sessions. Approaches to improving speech production skills in these small studies have been diverse, with the majority of studies reflecting behavioural approaches (structured: Lepper et al. 2013; Esch et al., 2005; Esch et al., 2009; contrasting structured and naturalistic: Paul et al., 2013; combining with Augmentive and Alternative Communication aids to target speech: Picture Exchange Communication System: Schreibman & Stahmer, 2014; Speech Generating Devices: Gevarter et al., 2016; Ronski et al. 2010). Non-behavioural approaches included music and/or rhythm based techniques such as auditory-motor mapping training (Wan et al., 2011; Chenausky et al., 2016, 2018) or melodic based communication therapy (Sandiford et al. 2013) and sensory-motor training (Rogers et al., 2006, Sweeney & Lebersfeld, 2016).

We sought to design and create an intervention to examine the causal relationship between speech skills and expressive language development. The intervention reported in this chapter employed two techniques novel to language interventions for autistic children: video modelling and cued articulation.

Video modelling is a technique whereby a target behaviour is demonstrated via pre-recorded video played to the learner via an electronic device, rather than through live demonstration. The person demonstrating the behaviour in the video (the model) can be a peer, an adult, or the learner themselves (video self-modelling). Videos are designed to accentuate important features of the behaviour and remove distracting extraneous stimuli, and video modelling interventions may

involve repetition of the stimuli to enhance learning. Several meta-analyses have concluded that video modelling can be effectively used to promote acquisition of a variety of academic, social, communicative and functional skills in autistic children and adolescents (Bellini & Akullian, 2007; McCoy & Hermansen, 2007; Wang & Koyama, 2014). To our knowledge, video modelling has not been investigated as a potential tool for speech production training, however it has been used to promote spontaneous requesting via speech generating devices (Copple, Koul, Banda & Frye, 2015) in participants with a similar profile to those in the current study.

Cued articulation (Passy, 2003) is one way of visually indicating how a speech sound is made, for those who do not find it easy to copy speech sounds. The rationale behind cued articulation is that each phoneme is accompanied by a hand gesture which provides a visual clue as to how and where the sound is made by the articulators, for example a 'p' sound starts with rounded lips that open when the sound is released, and the 'p' cued articulation gesture is index finger and thumb creating a circle which then opens as the sound is made. Unlike manual imitation, speech sound imitation cannot be physically prompted, and because much of it occurs inside the mouth, it can also not be viewed. Cued articulation has rarely been tested in research studies but has been widely used by Speech and Language Therapists (SLTs) in a variety of conditions including English as an Additional Language, hearing impairment, autism, and speech sound disorders (McLeod & Baker, 2014).

The intervention was devised to encourage children to practice speech sounds with a parent, in order to increase their speech sound repertoire. It aimed to take into account specific features of autism and adapt typical approaches to speech skill training accordingly. High quality intervention evidence for children with speech-motor difficulties is lacking (Morgan & Vogel, 2009). Nevertheless, widely delivered interventions frequently include a) the provision of high quality multi-modal models of sounds to be imitated and b) facilitating frequent practice of the sounds incorporating the principles of motor learning (Murray et al., 2014). For a

myriad of reasons, typical approaches may be problematic for autistic learners and need to be adapted.

Repeated modelling of sounds in the natural environment is designed to draw the child's attention to how to articulate a given sound, often supported by additional visual cues such as the cued articulation signs. If a child is minimally verbal, parent-child interactions may not afford as many natural opportunities for the parent to model the sound. Briefly presented multisensory social input (e.g. sound and lip movement) may be less precisely perceived by autistic individuals (Stevenson et al., 2015). By placing the speech sound model in a very structured repetitive video with no distractions in the background, we hoped to reduce the attentional load required to process the model.

Repeated practice is needed in order to master a specific motor skill. SLTs often achieve this by playing motivating interactive games with a child, e.g. a 'fishing for sounds' game where child and therapist take turns to lift up pretend fish with a magnet fishing rod, each fish having a sound symbol or picture. The person has to say the sound aloud when they have 'fished' it. Autistic children may find interactive games with an unfamiliar SLT aversive or if learning difficulties are present, play-related tasks could increase the cognitive demands of the task (e.g. child struggling with fine motor aspects of 'fishing for sounds' game). Simplifying the task, and removing the interactive aspect may thus benefit autistic children. Motivation is of course important, and it may be possible to replace the assumed social motivation with a child's special interests, to motivate them to continue with speech practice. An example would be using video clips as a reward after attempting the target sound.

Importantly, the intervention was designed to be simple, portable and requiring no additional materials or reporting, given that engaging children in less preferred activities may be challenging enough for parents. It was thus designed to be delivered via a smartphone application (or 'app'). There were several reasons for the decision to deliver the intervention via an app. Smartphones and tablets hold much promise as cost-effective, flexible and efficient delivery systems for a range

of educational interventions, and reviews have demonstrated their effectiveness for autistic learners across a host of skills (Kagohara et al., 2013; Grynszpan, Weiss, Perez-Dia & Gal, 2014; Ramdoss et al., 2011; Boyd, Hart Barnett & More, 2015; Ledbetter-Cho, O'Reilly, Lang, Watkins & Lim, 2018). Financial and logistical constraints would have made a face-to-face parent coaching style intervention impossible to deliver: participants were geographically disparate, each had a unique profile of speech targets, each had a unique schedule of intervention delivery. Using an app prompts parents to deliver individualised targets to a specific schedule. Furthermore, weekly probe data collection would have been difficult to achieve without the ability to easily video record and upload speech attempts alongside parent ratings, which is afforded by an app. Finally, using an app makes the intervention scalable at a low cost if it proves successful, whereas a parent-mediated intervention which required clinical supervision would likely be constrained by capacity under current service delivery models.

4.1.1. Aims

This chapter describes the design of a speech sound intervention using video-modelling and cued articulation, with two stages of formative evaluation and resulting improvements to the intervention app.

The central aim was to pilot the intervention and evaluate two important aspects of feasibility for this intervention:

1. Acceptability (defined by score on parent satisfaction questionnaire); and
2. Usability (defined by adherence to the intervention in minutes spent per week).

Additional exploratory analyses were carried out to provide further descriptive information regarding both of these aspects.

4.2. Method

This section first describes the intervention used in the pilot study, and how it was designed and modified with autism community input (Stage 1 and Stage 2 of consultation). In the second section, the pilot study methodology is described.

4.2.1. Intervention design

4.2.1.1. Design Process

Intervention development is ideally an iterative process rather than a linear one, as illustrated in Figure 7, adapted from Craig et al. (2006, p.8). Once an idea is developed and piloted, evaluation should inform any refinements to the intervention, which if necessary should be the subject of further piloting and evaluation.

Our iterative design process comprised:

- a) Initial design;
- b) Stage 1 consultation and app creation;
- c) Stage 2 consultation and associated improvements to app.

I initially designed the app in collaboration with a team of students from UCL's Computer Science department. In March 2017, before coding for the app had begun, I carried out Stage 1 of the consultation (described below). Once a working version of the app had been created in May 2017, I piloted it on a small convenience sample of users (Stage 2 of consultation, described below). Afterwards, an independent programmer was commissioned to carry out the recommended changes and solve highlighted technical problems, resulting in the prototype version of 'BabbleBooster', which was used for the pilot study. This version is described, followed by brief summaries of both consultation exercises and the improvements that resulted.

There is a growing awareness that high quality autism research should directly involve autistic individuals as partners within a participatory framework. Fletcher-Watson, Pain, Hammond, Humphry & McConachie (2016) advocated for “user-centred design with relevant stakeholders” in their description of the design process for an app-based intervention game designed for young autistic children, but discussed the challenges of facilitating full participation by the user-group. In some cases, necessary input is sought from family members and experts in ‘participation by proxy.’ Given my aim to create an intervention for minimally verbal autistic children, I engaged in participatory design with parents both during Stage 1 of the consultation and for the pilot study, as they are the principal agents of delivering this intervention, and best placed to advocate for their child’s communication needs.

4.2.2. BabbleBooster Description

BabbleBooster was designed to deliver predictable and repetitive speech models via video-modelling and with cued articulation. The app-play is parent-mediated, so parents are required to watch the stimuli with their children, encourage them to make the sound, and then provide feedback on the sound in order to trigger the reward videos. Reward videos are designed with a gradient response, so a ‘good try’ at a sound (an incorrect attempt) will result in a lesser reward than an accurate response. The families were encouraged to make or upload their own reward videos, based on their understanding of the individual child’s specific motivators.

BabbleBooster was designed specifically for use in a case series design, whereby each participant acts as their own control and the outcome variable is tested repeatedly both before and during the intervention period. Each participant is given a personal intervention schedule comprising A (baseline) and B (intervention) weeks. BabbleBooster thus functions in one of two ‘modes’ depending on whether the participant is in an A or a B week:

- **Test mode:** this is during the baseline data collection period. The intervention itself is not accessible but the test module is live. Once per week the participants are prompted by text message to complete the test module (9 trials testing 9 phoneme targets once each, generating a score of 0 to 9).
- **Training mode:** this is during the intervention period. Both the intervention and the test module are live. The participants are expected to carry out the intervention as per instructions, plus complete the weekly test module as above.

Each participant is likely to have a unique profile of speech skills, meaning that targets need to be individualised. Nine Probe Phonemes were allocated at the start of the intervention by following the 'Sound Target Protocol' (see Additional File 4). Three of these were allocated as target speech sounds for training. Thus each week the test module comprised nine single trials of the nine probe speech sounds. Children were therefore tested on three trained and six untrained speech sounds. The untrained sounds were used as a control (to investigate systematic relationship between training and any improvement) and to assess whether any improvements generalized to other sounds. Weekly test score was calculated as a percentage, representing the number of phonemes correctly produced out of nine.

For each of the three **target speech sounds**, and there is a set of learning stimuli for each which comprise:

- **mandatory content:** this is unchangeable content, such as the auditory model of the sound and the cued articulation video.
- **customisable content**, which can be added to, removed, changed as much as desired by the child (with help from the parent). For example, the app comes loaded with images of items beginning with 't' for the 't' target (e.g. tiger) but the child may have a favourite toy called 'Timmy' or a family friend called 'Tania' - images of these specific items can be transferred onto the app to create a more meaningful personalised set of stimuli. Example screenshots are provided in Figure 12.

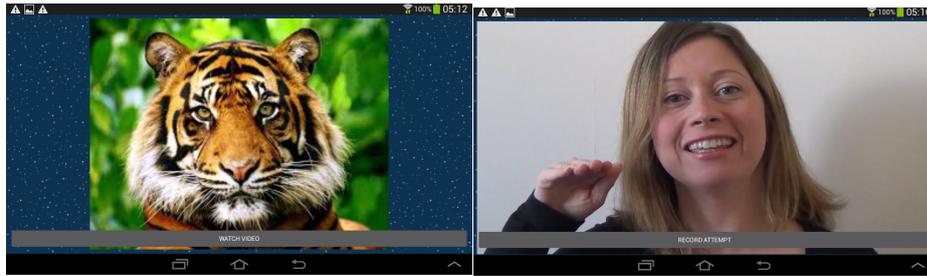


Figure 12. Learning stimuli (left: customisable; right: mandatory)

In training mode, after watching the learning stimuli, the child is prompted to attempt the speech sound. Children can use the video capture part of the app as a mirror whilst speech attempts are being recorded, and have the opportunity to play back and review their speech attempts. Parents then press one of three buttons to assign a rating to the attempt, in accordance with Table 7. Depending on the parent feedback, the child is either presented with a customizable reinforcement video as a reward, or another attempt begins. The app records progress made by the child and determines whether mastery criteria have been fulfilled and whether a new target can be selected or the existing target should continue.

Table 7: BabbleBooster parent rating buttons

Button	Meaning	Example	Consequence
Yes	Child has produced elicited sound accurately	child is asked to say /b/ and they say /b/	'Well done' video
Good Try	Child tried to make a sound but did not make the target sound	child is asked to say /b/ and they say /w/	'Good try' video
Try Again	Child does not attempt to make any sound	child is silent / shouts / cries	No video clip

Figure 13 depicts how a single 'trial' of the intervention works.

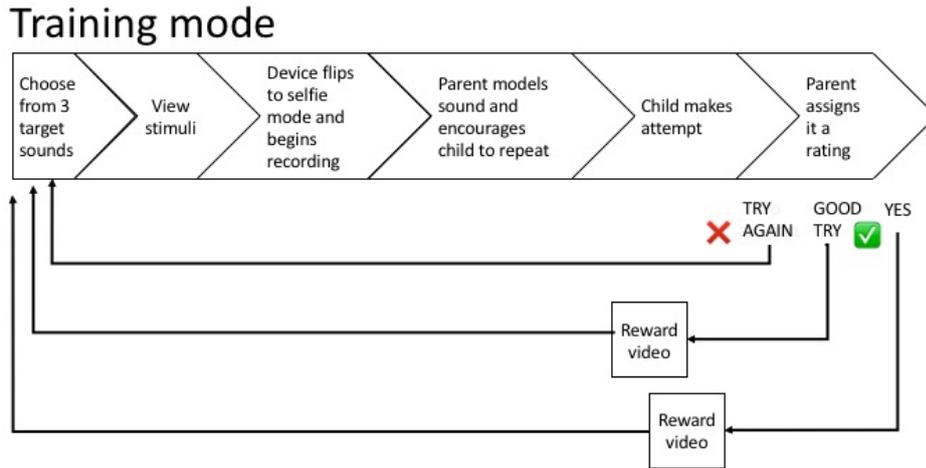


Figure 13. Depiction of one intervention trial

4.2.3. Consultation Stage 1

An initial version of the app was presented at a focus group in March 2017 with four parents whose autistic children have co-morbid language difficulties (referred to as Participants L, E, R and A). A fuller description of the focus group is provided in Appendix G. Their input contributed to the app prototype and is briefly summarized below:

Technology: all parents reported that mobile and tablet devices were inherently motivating for their children, with the most commonly used function being to access video content online (e.g. via YouTube). Content was often esoteric, user-uploaded and specific to the child’s special interests (e.g. people going on waterslides, opening toys).

Aim: all liked the idea of the app and the mirror function. Parents suggested having images that match the sounds would make it more functional. Parents would like to have input on the initial sound selection process.

Time commitment: all agreed five minutes per day is an achievable target.

Cued articulation aspect: only one parent had heard of this approach, but when her daughter was minimally verbal she had found it very helpful in progressing speech skills.

Video modelling aspect: all agreed this would be good. R remembers it being hard to get her son to look at her whilst she modelled language, and that is why she thinks he found PECs (a picture exchange communication method) easier than Makaton (a simplified form of sign language, requiring learners to copy manual signs from adult models).

Parent feedback on child productions: parents unanimously disliked the proposed red “no” button, reporting that their children were very sensitive to ‘getting things wrong’. They suggested changing it to a ‘try again’ button and altering the colours.

Reinforcing videos: all agreed that customisable content is a must-have feature of the app. Various ways of supporting parents to create content were discussed, for example providing a parent idea sheet or ‘how to’ videos.

In summary, changes to the app from this stage of consultation were:

- a) Rather than just providing the sound and a letter symbol in the sound modelling phase, this was changed to three images corresponding to the sound (e.g. for ‘b’: ‘baby’, ‘ball’, ‘biscuit’). These can be replaced or exchanged by parent customization.
- b) Rather than presenting parents with three choices for feedback buttons (‘yes’, ‘good try’ or ‘no’), this was changed to ‘yes’, ‘good try’ or ‘try again’ and red and green colours were removed.

4.2.4. Consultation Stage 2

A second feedback phase occurred prior to launching the pilot study. In May 2017 a convenience sample, which included parents with pre-verbal children with

additional needs, was invited to try the app, over the course of a week. A fuller description of this consultation stage is provided in Appendix H. This process highlighted technical glitches and generated further improvements to the lay-out.

Summary of changes from this stage:

- a) Addition of replay button so the attempt videos can be re-watched.
- b) Addition of in-app camera to take photos of stimuli directly from the customisation menu to aid customisation.

4.2.5. Intervention Pilot Study

4.2.5.1. Participants

As planned, participants were drawn from the prior longitudinal study described in this thesis (Study 1). Participants were assessed at Time 4 of the longitudinal study (Visit 1 of the intervention study) and screened according to the following criteria:

- parent reported fewer than 10 sounds; or
- parent reported fewer than 20 words; or
- during observation at Time 4 (Visit 1), fewer than five words spoken

Participants were thus 19 minimally verbal autistic children (3 girls, 16 boys) aged 47 to 74 months (mean=60, sd=7) with a confirmed diagnosis of autism. The following exclusions applied at initial screening: epilepsy; known neurological, genetic, visual or hearing problems; English as an Additional Language. Children were initially recruited via social media, local charities, independent therapists and a university-run autism participant recruitment agency, and all took part in a larger longitudinal study (see Chapter 3).

Parents reported 17 participants to be White, one to be Asian and one to be Mixed Race. The formal education levels of the primary caregivers were distributed as follows: eight completed high school, eight completed university education and three completed post-graduate studies or equivalent. Eighty-eight per cent of

parents reported that their child had an Education Health and Care Plan, a legal document that specifies special educational support required for the child, at Visit 1.

Ethical approval was obtained from the UCL Research Ethics Committee (Project ID 9733/001) and informed consent was sought from parents on behalf of each participant.

Figure 14 describes the process through which participants were selected for the study.

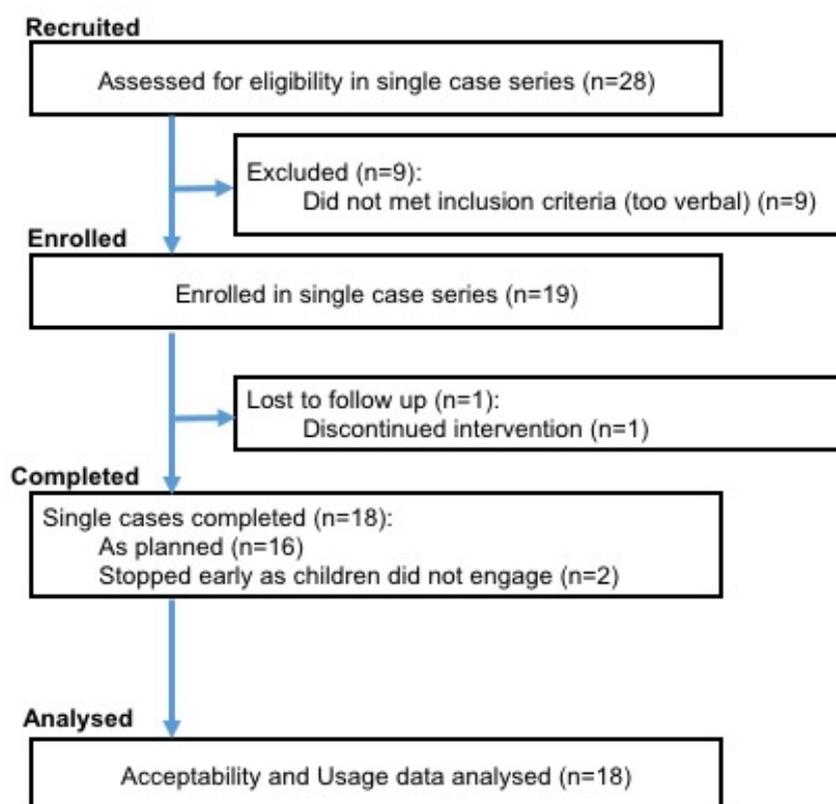


Figure 14. Recruitment flow chart for Study 2 feasibility analysis

4.2.5.2. Procedure

Children were visited in their homes in two sessions (Visit 1 and Visit 2), which were separated by four months each (mean=4.0, sd=0.3). Compensation of a small toy or £5 voucher was provided following each visit.

At **Visit 1**, each participant received a new Samsung Galaxy Tab A6 tablet containing the app⁵, unless parents expressed a preference to use the app on their own Android device (n=3). Parents were given a demonstration of the app by the experimenter, and an information pack explaining how to download and use the app. The Probe Phonemes were selected by following the 'Sound Target Protocol' (see Appendix I) and each parent-child dyad was informed of their randomly allocated intervention start date. Probe Phonemes are the nine sounds that are elicited each week in the baseline and intervention period. They also form the list from which an initial three target phonemes are drawn for the intervention. Probe Phonemes remained the same for each participant throughout the study, whereas the target phonemes for the intervention could vary over time according to specific mastery criteria. Probe Phonemes were not manipulated as part of the experiment, rather they were a necessary feature to accommodate the fact that each participant has a unique profile of speech related difficulties.

At **Visit 2**, parents completed a post-intervention questionnaire (See Appendix J) in order to objectively analyze the user experience of this intervention. It contains a grid of 10 questions regarding the usefulness and user-friendliness of the app, which can each score between one and four points, generating a score ranging between 10 and 40, with 40 representing the most positive rating of the app possible. Additionally, the questionnaire contained four open-ended questions regarding the strengths and weaknesses of the app.

At **both visits**, a battery of language-related measures were taken, some of which were designed as secondary outcome variables (Expressive Language by parent report, consonant inventory) and others related to a broader longitudinal study (of which these visits were Time-points 4 and 5). As part of this battery, at Visit 1 questionnaires on AAC use and educational placement were completed. All participants were free to take part in as much or as little additional therapy as they

⁵ One participant received a comparable second hand Nexus 7 tablet.

chose during the study, and this information was recorded via parent questionnaires at both visits.

Between Visits 1 and 2, text message reminders were sent to parents to remind them of the weekly obligation to complete the test module (a ‘probe day’), and if necessary missed probes were rearranged for the following day. On the intervention start date parents received a reminder text. Thereafter, parents were asked to play with the app for 5-10 minutes per day for five days a week. This resulted in children carrying out the intervention for between six and 13 weeks (8 possible outcomes, as illustrated in Figure 15).

Visit 1

Visit 2

A	A	A	B	B	B	B	B	B	B	B	B	B	B	B	B
A	A	A	A	B	B	B	B	B	B	B	B	B	B	B	B
A	A	A	A	A	B	B	B	B	B	B	B	B	B	B	B
A	A	A	A	A	A	B	B	B	B	B	B	B	B	B	B
A	A	A	A	A	A	A	B	B	B	B	B	B	B	B	B
A	A	A	A	A	A	A	A	B	B	B	B	B	B	B	B
A	A	A	A	A	A	A	A	A	B	B	B	B	B	B	B
A	A	A	A	A	A	A	A	A	A	B	B	B	B	B	B

Figure 15. All possible permutations of baseline (A) and intervention (B) weeks.

Throughout the baseline and intervention period, data from the app were uploaded regularly to a secure server accessed by the experimenter. These data comprised:

- a) information on time, type and duration of app usage (e.g. Participant 1 used the app for 35 seconds on test mode at 13:05 on 21/12/18);
- b) videos and parent ratings of probe trials each week; and
- c) videos and parent ratings of intervention trials during the intervention phase.

Additionally, as the participants are drawn from a previous longitudinal study (Study 1), further background measures, which were gathered between eight and 12 months prior to the beginning of this study, were also available to characterize the sample. A summary of all the relevant data collected across both studies is provided in Figure 16.

Data collection summary

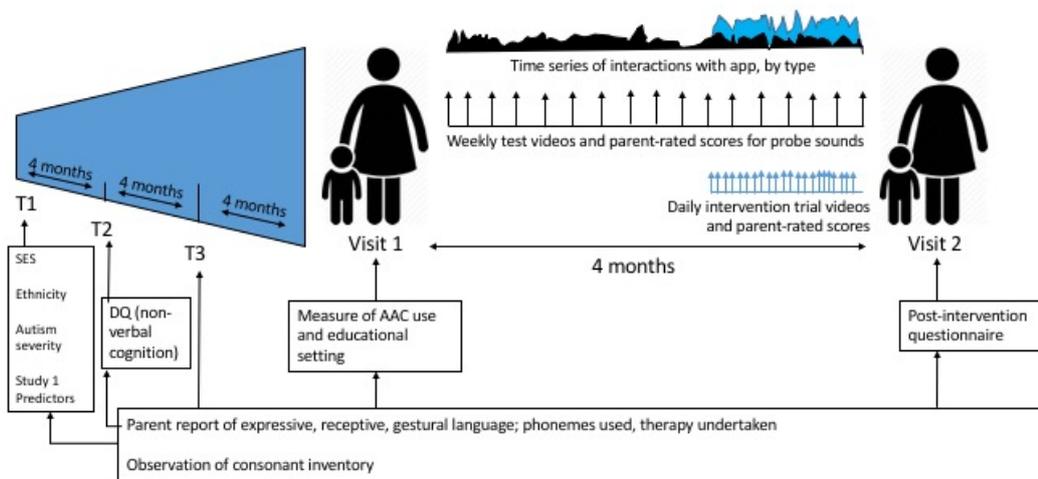


Figure 16. Summary of data collected from Study 1 and Study 2

Note: AAC: Augmentative and Alternative Communication; DQ: Developmental Quotient (developmental age/chronological age); SES: Socio-Economic Status

The accessibility analysis in this paper will focus on a) the post-intervention questionnaire and b) the information on time, type and duration of app usage. Analysis of efficacy measures such as weekly probes, pre- and post-intervention variables is provided in Chapter 5.

4.2.5.3. Analysis plan

Preregistered Questions

A feasibility trial is designed to answer the question ‘Can it work?’ rather than provide evidence of a treatment effect. One vital dimension of feasibility testing is

whether an intervention and its study procedures are suitable for and acceptable to the target population, which includes aspects such as retention, adherence and satisfaction (Orsmond & Cohn, 2015).

The analysis plan was pre-registered on Open Science Forum at <https://osf.io/9gvbs>, the acceptability and usability hypotheses were as follows:

1. **Acceptability:** More than half of the participants given the intervention will rate the app favourably via the feedback form, as defined by a score of more than 24/40 on a 10 question feedback form, where each question can be answered from one (not favourable) to four (highly favourable). This hypothesis will be tested by simply counting the proportion of parent-child dyads who score greater than 24 on the parent satisfaction measure.
2. **Usability:** More than half of participants given the intervention will comply with the intervention to a reasonable degree, as defined by an average of more than 12.5 minutes per week engaging with the app. This threshold was based on the instruction for parents to spend five minutes per day for five days a week using the app with their child. This totals an average of 25 minutes per week, which would define 100% compliance. Based on previous studies (Bagaiolo et al., 2017) I rated 12.5 minutes per week (50% compliance) to represent the lower threshold of 'reasonable compliance.' This hypothesis will be tested by counting the number of participants who spend a mean of > 12.5 minutes/week on the intervention, and dividing by the total number of participants.

Exploratory Analysis

Multiple other aspects of compliance were investigated using the available data to answer the following questions:

1. Did parents comply with the intervention schedule (i.e. did they begin the intervention on time)? This was evaluated by calculating the number of weeks of delay (versus scheduled intervention start) for each participant.
2. Did parents comply with the test schedule? This is important for the chosen intervention efficacy evaluation method and was evaluated by counting the proportion of planned weekly test trials that took place for each participant.
3. How many intervention trials per week did the parents do (i.e. did they spend five minutes per day completing just one trial)? This was a mean weekly trial count for each participant.
4. Given that the participants separated into 'high' users, who provided enough test data for analysis purposes, and 'low' users, who completed fewer than four weeks of test probes, we asked if any background variables or other factors could explain these groupings. This was evaluated via a series of t-tests comparing the values for each group.

Finally, written and verbal feedback regarding the app from parents was aggregated from various sources (texts, emails, written answers to open-ended questions on the App User Questionnaire, notes made by the first author from verbal comments made by parents at Visit 2 following the intervention. I analysed this data using thematic analysis (Braun & Clarke, 2006), in order to glean further qualitative information regarding acceptability and potential avenues for improvement.

4.3. Results

4.3.1. Acceptability: parent satisfaction

The acceptability questionnaire was completed by 89% of participants.

Table 8 outlines the key characteristics of each of the 19 participants and their acceptability score. Over half of participants were 'high' users (defined as providing greater than 66% of test trial data). Two participants dropped out during the trial, one was lost to follow-up and the remaining six participants engaged with the app but not enough to produce analyzable data (fewer than four weeks of test data and fewer than five intervention trials). Of these six participants, three did not find technology motivating; one had ongoing health problems; one had a technical issue with the app, and one made language progress via another therapy during the course of the intervention so did not engage with the app.

The pre-registered hypothesis that over half of participants would assign the app an acceptability score of over 24 was confirmed. In fact, participants gave the app a mean score of 29.5 out of 40 and only one participant rated it below 24. See Appendix K for score breakdown.

Table 8: Descriptive Statistics for whole sample

ID	Sex	Visit 1 Age (months)	Visit 1 RCDI (words)	Visit 1 ECDI (words)	User type	Device used	Acceptability score	Visit 1 Therapy hours per week: Total (SLT)
1	M	58.9	314	9	high	provided	35	9 (1)
2	M	56.4	38	0	high	provided	30.5	6 (0)
3	F	55.8	5	0	high	provided	29	39 (0)
4	M	59.5	282	1	high	provided	27	2 (1)
5	M	50.3	37	1	minimal	provided	29	0 (0)
6	M	62.5	103	6	d/o	provided	NA	0 (0)
7	M	60.4	290	0	high	provided	34	0 (0)
8	M	46.9	171	1	minimal	provided	22	1 (1)
9	M	54	NA	NA	l/f	own	NA	NA (NA)
10	M	58.7	55	0	high	provided	29	7 (1)
11	M	73.5	68	0	high	provided	35.5	2 (2)
12	M	53.6	212	19	high	provided	33	1 (1)
13	M	63.6	116	13	minimal	provided	25	0 (0)
14	M	72.8	404	8	light	provided	29	0 (0)
15	M	67.7	406	18	minimal	provided	29	1 (0)
16	F	68.3	245	0	d/o	own	26	0 (0)
17	M	55.1	189	0	light	provided	29	0 (0)
18	F	68.6	337	0	high	provided	35	0 (0)
19	M	61.8	8	5	high	own	25	0 (0)
Mean		60.4	182.2	4.5			29.5	3.7 (0.4)
Sd		7.3	137.9	6.3			10.0	8.9 (0.6)
Minimum		46.9	5	0			22	0 (0)
Maximum		73.5	406	19			35.5	39 (2)

Note: d/o: dropped out; l/f: lost to follow up; Acceptability score is out of 40; ECDI: Expressive Communicative Development Inventory; RCDI: Receptive Communicative Development Inventory

To supplement this result, I briefly summarise the qualitative feedback gathered via feedback forms and verbal comments made by parents at Visit 2.

The overall premise and main features of the app were well received. One parent wrote *“We are working on single sound production in ad hoc way and it is good to have a framework/system to focus us”*. Another wrote *“Previously have focused on whole words, this strips it back to a more basic skill, which I think is what we need.”*

Parents reported the app was quick to do, simple and accessible, facilitating practice little and often. Stimuli were clear and predictable, which parents felt was a strength. Parent quotes include that it was a *“short focused activity that we could fit into everyday life, simple and easy to use”* and *“my son enjoyed the predictability.”* Many parents reported that their children particularly liked the video-modelling stimuli.

Most parents reported that their children were specifically engaged by the mirror function (being able to watch themselves on screen during the speech trials and having the option to re-watch these videos, e.g. *“the selfie aspect was something I had not tried previously and my child responded well to it”*). One parent suggested that in a future version the videos could be side by side with the selfie screen during practice attempts. However, for a few parents, this feature was felt to have a negative impact in their child’s engagement with the app (one parent reported a more general issue that their daughter had with mirrors, and asked if the mirror function could be made optional). Another parent suggested masking the eyes as they felt their child did not like looking at their own eyes but would have found viewing the mouth useful

Feedback also highlighted five main areas that could improve acceptability:

1. **Better training and support with customization.** Although customization was a popular feature and reported to be easy to do, several parents said that they found it difficult to source the customized stimuli. This problem was compounded by the fact that most users were not using BabbleBooster on their ‘own’ phone and thus did not readily have access to their own photo library and usual apps. Several parents suggested in a future trial that I make a library of popular videos and images. One parent reported

frustration that the app could not interface directly with YouTube, since that was where all her child's videos were and it had the functionality that her child was used to (e.g. volume controls, fast forward buttons).

2. **Modifications to the test mode to make it more engaging.** The test mode did not have any built in reward videos and many parents said this made it difficult to engage their children, particularly as they were first exposed to only the test mode, during the baseline phase. Furthermore, several parents reported that the probe contained too many targets (nine) and three would have been more manageable.

3. **Solving technical problems and limitations.** Technical problems were an impediment to participation in that for a couple of users, the app would take a long time to initiate or would fail to login. This made it hard to plan therapy time, and left the child and parent feeling frustrated. A few users reported that crashes led to the need to re-customise the stimuli, which was time consuming. There were reports of recordings stopping mid-video, leading to a lack of reward, and frustration. From a research perspective, I estimate that some data has been lost through technical faults, preventing analysis of the full dataset. For 15% of the parent ratings submitted in the test phase, the accompanying video is missing due to technical problems. It is not possible to estimate how much more data may have been lost due to incomplete trials. In this trial it was not financially feasible to provide the app in OS (suitable for iphones/ipads) as well as Android format, and this had numerous disadvantages. Two children were motivated by technology but had aversion to the new device because it wasn't their usual one and did not have all the apps they had on the other one. The new device became linked with work/demands and thus not motivating.

4. **Introduction of more variety and visual interest in intervention presentation.** One aspect of this feedback is similar to point 3 above: parents requested that the app look and sound more game-like, e.g. tones, jingles, cartoons, glitter effects, colours, animations. A specific suggestion

was the incorporation of letters in the app itself (some members of this cohort had a special interest in letters). Some parents reported an initial high level of engagement and then a fatigue effect, which was revived once targets were substituted.

5. **Introduction of progress feedback for parents.** Many parents stated they would like quantitative progress feedback to encourage them to continue with the activities, e.g. % attempts that were correct, minutes spent using the app, progress towards goals.

In sum, a more 'game-like' app that can run on participants' own devices and is easier to customize with a wider range of personally relevant stimuli are key factors that would enhance user experience and could be implemented in future trials.

4.3.2. Usability

In order to consider whether the second pre-registered hypothesis can be confirmed, I have presented usage data in Table 9.

Table 9: Usability Data for 'high' user group

ID	Planned weeks of intervention	Delay to intervention start (weeks)	Actual weeks of intervention ⁶	% Weeks adhering to intervention	% Total test trials completed	Intervention trials / week	Mins / week during intervention	Mins / week during baseline
1	12	3	6	50%	69%	18	9.04	8.64
2*	10	3	5	50%	71%	0.6	8.72	6.25
5	7	0	4	57%	71%	9.8	11.61	6.38
6	8	0	7	88%	94%	3.7	3.88	2.04
7	11	1	7	64%	76%	6.4	5.32	2.21
10	12	0	9	75%	88%	15.9	35.85	14.47
12	9	1	7	78%	75%	14.1	13.65	2.92
14	7	3	3	43%	88%	76	20.74	8.06
15*	6	2	2	33%	100%	1	1.44	4.02
17	13	1	10	77%	83%	6.3	11.63	11.11
Mean	9.5	1.4	6.0	61%	82%	15.2	12.2	6.6
Sd	2.5	1.3	2.5	18%	11%	22.2	9.9	4.1
Min	6.0	0.0	2.0	33%	69%	0.6	1.4	2.0
Max	13.0	3.0	10.0	88%	100%	76.0	35.8	14.5

Participants marked with * completed very few trials in the intervention but were included in this group due to their adherence to the test schedule.

From this table it is apparent that only three participants used the app for longer than 12.5 minutes per week of intervention (fewer than half of the expected time on task), therefore the hypothesis is not confirmed. Time spent on intervention and number of trials per week of intervention were both highly variable.

Other usability metrics are also presented, notably that of the 'high' users, 82% of all test trials were completed, indicating that compliance with the test element of the trial was good in this subgroup. It was also not time consuming, taking a mean

⁶ Actual weeks = planned weeks – delay – missed weeks (where intervention not used at all for 1 week)

of 6.6 minutes to complete per week. Adherence to the intervention schedule was also reasonable at 61%, with a mean delay to starting of 1.4 weeks.

4.3.2.1. Characteristics of 'high' user group

Given that approximately half of participants engaged successfully with the app to some degree (n=10), and nine participants did not, I present an exploratory side by side analysis of group characteristics in Table 10, to determine whether background characteristics like family socio-economic status, or child symptom severity influenced use of the app. There were no significant group differences on any variable. This analysis also highlighted the limited amount of special clinical services these families were receiving, on average fewer than two hours per week. The total therapy hours/week variable is skewed by one high value (57 hours vs. mean of 1.3 hours/week for all other participants). If this outlier was removed the means would be even more comparable.

Table 10: Descriptive statistics describing the demographic features of the High vs. Low user groups

	'High' group					'Low' group					T	P-value
	n	mean	sd	min	max	n	mean	sd	Min	max		
Age at Visit 1 (months)	10	60.7	6.0	53.6	73.5	9	60.1	8.9	46.9	72.8	-0.17	0.87
Parent SES	10	1.8	0.7	1	3.5	9	1.6	1.0	1	3.5	-0.35	0.73
Autism symptom severity (-12 months)	10	43.0	5.1	35	49	9	42.3	4.9	37.5	52.5	-0.31	0.76
Non-verbal cognition DQ (-8 months)	10	0.4	0.1	0.13	0.56	9	0.3	0.1	0.18	0.52	-0.62	0.54
Receptive language at Visit 1 (words)	10	160.9	137.9	5	337	8	208.9	136.0	37	406	0.74	0.47
Expressive language at Visit 1 (words)	10	3.4	6.3	0	19	8	5.9	6.7	0	18	0.80	0.44
% AAC user	8	60%				8	60%				0.00	1.00
Total therapy (hours/ week (Visit 2))	10	7.0	17.5	0.0	56.6	8	1.1	2.1	0.0	6.0	-1.05	0.32
SLT therapy (hours/week) (Visit 2)	10	0.6	0.7	0.0	2.0	8	0.2	0.3	0.0	1.0	1.55	0.14

Note: AAC: Augmentative and Alternative Communication; DQ: Developmental Quotient (developmental age/chronological age); SES: Socio-Economic Status; SLT: Speech-language Therapy.

4.4. Discussion

This is an acceptable intervention as judged by the pre-registered analysis of post-intervention questionnaires. Qualitative results reveal strengths in the study and app design and areas for further development that largely focus on solving technical issues, ease of customization and gamification. This finding indicates that the app intervention has numerous features which parents and children liked and fostered engagement, and it has potential for future development.

Usage figures, however, presented a more mixed picture. Participants polarized into those who engaged with the app to a minimal degree (n=9) and those who adhered well to the test schedule (n=10). Of these 10 children, intervention usage figures were highly variable (in minutes spent and trials per week), but reasonable overall adherence to the intervention schedule was observed. Due to the unique features of this intervention, it is difficult to compare these usage figures to others in the literature. Adherence to parent-mediated autism interventions tends to be evaluated on criteria such as training session attendance (Dababanah & Parish, 2016), or how many learnt strategies are employed by parents at subsequent observations (e.g. Shire, Shih & Kasari, 2018). App-based autism interventions are usually designed to be independently accessed by users (e.g. Fletcher-Watson et al., 2015). The shared app-play here resembles more closely the off-line 'homework' allocated by speech-language therapists for children with speech sound disorders, however, the developmental and behavioural profile of children in this cohort is more challenging. A useful comparator is Stockwell et al. (2019) which describes a feasibility trial of an app-based parent-communication training for children with motor and communication disorders. Parents were required to upload and annotate videos of themselves interacting with their child at regular intervals, and received remote coaching on their use of key strategies. Attrition rate was 44% and participation levels fell short of the target (target=39 sessions, median=26, range=5-33). Parents reported that they found the intervention useful but cited time pressures and technical problems amongst reasons for lower engagement.

Many who found the app acceptable still dropped out or engaged only minimally, and it is important to consider the reasons why this happened.

This was a pilot study run on a minimal budget and consequently numerous practical challenges were encountered, particularly in resolving technical difficulties. Loss of data and frustration leading to avoidance also explains why some children and families were minimal users. I believe that these problems would be surmountable if a professional app company were to be engaged. This pilot has highlighted the importance of having enough memory capacity on the chosen device, which must be weighed up against cost considerations for any future trial. A lack of cross platform approach led to limitations in device choice, without doubt using participants' own phones would have been more effective. In a future trial I would strongly recommend a cross-platform approach, so that participants can use their own devices.

Thematic analysis of qualitative feedback received from parents highlighted several key areas for improvement. Parents reported that additional support with customization would be beneficial. Some of the difficulties stemmed from parents not using their usual devices for the intervention (due to the need to use a specific android device). This meant that they did not have direct access to their personal photo libraries. I had posited that with cloud based file repositories and online access to limitless content (e.g. YouTube or Google images), this would not be problematic, however it appears to have influenced the degree of customization which took place. The only official gauge of how much customisation occurred is by measuring minutes spent on the customising screen, however this gives a limited impression of how much the stimuli were changed. In a future trial I would recommend the data capture to give more detailed information of this, or that parents record their customisation activities in a diary format. In addition, a library feature and better interfaces with popular video sharing apps would enhance the usability for parents, although a preliminary feedback exercise may be necessary to ascertain which content would appeal to users.

A second area for improvement was the need for gamification, particularly with the test module, which parents found less engaging than the intervention. Overall the test phase was a stumbling block for many users, and could have been responsible for several of the participants dropping out or not adhering to the intervention. It comprised nine probes, which parents suggested was too many. This number was chosen to provide enough items to demonstrate any improvement within participant over time, incorporating trained and untrained items. It was designed to be a 'cold probe' in order to provide evidence of an improvement in the target skill if there was one, therefore no feedback was provided on speech attempts. An improvement would be to incorporate non-contingent rewards into the test phase, in order to ensure the children's first exposure to the app was associated with the fun aspects that appear later in the intervention phase (i.e. reward videos). As BabbleBooster was a low-budget prototype there was little scope to 'gamify' the app by incorporating visual effects such as spinning images and sound effects, but these may have also helped engage children from the outset.

Gamification of the intervention trials was also called for by parents. This would be a key area for refinement if this app is developed further. Among potential solutions are: to intersperse targets with mastered items (although this may have to be a non-speech task depending on the children's ability level) or to have more targets but rotate them frequently, perhaps at the syllable level so instead of just working on 'b', work on 'bee', 'boo' and 'bah.' The need to individualise and differentiate activities is common in app-assisted autism interventions (Powell, Waas, Erichsen & Leekam, 2016).

The final recommendation from parents was to incorporate a feedback mechanism, in order to inform and motivate families during the intervention. This was an initially planned feature that had to be disabled in the final version of the app due to cost constraints (a reworking of the app by the independent developer to resolve technical issues identified at Stage 2 had caused the feedback mechanism to stop working and no further funds were available to reinstate it). In future trials this should be reinstated.

Parents of autistic children experience higher levels of stress (Baker-Ericzen, Brookman-Frazee & Stahmer, 2005; Estes et al., 2013; Hayes & Watson, 2013) and thus fitting an additional therapy task into daily life could also be challenging. The occurrence of family illness, carer chronic health conditions, siblings with additional needs and difficult transition periods between educational settings and school holidays, were among the many barriers to adherence faced by this cohort. Parent-mediated speech and language therapy is often suggested, and digital tools such as BabbleBooster are designed to make this more feasible; however, we must be realistic that even this will be too much for some families, given their circumstances. Relatively few studies have analysed parent intervention adherence in families with a minimally verbal autistic child (e.g. Shire et al., 2014) and this is crucial to understanding what intervention approaches are likely to enhance parent co-operation.

Finally, we should recognise that app-based therapy is also not for everybody, and that is especially true in this very heterogeneous group. Three of the children's families reported that they were just not interested in technical devices. In these cases, future studies could consider whether the principles of the intervention design could be applied through different media. An improvement for future studies could be to evaluate technology use, familiarity and preferences among participants (parents and children) prior to an app trial. In the case of this study children were recruited from a larger longitudinal study using pre-registered inclusion criteria, which did not include information regarding technology preferences, although the information materials available to parents as part of the consent process did explain that the intervention would be app-based.

None of the background measures available in this study associated with 'high'/'low' group membership, although given small sample size this could reflect low power to detect effects. Future studies could seek to identify other factors which could be important predictors of intervention compliance in this population, such as parent physical and mental health, employment, confidence in using technology or delivering therapy.

Considering the challenges identified above, a future trial could improve the technological and motivational aspects of the app identified in Section 3.1. It is also possible that asking other significant adults (such as grandparents, learning support assistants, etc.) to use the app with the child in different settings could be an acceptable solution, in families where adherence to intervention is not feasible. More detailed predictions could be made regarding usability metrics, and data should be collected on customisation activities, in order to gauge whether degree of engagement with customisation was a factor in subsequent acceptability and usage scores.

4.4.1. Limitations

Like many feasibility studies I evaluated participant satisfaction using a bespoke questionnaire, tailored to the key components of the intervention (e.g. Mallet et al., 2016; Williams, Hastings & Hutchings, 2020). There are thus no appropriate benchmarks or norms available for our acceptability measure. Future studies could combine our highly informative bespoke measure with a commonly used generic intervention evaluation measure such as the Behavior Intervention Rating Scale (BIRS; Elliott & Treuting, 1991). The BIRS is a 24-item inventory using a 6-point Likert-type scale (“Strongly Disagree” to “Strongly Agree”) and addresses Acceptability and perceived Efficacy. Secondly, no fidelity measures were taken during the parent training aspect of this study (e.g. checklists of training topics or video coded analysis of parent training session). This was deemed unnecessary given the simplicity of the app and provision of a detailed manual, but it may be useful in a future study. Finally, thematic analysis of qualitative feedback from parents was evaluated subjectively by author. In future studies, an experimenter from outside the study could gather qualitative feedback using semi-structured interviews in order to reduce bias, and themes derived from transcribed interviews could be reviewed by another experimenter to ensure convergence.

The study aimed to incorporate user-centered design into the creation of the app, via a focus group (Consultation Stage 1). This phase did not lead to significant

changes to the app yet a wealth of proposed changes resulted from the pilot. This may suggest an ineffective consultation process, perhaps it was not done in sufficient depth, or at the right time in the design process. These aspects were constrained by the timeline and budget for the study. Some aspects in need of improvement (such as technical problems which only became apparent after several weeks of video downloading) could not have come to light until the app was in daily use.

4.4.2. Conclusion

This study reports a first attempt to develop and pilot a customisable app to develop speech production skills in minimally verbal autistic children, using video modelling and cued articulation to demonstrate where and how speech sounds were made, and video capture to record the child's production efforts. Overall, parents reported that a structured focus on improving speech skills was welcome, and reported that the app and the intervention design were acceptable. Nevertheless, parent compliance with the intervention schedule was highly variable and parents delivered about half of the recommended trials. While technical issues with software and device may explain some of this, the demands of family life may make parent-mediated interventions more challenging for this population. A better understanding of how best to facilitate engagement in therapies is a priority for future research.

5. Study 2: Intervention Efficacy

A randomized case series approach to testing efficacy of a speech production intervention for minimally verbal autistic children

5.1. Introduction

In Chapter 2 I reviewed intervention research describing trials whose primary aim was to improve expressive or spoken language in autistic participants, with a focus on those who were minimally verbal. One key finding was the lack of robustly designed and adequately powered studies (e.g. Brignell et al., 2018). In order to improve the evidence base for language interventions in autism for minimally verbal children, high quality studies are required, yet the financial and logistical challenges of recruiting a large enough sample can be prohibitive, and at an early stage of design a large RCT may not be appropriate (Craig et al., 2006). In this chapter I describe an alternative study design suitable for smaller samples, the randomized case series, and illustrate its use and analysis with data collected as part of the BabbleBooster feasibility trial described in Chapter 4.

The Randomized Controlled Trial (RCT), in which a large group of participants is randomly allocated either to receive the treatment or to a control condition, is considered the gold standard method with which to evaluate the efficacy of intervention trials (Sibbald & Roland, 1998; Kendall, 2003). Despite widespread adoption of RCTs in neurodevelopmental conditions, certain circumstances can make implementing an RCT difficult: the target population may be rare, difficult to recruit in sufficient numbers, and/or extremely heterogeneous (e.g. individual targets may need to vary by participant). RCTs are also costly to implement, and thus only appropriate once an advanced stage of intervention development has been reached, following the incorporation of prior rounds of piloting and feedback (Craig et al., 2006).

An additional pitfall of any between-subject design such as RCTs, is their reliance on single time-point measurements of pre- and post-intervention performance. This requires the same outcome to be measured on only two occasions and compared. In an emerging skill, or for a population with highly variable test performance due to attentional or behavioral factors, this method risks over- or underestimating an effect. The assumption that grouping participants at random will 'equal out' this variance may only be true in participants with a homogenous profile, which is rarely the case in neurodevelopmental conditions. Dense sampling, in which there is repeated assessment of the outcome measure both before and during the intervention, may provide a more robust measurement method in populations with high heterogeneity or where individual differences are of special interest (Wilson, 2011).

5.1.1. Single Case Design

An alternative experimental design which incorporates repeated measurements, is the Single Case Design, in which each participant serves as their own control and multiple measurements are taken across at least two experimental phases, usually baseline and intervention. Single Case Designs may be a viable alternative when RCTs are not feasible (Kourea & Lo, 2016). Single Case Designs come in many formats, predominantly either a phase design, where baseline and intervention measurement occasions are grouped together in sequential blocks, or an alternating design, where intervention and baseline sessions are interspersed. Features of the intervention usually guide design choice: alternating designs are best suited to interventions that work only while they are ongoing and do not have a lasting effect (e.g. tick chart for target behaviour in class), whereas phase designs suit interventions where skills are built up and retained over time.

Single Case Designs are a widely accepted source of evidence in a number of fields such as education (Shadish, Hedges, Horner & Odom, 2015), medicine (Vohra, 2016) and psychology (Kazdin, 2019). Despite the advantages of being low-cost, easy to implement and extremely flexible, Single Case Designs have been historically viewed as methodologically inferior (Concato, Shah & Horowitz,

2000). One reason for this is the lack of statistical tests available to evaluate their results, since Single Case Designs violate the parametric assumptions of independence of observations and random sampling from the normal distribution. Single Case Designs were thus historically analysed by visual inspection alone, in which observations of the outcome variable are graphed over time and aspects such as level, trend and variability are compared between experimental conditions. This approach incorporates the richness of the data whilst remaining simple and accessible (Heyvaert, Wendt, Van den Noortgate & Onghena, 2015). However, the lack of objective decision guidelines leaves this approach vulnerable to bias and inconsistency between researchers (Matyas & Greenwood, 1990; Parsonson & Baer, 1992; Ninci, Vannest, Willson & Zhang, 2015).

There has been a renewed interest in Single Case Designs, based on numerous innovative quantitative approaches to their analysis, which go beyond visual inspection (Manolov & Moeyaert, 2017). New methods enable researchers to use Single Case Designs to robustly test functional relationships between interventions and outcomes, and to compute effect sizes for cross-study comparison and inclusion into meta-analyses.

A growing recognition of the value of Single Case Design when these analytic approaches are incorporated, has led to new standards being established for Single Case Designs (Shamseer et al., 2015; Vohra et al., 2015; Tate et al., 2016). Replication of effects is crucial (Kratochwill et al., 2010; Horner et al., 2005), and can be achieved in various ways. For instance, using a single participant with three different exposures to or withdrawals of an intervention (ABAB design), or using three participants who each begin an AB phase intervention at staggered start time-points (multiple baseline design).

An array of books, special journal issues, tutorials and simulations have been published in the past decade, all proffering new ways to statistically analyse Single Case Designs (see summary in Manolov & Moeyaert, 2017). A clear standard approach is yet to emerge. Furthermore, despite the heavy output of methods papers, published studies employing any of these methods are still rare. The

randomization test (described below) is one innovative approach that has been employed in several Single Case Designs (Hoogeboom et al., 2012, Schulte & Walach, 2006; Hwang, Levin & Johnson, 2018; Wenman et al., 2003; Calet, Pérez-Morenilla & De los Santos-Roig, 2019; Alfonsson, Englund & Parling, 2019). The between-case standardized effect size (described below) has been used in a recent meta-analysis (Barton, Pustejovsky, Maggin & Reichow, 2017). To our knowledge, a practical application that combines these methods has not yet been carried out to evaluate interventions in autistic populations.

Development of functional speech by age five is one of the strongest predictors of positive outcome in autism (Howlin, 2005; Szatmari et al., 2003), yet 14-29% of autistic individuals remain minimally verbal. A recent Cochrane review highlighted the dearth of high quality evidence in language interventions for minimally verbal autistic individuals (Brignell et al., 2018). Like previous reviews of spoken language interventions in autism (e.g. Hampton & Kaiser, 2016), it focused solely on evidence from group studies. Systematic reviews incorporating Single Case Design evidence have either been unable to generate an effect size at all (Lane et al., 2016, Mulhern et al., 2017) or used the Percentage of Non-overlap statistic (Kane et al., 2010), which is considered limited due to ceiling effects (Parker, Vannest, Davis & Sauber, 2011) and confounds with baseline length (Allison & Gorman, 1993). Lane et al. (2016) assessed naturalistic spoken language interventions in autism for methodological quality and found only half the Single Case Design studies (24 studies, n=45) were of adequate quality. In summary, robust analysis measures and quality standards are still sorely lacking in the Single Case Designs describing language interventions in autism.

The goal of this paper is to demonstrate a practical application of two of the commonly cited metrics for statistical analysis of Single Case Designs: (1) the randomisation test, and its subsequent pooling across participants, and (2) a standardised effect size accounting for between-participant variance Between-Case Effect Size (BCES). These metrics are complementary to and independent of one another. A thorough comparison between these and other potential methods are beyond the scope of this Chapter, however I will briefly describe them, explain

why they were chosen, and address common criticisms of these methods. An in-depth mathematical and theoretical explanation of why these methods are appropriate can be found in Shadish et al. (2014), Shadish, Hedges and Pustejovsky (2014), Shadish, Zuor and Sullivan (2014) and Hooton (1991).

5.1.2. Randomization Test

Randomization is a cornerstone of good experimental design as it reduces extraneous confounds and increases internal validity (Barton, 2006). Single Case Designs can also incorporate random assignment, and functional relationships can subsequently be statistically tested via the Randomization Test devised by Fischer in 1935 (Rvachew & Matthew, 2017; Kratochwill & Levin, 2010). This is done by randomly selecting the intervention schedule for a given Single Case Design from a pre-determined number of permissible schedules. The scope of this random assignment varies by Single Case Design type: in an alternating design, intervention allocation can be completely randomized (e.g. producing the sequence ABBABABBBBAABA, where A=baseline measurement occasion and B=intervention measurement occasion), whereas in a phase design the baseline and intervention measurement occasions must be grouped together in phases (e.g. AAAAAABBBBBBBBBB). The number of permutations from which the allocated schedule is chosen will vary by design type, number of measurement occasions and any further constraints (e.g. a minimum baseline period before intervention is introduced in a phase design).

So long as the intervention schedule was randomly allocated from a number of possible permutations, a Randomization Test can be performed by computing a test statistic (e.g. the mean difference score of A versus B occasions) for each permissible permutation, via resampling. This yields a distribution of all possible test statistics given the actual data. The test statistic from the allocated schedule is then compared to this distribution to derive a p-value, which suggests the likelihood of this test statistic occurring by chance under the Null Hypothesis (assuming no intervention effect). A visual example with hypothetical numbers is

provided in Figures 17 and 18 to illustrate a Randomization Test for a simple AB phase design.

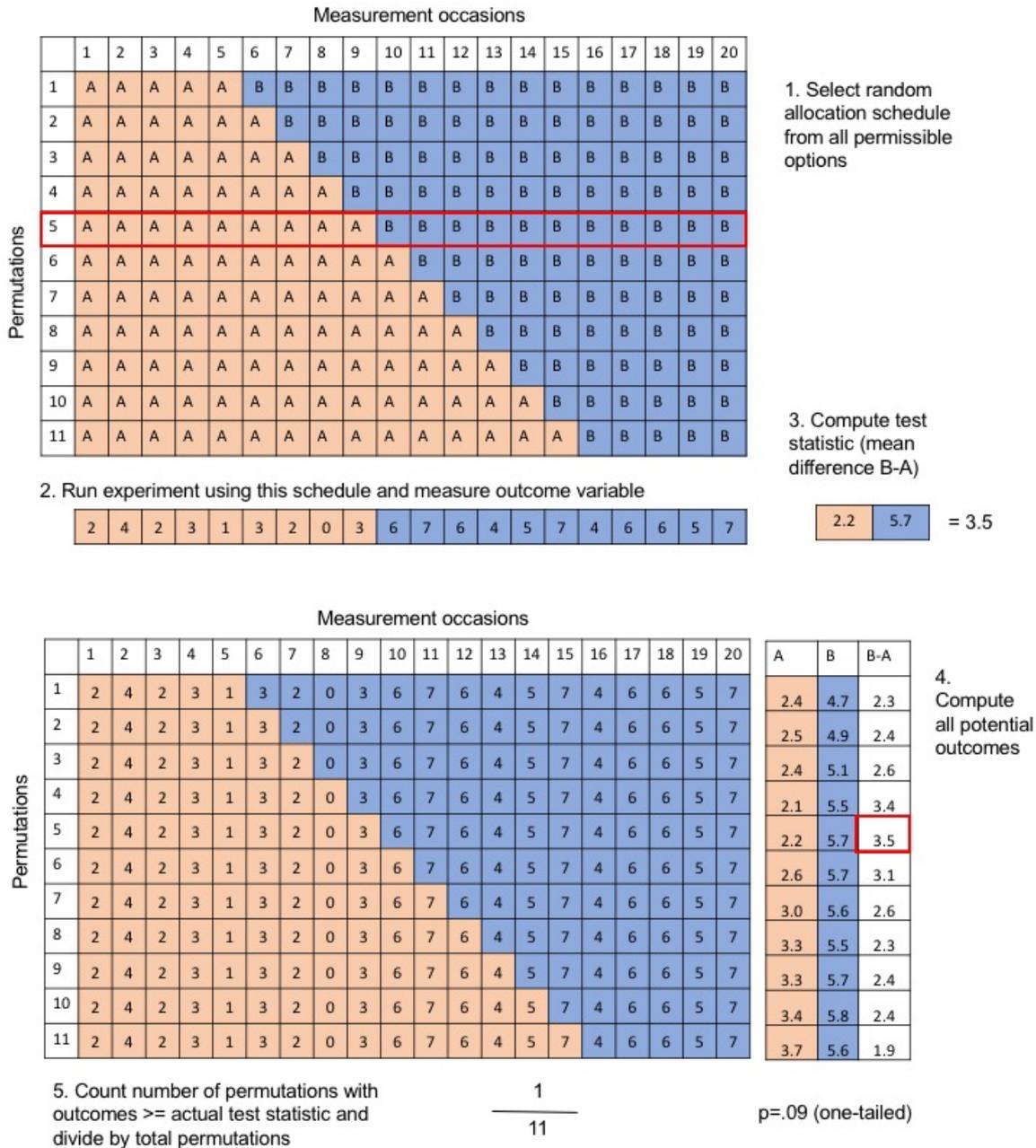


Figure 17. Steps needed to calculate a Randomisation Test

1) Random selection of intervention schedule; 2) repeated measurement of outcome variable; 3) calculation of mean difference between intervention and baseline scores; 4) compute all potential mean differences (one for each permissible intervention schedule); 5) compare the actual mean difference with all possible outcomes to obtain a rank, e.g. the greatest mean difference out of 11 possibilities, which corresponds with a p-value of 1/11 or .09.

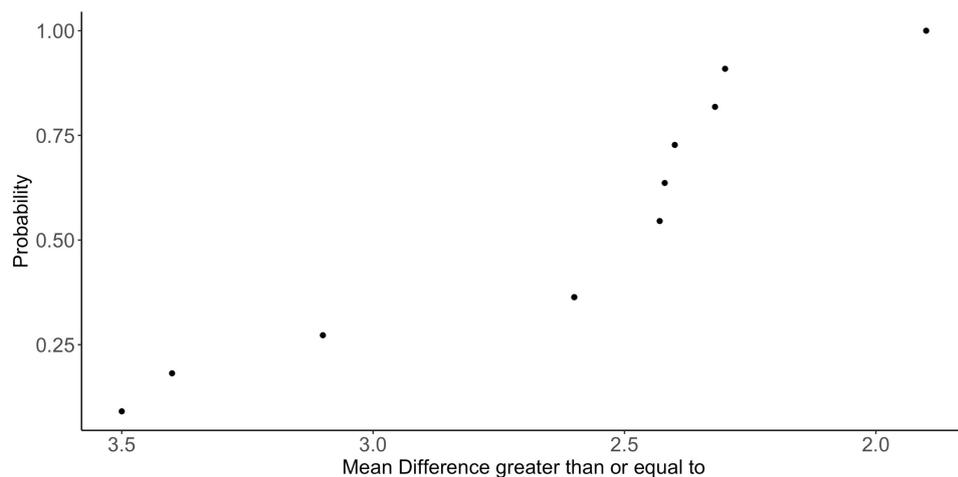


Figure 18. Probability distribution of all possible mean differences.

Plots the mean difference for each of 11 permissible permutations in rank order, against the likelihood of the mean difference being at least as great, e.g. all mean differences are greater than 1.9, $p=1$ that any of the 11 selected at random will be at least 1.9. Only 2 are greater than or equal to 3.4, therefore the associated $p=2/11$ or 0.18.

The Randomization Test can enhance the power to detect a significant effect, particularly when combined with replication via multiple baseline (Ferron & Sentovich, 2002). Conceptually, random assignment strengthens internal validity by counteracting the threats of maturation and history (Heyvaert et al., 2015). The Randomization Test is not linked to a specific test statistic so if the mean difference is not appropriate, there is flexibility to use a different metric, such as the Immediate Treatment Effect Index (ITEI), which compares the last three datapoints of the baseline phase with the first three datapoints of the intervention phase (Michiels, Heyvaert, Meulders & Onghena, 2017). As a nonparametric test, the Randomization Test is robust to violations of certain assumptions that are difficult to meet in Single Case Design research, namely independence of observations and random sampling from a normal distribution (Hooton, 1991). Single Case Design observations usually have a degree of serial dependency, or autocorrelation, and can display trends (Solomon, 2014). The Randomization Test can accommodate linear trend better than a group design (Michiels & Onghena, 2018).

Despite these advantages, randomization remains rare in Single Case Designs (Heyvaert et al., 2015). One criticism is that the Randomization Test's power to detect an effect diminishes in the presence of certain non-linear trends such as a delayed intervention effect, a learning curve or an extinction burst (Wilson, 2011; Sierra, Solanas & Quera, 2005, Levin, Ferron & Gafurov, 2017). Another reason the Randomization Test can be perceived as problematic is because random-assignment of intervention start point not always possible or desirable. The pre-determined introduction point of an intervention is at odds with response-guided experimentation (Kazdin 1980), and can be challenging if it is not known how long a stable baseline will take to achieve. Rvachew & Matthews (2017) also highlight the ethical dilemma of potentially giving some participants a very long baseline with many repeated measurement obligations prior to receiving the intervention. However, each participant does receive some exposure to both conditions, unlike in an RCT where participants may be assigned to the control group and not receive any of the intervention.

As is evident from the example in Figure 17, if there are only 11 possible permutations for a given participant, the lowest achievable p-value for a single case design is .09, or 1/11, assuming a one-tailed analysis. A single AB phase Single Case Design alone is unlikely to have adequate power (Michiels & Onghena, 2018; Haardörfer & Gagné, 2010). Ways to increase power include increasing the number of measurement occasions, or replicating the result by pooling results across participants.

P-values derived from individual Randomization Tests can be pooled across participants in a case series or multiple baseline design, to determine the likelihood of these p-values occurring by chance, using Stouffer's Z statistic (Rvachew & Matthews, 2017).

" a very small experiment can be replicated across subjects and probability values pooled even when the individual P values are greater than 0.05" (From Rvachew & Williams, 2017, p.5)

Alternatively, a single Randomization Test can be derived for the case series as a whole, however, there is a lack of clarity on the optimal way to execute random assignment in this method (Ferron & Sentovich, 2002, Levin, Ferron & Gafurov, 2017, 2018).

5.1.3. Effect sizes

The Randomization Test tests the significance of a functional relationship between the intervention and a change in the outcome variable, but does not inform us as to the magnitude or variability of this effect. Effect sizes not only convey this important information, but due to their standardization, enable the comparison of effects across studies. Effect sizes are increasingly considered important for interpreting intervention results and determining evidence based practice (Wilkinson & the Statistical Inference Taskforce, 1999). RCTs have an established standardised effect size, Cohen's *d* (Cohen, 1977), which can be adjusted to Hedges *g* (Hedges, 1981) for small samples. The unit of comparison is standard deviations of outcome variable. Effect sizes historically developed for Single Case Designs cannot be standardized in the same way and do not account for between participant variance, in the way that Cohen's *d* does in a group study (see Odom, Barton, Reichow, Swaminathan & Pustejovsky, 2018 for a summary of previous approaches and their failings).

The importance of determining a robust effect size for Single Case Designs is increasingly recognised (Shadish et al., 2015), as few Single Case Designs currently report effect sizes or their variances (Jamshidi et al., 2018). Many effect size metrics have been proposed (Manolov & Moeyaert, 2017), yet there is no common consensus on the best approach. Approaches using regression coefficients as effect sizes have also been devised (Moeyaert, Ferron, Beretvas, & Van Den Noortgate, 2014; Shadish, Zuur & Sullivan, 2014). These are able to account for linear or nonlinear trends in the data as well as for dependent error structures, however they are more complicated to implement and interpret, when compared to mean difference based approaches (Heyvaert et al., 2015). Other approaches have been developed and tested using a Bayesian framework (Jones,

2003; Swaminathan, Rogers & Horner, 2014; de Vries, Hartogs & Morey, 2015, Odom et al., 2018), however implementation is similarly complex. Non-parametric approaches have been proposed such as the Randomization Test Inversion, which exploits the equivalence between a hypothesis test and a Confidence Interval to create an effect size based on the Randomization Test (Michiels et al., 2017), but this is yet to be robustly tested. Tau-U, based on the tradition of examining non-overlap between experimental conditions, combines existing non-parametric tests Mann-Whitney U and the Kendall Rank Correlation coefficient (Parker et al., 2011).

5.1.4. Between-Case Effect Size

In the current study I focus on the Between-Case Effect Size (BCES) devised by Shadish, Pustejovsky and colleagues (Hedges, Pustejovsky & Shadish, 2012, 2013; Pustejovsky, Hedges & Shadish, 2014), illustrated in Figure 19. The BCES is easy to interpret, has been tested in simulations (Hedges et al., 2012), meta-analyses (Barton et al., 2017), tests of practical applicability (Odom et al., 2018) and comparisons with other approaches (Odom et al., 2018; Shadish, Rindskopf and Boyajian, 2016). It is accessible to non-statisticians, given the straightforward conceptualisation (based on Cohen's d) and the availability of several R packages (Bulte & Onghena, 2009, 2019) and primers (Hedges et al., 2012, 2013; Valentine, Tanner-Smith, Pustejovsky & Lau, 2016) to aid calculation.

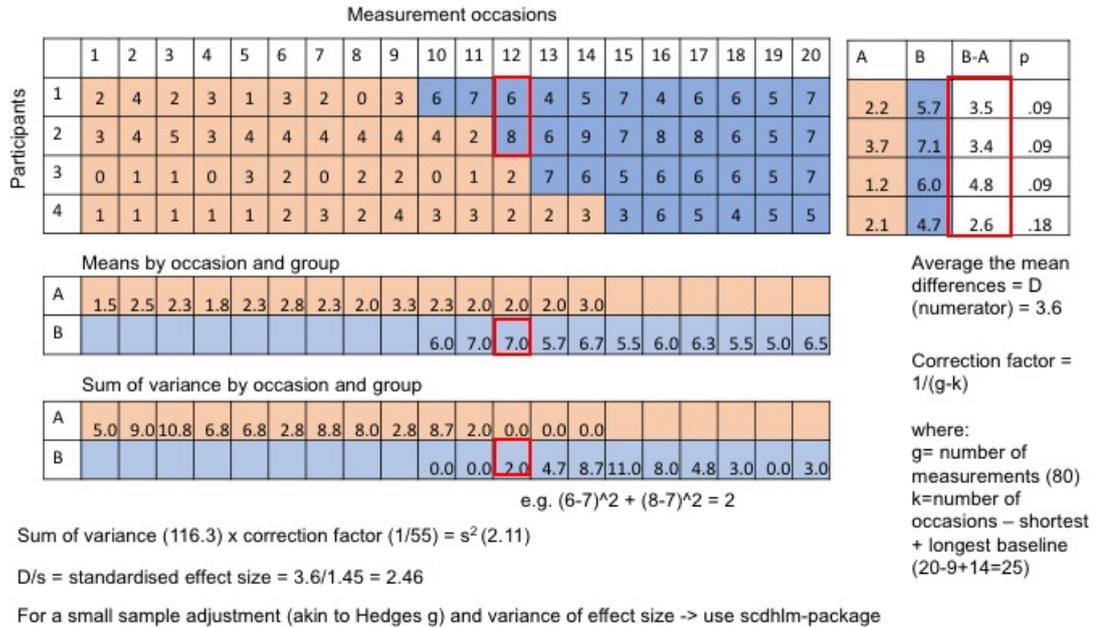


Figure 19. Calculation of unadjusted Between-Case Effect Size

For each measurement occasion, group scores into occasion type (baseline or intervention) and calculate variance; sum all the variances and multiply by a correction factor; take the square root to calculate the denominator (s); numerator is the average mean difference across participants (D); effect size is D/s.

5.1.5. Aims

In the current study, I use an AB phase Single Case Design to evaluate a parent-mediated app-based speech production intervention for minimally verbal autistic preschoolers (n=19). I describe the methods, analysis, and pitfalls to implementing this approach in a population of children that is difficult to recruit and have highly variable patterns of language growth (Chapter 3).

To my knowledge, random assignment and between-case effect size analysis have not previously been applied to a Single Case Design targeting expressive language growth in minimally verbal autistic children. Given the difficulty to recruit this highly variable target population, a Single Case Design was chosen. Single phase was considered the most appropriate format (rather than phase reversal or alternating), since the aim of the intervention is to teach speech sound skills, which once acquired should remain part of the child's speech sound repertoire. Employing an app-based intervention facilitated remote, repeated sampling of the outcome

measure, which is a core component of Single Case Design. Indeed, the practicality of repeated sampling, and the ability to introduce blinding or independent validation into this process is a key challenge in Single Case Designs (T. Smith et al., 2007), which using an app can overcome.

The overarching goal of this Chapter was to illustrate how Single Case Designs with random-assignment can be used to evaluate an app-based intervention for minimally verbal autistic children, delivered by parents. The analysed outcome measure was weekly performance on an in-app test of trained and untrained speech sounds.

Key goals of this Chapter were:

1. To establish whether parents could effectively rate their child's speech production attempts in order to facilitate remote dense sampling using the app.
2. To investigate whether the resulting ratings could be used to compute a randomization test statistic.
3. To compute a between-case effect size to summarise the intervention outcome data.

I also considered what else the outcome data could tell us about performance on the intervention, specifically:

1. Will there be a demonstrable link between those sounds that were targeted and any increase in speech skills (providing evidence that improvements are specific and not simply the result of maturation by chance)? If so, a Chi squared analysis would reveal a significantly greater proportion of successfully elicited trained vs. untrained sounds in the final week of the intervention.

2. Will improvements be limited to those sounds being targeted via the intervention, or will improvements in trained sounds generalise to improvements in untrained sounds? Generalization would be reflected by a significant positive correlation between number of trained and untrained sounds successfully elicited in the final week.

Finally, I considered how to combine post-test measures taken immediately after the intervention, with extensive data I had gathered in the 12-months prior to the beginning of the Single Case Design experiment. My question was will the intervention lead to a significant increase, relative to the long pre-intervention baseline period, in any measurable broader speech related variables in participants, such as number of consonants occurring in a natural speech sample, or parent-reported expressive language. All objectives and hypotheses listed above were pre-registered⁷ (<https://osf.io/9gvbs>).

⁷ In light of non-significant main findings, the final section of pre-registered analyses were not carried out, as these sought to identify potential moderators of success.

5.2. Methods

5.2.1. Study Design

The study utilized an AB phase design with randomized baseline allocation. Each participant underwent the intervention, however the number of weeks of baseline testing (A weeks) and the number of weeks of subsequent intervention (B weeks), were determined randomly for each participant.

Constraints on randomisation were as follows:

- each participant had a minimum of 3 weeks baseline (A) weeks
- each participant had a minimum of 6 weeks intervention (B) weeks

These constraints were determined due to the limited timeframe available for the intervention (16 weeks), and prioritising intervention weeks whilst retaining a long enough minimum amount of A weeks for a baseline to be established (Kratochwill et al., 2010).

From these constraints we can say that each participant followed a profile outlined in Figure 20:

Visit 1										Visit 2					
A	A	A	4	5	6	7	8	9	10	B	B	B	B	B	B

Figure 20. Intervention schedule for Study 2.

Shaded numbers indicate weeks that could be allocated to A or B weeks, depending on the randomly selected start week.

There are thus eight possible intervention schedules, as illustrated in Figure 15.

5.2.2. Intervention

BabbleBooster was designed to deliver predictable and repetitive speech models via video-modelling and cued articulation. The app-play is parent-mediated, so parents are required to watch the stimuli with their children, encourage them to make the sound, and then provide feedback on the accuracy of the production attempt in order to trigger the reward videos. Reward videos were designed with a gradient response, so a 'good try' at a sound (an incorrect attempt) will result in a lesser reward than an accurate response. The families were encouraged to make or upload their own reward videos, based on their understanding of the individual child's specific motivators. Acceptability data and discussion of the genesis of the app prototype is discussed in Chapter 4.

5.2.2.1. Participants

Participants are as previously described in the feasibility study in Chapter 4. They were assessed at Time 4 of the longitudinal study (Visit 1 of the intervention study) and screened according to the following criteria:

- parent reported fewer than 10 sounds; or
- parent reported fewer than 20 words; or
- during observation at Time 4 (Visit 1), fewer than 5 words spoken

Participants were thus 19 minimally verbal autistic children (3 girls, 16 boys) aged 47 to 74 months (mean=60, sd=7) with a confirmed diagnosis of autism. The following exclusions applied at initial screening: epilepsy; known neurological, genetic, visual or hearing problems; English as an Additional Language. Children were initially recruited via social media, local charities, independent therapists and a university-run autism participant recruitment agency, and all took part in a larger longitudinal study (see Chapter 3).

Parents reported 17 participants to be White, one to be Asian and one to be Mixed Race. The formal education levels of the primary caregivers were distributed as follows: eight completed high school, eight completed university education and

three completed post-graduate studies or equivalent. 88% of parents reported that their child had an Education Health and Care Plan, a legal document that specifies special educational support required for the child, at Visit 1.

Ethical approval was obtained from the UCL Research Ethics Committee (Project ID 9733/001) and informed consent was sought from parents on behalf of each participant.

Simulations suggested adequate power (see Appendix L).

Figure 21 describes the process through which participants were selected for the study.

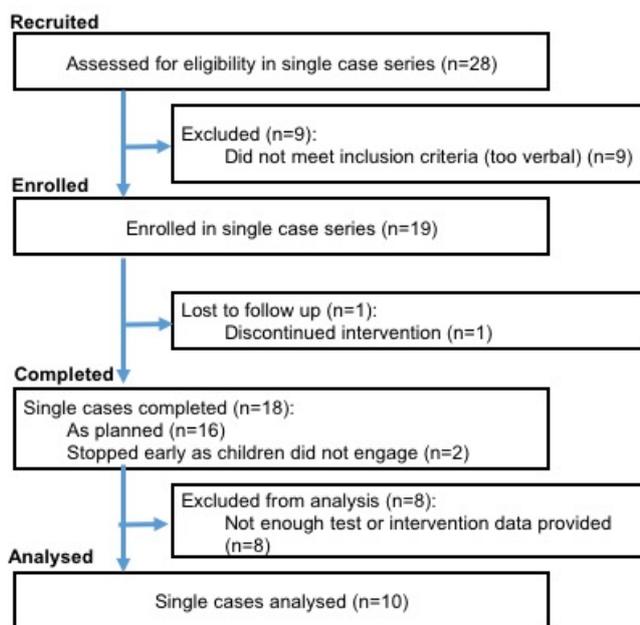


Figure 21. Recruitment flow chart for Study 2 efficacy analysis

5.2.2.2. Procedure

Data collection

The data collection procedure is as described in Chapter 4 and summarised in Figure 16, with further description of individual variables in Table 1.

Children were visited in their homes in two sessions (Visit 1 and Visit 2), which were separated by 4 months each (mean=4.0, sd=0.3). Compensation of a small toy or £5 voucher was provided following each visit.

At **Visit 1**, each participant received a new Samsung Galaxy Tab A6 tablet containing the BabbleBooster app⁸, unless parents expressed a preference to use the app on their own Android device (n=3). Parents were given a demonstration of the app by the experimenter, and an information pack explaining how to download and use the app. Secondly, the Probe Phonemes were selected by following the 'Sound Target Protocol' (see Appendix M) and each parent-child dyad was informed of their randomly allocated intervention start date. Probe Phonemes are the nine sounds that were elicited each week in the baseline and intervention period (outcome variable). They also formed the list from which an initial three target phonemes were drawn for the intervention. Probe Phonemes remained the same for each participant and were not manipulated as part of the experiment, rather they were a necessary feature to accommodate the fact that each participant had a unique profile of speech related difficulties.

At **both Visit 1 and 2**, a battery of language-related measures was completed, some of which were designated as secondary outcome variables (Expressive Language by parent report, consonant inventory) and others related to a broader longitudinal study (of which these visits were Time-points 4 and 5). Questionnaires on AAC use and educational placement were completed at Visit 1 (see Appendix M). All participants were free to take part in additional therapy as usual during the study period, and details about therapy provision was recorded via parent questionnaires at both visits (Appendix E).

Between Visits 1 and 2, text message reminders were sent to parents to remind them of the weekly probe day, and if necessary, missed probes were rearranged for the following day. Parents also received a reminder text on the intervention start date. Thereafter, parents were asked to engage their child in play with the

⁸ One participant received a comparable second hand Nexus 7 tablet.

app for 5-10 minutes per day for 5 days a week. This resulted in children carrying out the intervention for between 6 and 13 weeks (8 possible allocations, as illustrated in Figure 15). Throughout the baseline and intervention period, for each test attempt, all pertinent information was uploaded to the server (date stamp, phoneme, attempt number, parent rating (either "correct", "incorrect attempt" or "no attempt") and a video clip of the attempt. Parents pressed one of three buttons to assign a rating to the attempt, in accordance with Table 7 (Chapter 4).

Data collected prior to Visit 1: As the participants were drawn from a previous longitudinal study (see Chapter 3), further background measures, which were gathered between 8 and 12 months prior to the current study, were also available to characterize the sample.

Table 13 displays descriptive variables for the intervention group.

5.2.2.3. Measures

Primary outcome measure: elicited phoneme weekly score

Each child received a probe score out of 9 for each of the 16 weeks between Visit 1 and Visit 2, which was used to generate a mean baseline probe score and a mean intervention probe score, as well as the mean difference between these two measures.

Missing data

In the pre-registered analysis, I planned to impute missing data following Enders (2010), transforming metrics before imputation if necessary (von Hippel, 2009). I planned to use 40 imputed data sets in order to minimize bias in parameter estimates (Graham, 2009). For the primary outcome variable, I made a distinction between participants who did not engage meaningfully with the testing regime and those who did ('high-users' who provided more than 66% of data). Results were reported for high-users only, both on the basis of the incomplete dataset and

pooled estimates from 40 multiply imputed datasets, created using the Amelia package in R (Honaker, King & Blackwell, 2011).

Reliability

The recorded weekly score is derived from parent ratings. To assess reliability of parent scores 20% of the probes were coded by a qualified Speech and Language Therapist, who was not involved in the study, and was blind to the intervention targets and individual assessment point.

To calculate the reliability of the parent ratings, I derived a list of the filenames of all available video clips downloaded from the BabbleBooster server for the 10 analysed participants (n=1,120). This number did not correspond with the total number of parent ratings (n=1,248) due to the loss of some videos due to technical problems with the devices used. For coding purposes, data from incomplete weeks were also removed (n=113). Videos were not selected completely at random: the sample needed to include at least 2 complete weeks of data for each user (n=214 videos) since the variable I was comparing across raters was the weekly score. Weeks were chosen at random from the available weeks and comprised at least one A and a B week⁹. For each video clip, the coder was told which sound the child was attempting and told to rate '0, 0.5 or 1' to correspond with no attempt, incorrect attempt or correct attempt.

This process generated two to three randomly selected weekly scores for each of the 10 'high use' participants, which were used to compute an Intra-class correlation coefficient to denote inter-rater agreement. This was computed intra-class correlation ICC() command in the psych R package (Revelle, 2018). An agreement of .85 or higher was considered an acceptable level of agreement (Koo & Li, 2016, suggest >.75 represents good agreement).

⁹ Not possible for one participant due to technical problems with uploading in initial weeks

Secondary outcome measures

Two secondary efficacy measures, consonant inventory and expressive language, were selected to reveal any broader changes in speech production skills and/or expressive language between Visit 1 and Visit 2. There were very few missing data points for these variables (just one consonant inventory variable at Visit 2 due to the child being ill on the assessment day) and thus no multiple imputation was required. Unlike the primary outcome measure, data was available for the whole group (intention to treat basis), so analyses were performed for this group (n=18) as well as the 'high user' group (n=10).

Consonant inventory

Consonant inventory was obtained from a natural speech sample recorded during a play interaction with the experimenter at each Visit. All speech samples were coded in ELAN software, in accordance with CSBS guidelines described in Appendix A and as described further in Section 3.2.3. Consonant inventory was also available for each child for three time points prior to Visit 1, spaced four months apart, due to their inclusion in a prior longitudinal study (Chapter 3). This enables an additional comparison of the average change over three 4-month periods in the 12 months prior to the intervention, compared with the average change in the 4-month period during which the intervention took place.

Expressive language

Expressive language was measured by parent report using the raw score from the Oxford CDI (Hamilton et al., 2000) at Visit 1 and Visit 2. This represents the parent-estimated size of the child's expressive vocabulary. As above, data are also available for each child for three previous time points spaced four months apart, enabling us to evaluate progress during the intervention period within the wider developmental context.

Attrition and adherence

I report how many participants submitted enough weekly data to be included in the analysis of primary outcome (>66% of test data-points), as well as completeness of data (proportion of missing data). Additionally, given its importance to accuracy of the randomization test, I report adherence to allocated intervention start date for each participant.

5.2.3. Analysis plan

Multiple methods were used to assess the effectiveness of the intervention.

5.2.3.1. Randomisation Test

The statistical model used to analyse the significance of a positive change in the primary outcome variable (elicited phoneme test score), was the randomized phase design with resampling as outlined in Rvachew & Matthews (2017). This is a one-tailed analysis, and was calculated in R (R Core Team, 2017) using the script detailed in Appendix N. The anonymised dataset is available upon request.

For each participant I derived a mean A week score (baseline period) and a mean B week score (intervention period); subtracting A from B weeks to yield a mean difference score. I then generated the range of all eight possible mean difference scores (assuming the intervention had started in either week 4, 5, 6, 7, 8, 9, 10 or 11) and comparing them in size to the actual mean difference obtained. If the intervention had no effect (the null hypothesis), there would be a 1/8 chance that the obtained mean difference would be the greatest score when compared to all the 7 other outcomes. The relative ranking of the actual mean difference is translated into a p-value, for example, if there are 8 possible comparisons, and the mean difference is the greatest, this equates to a p-value of 1/8 or .125.

The next stage is to pool the p-values across participants, to gauge the consistency of any treatment effects. This was done using the `sumz` function in the `MetaP` Package in R (Dongliang, 2009; Dewey, 2019), which uses Stouffer's z-trend procedure to generate a p-value to denote the likelihood of achieving a series of

p-values merely by chance. I used a p-value of less than .05 for significance testing for the meta-analysis of p-values. The initial p-values derived for each case were generated assuming a one-tailed test.

5.2.3.2. *Between-case Effect Size*

Between-case Effect Size was calculated for the case series using the 'scdhlmm' package (Pustejovsky, 2016) and following the guidelines set out in Valentine et al. (2016). Thus performing the command `MB_effect_size()` generated an adjusted d statistic as well as its variance.

5.2.3.3. *Secondary outcome measures*

For the secondary measures, a paired t-test was carried out to compare the mean 4-month change in the 12 months prior to intervention, with the actual 4-month change in the 4 months between Visit 1 and Visit 2, for both consonant inventory and expressive language. These were one-tailed tests with a .05 significance threshold.

5.2.3.4. *Specificity of improvement to trained sounds*

To analyse whether a sound targeted during intervention was significantly more likely to be accurately elicited following the intervention, I conducted a Chi squared analysis of all probe data from week 16. This allowed examination of the difference in distribution of trained and untrained responses (correct/incorrect). When no data was available from week 16 ($n = 6$), I took the data from the last available week. The analysis was also one-tailed with a .05 threshold. Ideally the result would be adjusted due to the clustered nature of individual observations (each participant was tested on nine sounds). Various methods have been proposed (Gönen, 2004; Obuchowski, 1998; Yang, Sun & Hardin, 2010) but there is no freely available software to compute this adjustment.

5.2.3.5. Generalisation

To analyse whether greater success on targeted phonemes was related to greater success on untargeted phonemes (a generalization effect) I determined the correlation between this pair of variables (using raw score) on a group level. A significant positive correlation (Pearson's $r > .3$; $p < .05$) would support the hypothesis that improvements in trained sounds have extended to improvements in untrained sounds.

5.3. Results

5.3.1. Reliability of parent ratings of speech production attempts

The intra-class correlation coefficient for speech production ratings by parents compared with those by an independent rater was .84 when scores of 0, 0.5 and 1 were considered (0=no response, 0.5= incorrect attempt and 1=correct). When scores were re-categorised to reflect a binary correct/incorrect split (scores of 1 and 0 respectively, with an incorrect attempt scoring 0 instead of 0.5), this figure rose to .95. In light of this, scores of 0 and 1 were used in all subsequent analyses, rather than 0, 0.5 and 1, as originally planned. Individual weekly scores from the reliability analysis are plotted in Figure 22 to demonstrate the level of consistency achieved.

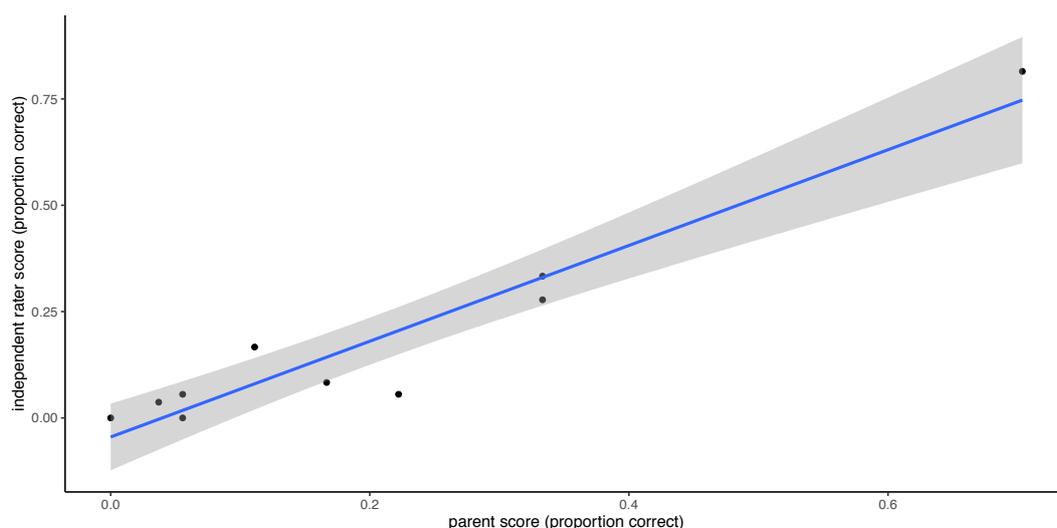


Figure 22. Reliability of parent-rated versus clinician-rated weekly scores

5.3.2. Can weekly scores be used to evaluate efficacy using the Randomization Test?

Attrition for the randomization test was 47%, as of the 19 original participants, only 10 were classified as 'high' users of the app, insofar as they completed >66% of test trials. Amongst these high users, the mean number of test trials completed was 82% (sd=11%, range=69-100%). The results below show that it has been

possible to calculate efficacy measures using the data collected from these 10 participants despite the missing data points. Comparison of allocated intervention start date and actual intervention start date revealed a mean delay of 1.4 weeks (sd=1.3, range=0-3). A more in-depth description of user characteristics, acceptability and usability data for this project are reported in Chapter 4.

5.3.3. Efficacy results from BabbleBooster

5.3.3.1. Randomization Test

Figure 23 presents the individual weekly probe scores of each participant (score out of 9 expressed as a percentage). The vertical line represents the allocated start week for intervention and the dashed line is the actual start week.

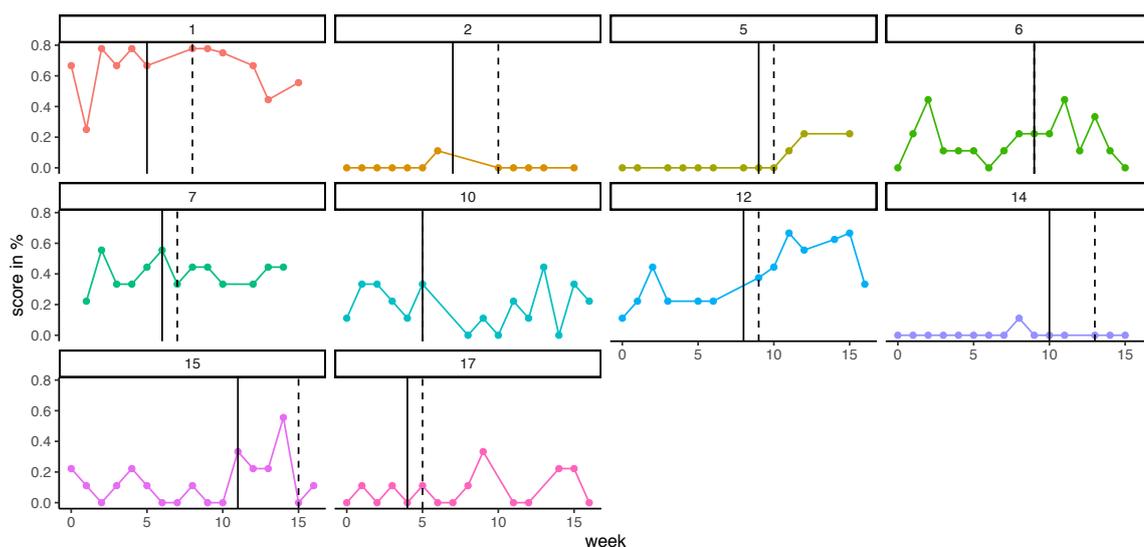


Figure 23. Weekly scores on elicited phoneme test, by participant (as %)

These scores were used to compute the mean difference score for each participant and compare it to the distribution of potential outcomes. Intervention was deemed to commence at the actual rather than allocated start date. Table 11 reports each participant's mean score and sd for A and B weeks, the mean difference between them, and the corresponding rank and p-value associated with that mean difference. A non-significant Stouffer's Z statistic was calculated for this range of

p-values ($z = .314$, $p = .38$), indicating that they were not significantly different from p-values expected by chance. In accordance with the pre-registration, this procedure was also re-run using multiply imputed values, also generating a non-significant result ($z = -.115$, $p = .91$).

5.3.3.1. *Between Case Effect Size*

The Between-Case Effect Size for the above data, adjusted for small sample size, is 0.161 with a variance of 0.011. This small effect size is consistent with the non-significant main finding.

5.3.3.1. *Specificity and Generalisation*

Two analyses were performed using the elicited phoneme score on the final week of data collection (which was week 14-16 depending on missing data, week 16 ($n = 4$), week 15 ($n = 5$), week 14 ($n = 1$)).

To determine if there had been a specific improvement to trained phonemes, a Chi squared of trained versus untrained phoneme scores in the final week of intervention is displayed in Table 12 ($\chi^2(df = 1) = .904$, $p = .34$).

To determine if improvement in trained phonemes had generalized to untrained phonemes, a Pearson's correlation analysis of individual scores for trained vs. untrained phonemes in the final week of data collection was performed ($r = .25$, $p = .48$).

Table 11. Comparison of A and B week elicited phoneme scores

ID	A week mean (sd) elicited phonemes (proportion correct)	B week mean (sd) elicited phonemes (proportion correct)	Mean difference (B – A weeks)	Rank	p-value
1	0.642 (.196)	0.660 (.137)	0.019	3	0.375
2	0.016 (.042)	0.000 (.000)	-0.016	5	0.625
5	0.000 (.000)	0.139 (.106)	0.139	2	0.250
6	0.148 (.136)	0.206 (.149)	0.058	2	0.250
7	0.407 (.135)	0.397 (.059)	-0.011	7	0.875
10	0.222 (.111)	0.178 (.159)	-0.044	4	0.500
12	0.241 (.109)	0.525 (.140)	0.284	3	0.375
14	0.009 (.034)	0.000 (.000)	-0.009	6	0.750
15	0.148 (.087)	0.056 (.191)	-0.093	1	0.125
17	0.044 (.061)	0.110 (.122)	0.065	4	0.500

Table 12. Distribution of correct and incorrect scores for trained and untrained phonemes

	Correct	Incorrect	Total
Trained phonemes	6	24	30
Untrained phonemes	9	51	60
Total	15	75	90

5.3.3.2. Secondary Outcome variables

For the secondary outcome variables, I collected data from 18 participants (all except the one lost to follow-up). I have therefore performed the following analyses on an 'intention to treat' basis (n=18). However, since only 10 participants engaged

meaningfully with the intervention, I also present exploratory analyses for this smaller group (n=10).

Descriptive statistics are summarised for the intention to treat group in Table 13. Figures 24 and 25 show the outcome variables Expressive Language and Consonant Inventory over 5 time-points. Time 1 occurred 12 months prior to Visit 1, and time-points are evenly spaced by 4 months (sd=0.3). Visit 1 thus corresponds with Time 4 and Visit 2 is Time 5. The red dot indicates the Time 5 measurement of each variable, which is the one we would hypothesise to demonstrate any treatment effect.

Table 13. Descriptive variables

Measure	n	Mean	sd	Min	Max
Age at Visit 1 (months)	18	61.6	7.5	47.6	74.6
Receptive Language Visit 1 (words)	18	182.2	135.2	5.0	406.0
Expressive Language Visit 1 (words)	18	4.5	6.4	0.0	19.0
Consonant inventory Visit 1 (sounds)	18	6.4	3.6	1.0	13.0
Age at Visit 2 (months)	18	65.7	7.3	52.2	78.3
Receptive Language Visit 2 (words)	18	195.0	141.9	5.0	417.0
Expressive Language Visit 2 (words)	18	11.6	26.3	0.0	90.0
Consonant inventory Visit 2 (sounds)	17	5.2	4.4	0.0	16.0
Autism Symptom Severity at Time 1 (raw score)	19	42.7	4.9	35.0	52.5
NVIQ at Time 2 (DQ)	19	0.36	0.13	0.13	0.56
Communicative intent at Time 1 (total comm. Acts)	19	20.1	11.5	0.0	45.0
Response to Joint attention at Time 1 (% correct)	19	18%	29%	0%	100%
Parent child interaction at Time 1 (% leads)	19	52%	17%	18%	84%
Consonant inventory at Time 1 (raw score)	19	2.1	2.1	0.0	8.0
Phonetic repertoire at Time 1 (raw score)	18	7.9	7.0	0.0	26.0
Alphabet score at Time 1 (% correct)	19	10%	30%	0%	100%
Total weekly therapy (hours) at Time 1	18	4.2	5.5	0.0	22.0

SD: standard deviation; NVIQ: non-verbal intelligence quotient; DQ: developmental quotient (developmental age/chronological age)

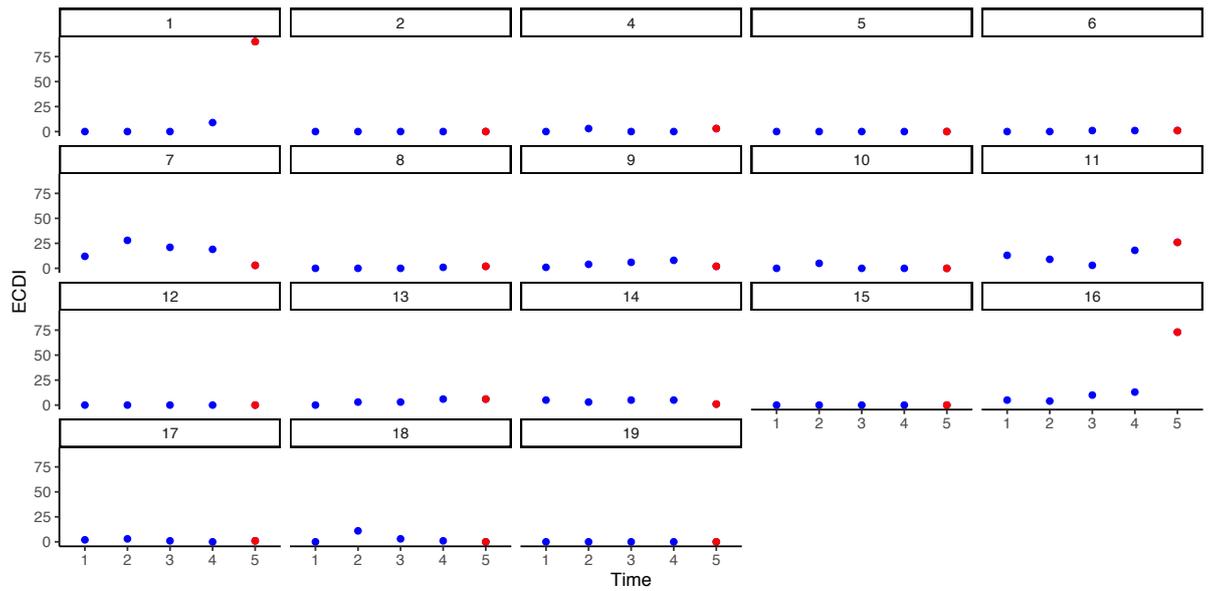


Figure 24. Expressive language over 16 months, by participant

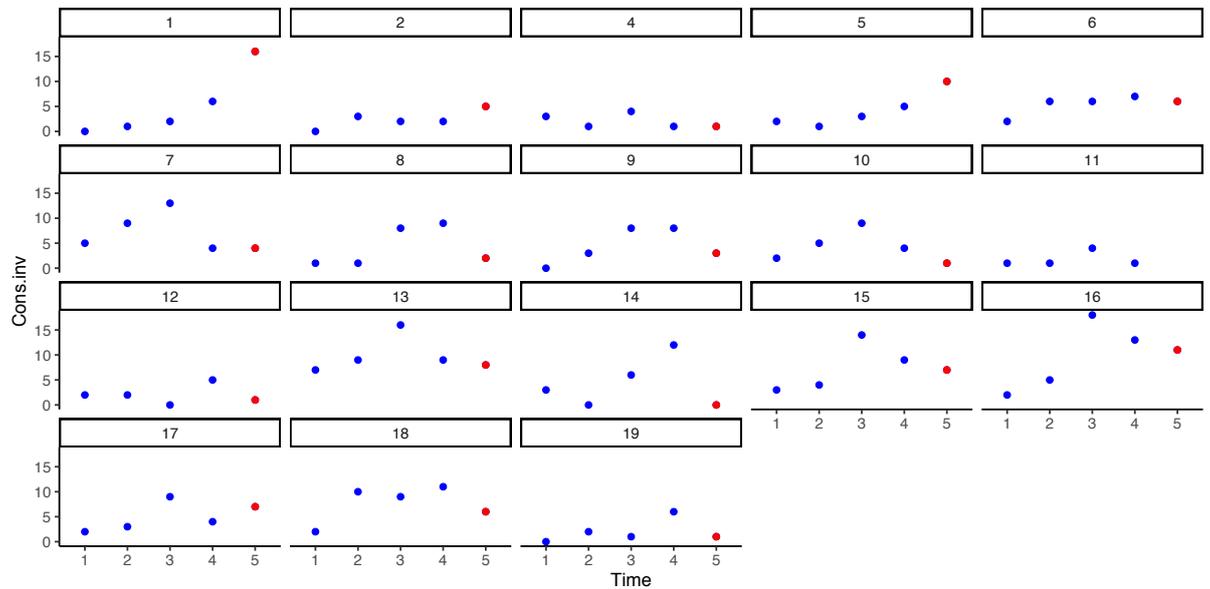


Figure 25. Consonant inventory over 16 months, by participant

Table 14 summarises the average 4-month change in both outcome variables for each participant for two relevant periods a) the 12 months prior to the intervention period and b) the 4-month period between Visit 1 and Visit 2. Both outcome variables were transformed due to skew – square root for Consonant Inventory and log10 for Expressive Communicative Development Inventory.

Table 14. Average change in secondary outcome measures for 12 months before Visit 1 and the 4 months before Visit 2

ID	ECDI average change 12 months before Visit 1	ECDI average change between Visit 1 and 2	Mean difference in average change in ECDI	Consonant inventory average change 12 months before Visit 1	Consonant inventory average change between Visit 1 and 2	Mean difference in average change in consonant inventory
1	0.33	0.96	0.63	0.82	1.55	0.73
2	0.00	0.00	0.00	0.47	0.82	0.35
4	0.00	0.60	0.60	-0.24	0.00	0.24
5	0.00	0.00	0.00	0.27	0.93	0.65
6	0.10	0.00	-0.10	0.41	-0.20	-0.61
7	0.06	-0.70	-0.76	-0.08	0.00	0.08
8	0.10	0.18	0.08	0.67	-1.59	-2.25
9	0.22	-0.48	-0.69	0.94	-1.10	-2.04
10	0.00	0.00	0.00	0.20	-1.00	-1.20
11	0.04	0.15	0.11	0.00	NA	NA
12	0.00	0.00	0.00	0.27	-1.24	-1.51
13	0.28	0.00	-0.28	0.12	-0.17	-0.29
14	0.00	-0.48	-0.48	0.58	-3.46	-4.04
15	0.00	0.00	0.00	0.42	-0.35	-0.78
16	0.12	0.72	0.60	0.73	-0.29	-1.02
17	-0.16	0.30	0.46	0.20	0.65	0.45
18	0.10	-0.30	-0.40	0.63	-0.87	-1.50
19	0.00	0.00	0.00	0.82	-1.45	-2.27
mean	0.07	0.05	-0.01	0.40	-0.46	-0.88

Note: ECDI: Expressive Communicative Development Inventory

Paired t-tests were calculated comparing the pre- and post-intervention average change data reported in Table 14 for Expressive Communicative Development Inventory ($T(17) = .140$, $p = 0.89$) and Consonant Inventory ($T(16) = 2.803$, $p = .01$). This latter significant value represents a decrease in the rate of Consonant Inventory growth, rather than decrease in absolute values over 16 months.

When the smaller intervention-compliant group is considered alone ($n = 10$), these values are both non-significant at $T(9) = .202$, $p = .85$ (Expressive Communicative Development Inventory) and $T(9) = 1.281$, $p = .23$ (Consonant Inventory).

5.4. Discussion

This chapter sought to describe and illustrate two powerful techniques for statistical analysis of Single Case Designs, which can be employed where the gold standard RCT may be difficult to implement. I used data from a brief intervention, which aimed to increase speech production skills in minimally verbal autistic children. The design and aims of the app were viewed positively by parents and 18 out of 19 families engaged on some level with the app. Nevertheless, neither the randomization test nor the between-case effect size demonstrated a significant functional relation between the intervention and the speech production outcome variable.

There are several likely reasons why change was not seen in the primary outcome measure. One factor that has become clear since the study was designed is the sheer volume of trials needed to effect even a tiny change in expressive language in this population. Comparable studies often report a large number of sessions, which are typically delivered by graduate students or clinicians. For example, Esch et al. (2009) employed an average of 50 10-minute sessions of Stimulus-Stimulus Pairing intervention to teach two specific words or syllables to three minimally verbal autistic children, but improvements were not generalised to non-target words. Chenausky et al. (2016) compared auditory-motor mapping therapy (AMMT) with traditional speech repetition therapy (SRT) in a matched group study (n=14). After 30 sessions of 45 minutes, participants were measured on change in % syllables approximated (29% improvement for AMMT vs. 3.6% for SRT). This study was limited by a 16-week timeframe which had to also include a baseline of a variable length, thus limiting the number of weeks of intervention. Future studies could use a longer case series design to determine optimal treatment intensity and duration.

A second consideration, is that this was the first time the app was formally trialled on a sizeable scale. Therefore, despite having undergone various collaborative stages of design and review, technical difficulties did arise and several key aspects of design were identified for improvement (Chapter 4). Based on parent feedback,

I expect that some attrition was related to frustration with technical difficulties. Adequate power was assumed given the starting sample, however attrition resulted in a much lower power to detect significant effects.

Furthermore, even for those participants who submitted enough data to be retained in the analysis ('high users', $n=10$), intensity of use, as measured by number of intervention trials per week or minutes spent on app per week, may have been higher if the app had been at a more polished stage of design. Chapter 4 considers ways in which the app could be improved for a future trial, in order to maximise both retention and adherence to the intervention. Key changes included: better training and support with customization features of the app; making the test module more visually engaging and rewarding for children; reducing the number of sound targets worked on at once; gamification of the intervention stimuli and inclusion of feedback to motivate and inform the parents on intervention progress. Finally, due to financial constraints the app was only available in Android format, thus most participants were not using the app on their normal phone. In future research, making the app available across all phone platforms (Android and iOS) should lead to an increase in parent engagement with the app.

Some parents ($n=3$) reported that their children were not motivated by apps or tablet use in general. A future trial could consider offering the intervention materials in an off-line format as well, so that such children could still take part. It is also worth considering the difficulty that some families have adhering to any form of intervention. Having an autistic child can place stress on a family (Baker-Ericzen et al., 2005; Estes et al., 2013; Hayes & Watson, 2013). The occurrence of family illness, carer chronic health conditions, siblings with additional needs and difficult transition periods between educational settings and school holidays, were among the many barriers to adherence faced by this cohort. Achieving 0% attrition and 100% adherence in a real-world context is therefore not a realistic goal for an intervention like this.

Unusually, this Single Case Design includes data for expressive language measures for the 12 months before the case series started. Analysis of changes in

reported expressive vocabulary and observed consonant inventory show no significant change in expressive language measures for either the intention to treat or the 'high user' group following the intervention, which is in line with the main non-significant finding. For most of the 19 participants in this study, we document a period of 16 months with little meaningful change in expressive language. This highlights the extreme degree of difficulty in advancing spoken language skills and thus it is perhaps not surprising that a brief 'light touch' parent-mediated intervention did not affect meaningful change.

An exception to this is the observed reduced pattern of growth in Consonant Inventory between Time 4 and Time 5, when compared to Time 1 to Time 4. Although a possible interpretation is that children's speech skills are getting worse over time, a more likely explanation is due to differences in circumstances between time points, which could impact child performance during a naturalistic communication sample. Over time, children's familiarity with the experimenter increased, novelty value of the stimuli decreased, whilst some measurement occasions involved additional tasks, placing extra demands on the child. I postulate that the pattern of rising and falling consonant inventory may represent these differences. This highlights the need for robust multidimensional measures of speech skills for this population, but as the main efficacy test did not rely on these data it has not been explored further. McDaniel, Woynaroski, Keceli-Kaysili, Watson and Yoder (2019) recently controlled for the number of communication acts when relating speech skills (number of canonical vocalisations) to expressive language. In future studies, consonant inventory could be adjusted for the number communication acts to reduce the impact of confounding circumstances.

The current study has shown that the chosen study design (multiple baseline with random assignment) and statistical approaches (Randomization Test and BCES) are feasible with real-world data, as generated by a sample of 10 participants. Prior to this study we did not know whether an app could be used to elicit and record speech production attempts, or if parents would be able to accurately rate those attempts online. We did not know how variable or auto-correlated such attempts would be, in children who met criteria for minimal language, nor what percentage

of recruited families would be able to meet the demands of the testing regime and comply with the intervention schedule.

Parents were able to accurately rate speech attempts of their children online in test conditions following a brief training; agreement with a trained speech-language therapist blind to assessment period exceeded .85. This means that one can have confidence in parent ratings, and they could be used to evaluate the intervention, enhancing the scalability of the app. I have also shown that incorporating a 'good try' rating for speech attempts that are incorrect tries, was not accurate enough for use as part of this paradigm (ICC=.84). Parents and independent raters differed in their perception of a child's attempt being 'incorrect' (no attempt) or 'good try' (incorrect attempt). Individual differences in child's baseline skills and visual attention patterns may make objectively scoring 'good tries' an impossible task. For example, one child may vocalise frequently in a manner which does not seem deliberate: coding all of these vocalisations as 'good try' may not accurately reflect deliberate attempts at the speech sound. A different child who vocalises very infrequently may utter a barely audible attempt, which may be perceived as a 'good try' by the parent but not by an independent rater, who is unfamiliar with the child.

This study has identified that important considerations for future research are participant attrition, missing data and adherence to intervention start date. The pre-registered analysis did not specify how to deal with these issues and future studies would benefit from a clear protocol set in advance to eliminate bias. Due to the design of this study, those not engaging with the app could not be replaced; however, in future studies replacement could be used to manage attrition, since the design does not need baselines to be sequential.

5.4.1. Limitations

Firstly, attrition resulted in a smaller than expected sample size, which limits statistical power to detect and intervention effect. Secondly, for the secondary pre- and post test analysis I used observed and elicited measures derived from home visits. The data generated in such contexts is more vulnerable to measurement error and confounding factors, due to poorer control of the testing environment, e.g. presence of pets, siblings, television screens and other distractions. However, home visits are preferred by families of children with complex needs and facilitate their participation, thus creating a broader representation of families within the study. Therefore, greater ecological validity was judged to be worth the trade-off with experimental precision. Finally, these findings are limited by the technological difficulties experienced during the BabbleBooster trial, and the limited range of intervention weeks. Whilst we can speculate that a longer study with a more polished app would have had a greater impact on the targeted proximal and distal variables, this remains to be determined.

5.4.2. Conclusion

A key goal of this paper was to outline how to exploit the benefits of single case design, by using random-assignment and the randomization test, as well as a between-case effect size to measure functional relationships between the introduction of an intervention and the outcome variable. The current study demonstrates that this is a robust method for rare, heterogeneous groups. The BabbleBooster intervention did not lead to meaningful change in spoken language skills on this occasion, however dosage and intensity have been identified as important considerations for future studies.

6. General discussion

In this thesis I explore the hypothesis that for some minimally verbal autistic children, their expressive language development is influenced by an additional co-morbid speech-motor difficulty. To do this I had two main objectives:

1. To extend work testing the hypothesis that early speech skills predict later expressive language in minimally verbal autistic children. I aimed to do this by documenting expressive language development in a group of young autistic children with minimal verbal skills and relating it to previously identified predictors, which included consonant inventory, a measure of speech production skills. I also sought to evaluate a novel phonetic composite in order to test the utility of a more comprehensive and accurate measure of speech skills in predicting expressive language.
2. To test the hypothesis that providing minimally verbal autistic children with a novel speech skills intervention could improve their speech skills and in turn positively influence their broader expressive language skills. A literature review revealed that evidence-based speech-focussed interventions aimed at minimally verbal autistic children are currently lacking, so the first step in seeking evidence for my hypothesis was to design and evaluate a parent-mediated intervention, aimed at improving speech production skills.

Below I outline the main empirical findings of my thesis, before discussing the contribution of these findings to our understanding of the role of speech skills in expressive language for minimally verbal autistic children. Practical and theoretical implications are discussed, as well as future directions for this research.

This thesis contributes further evidence for the important contribution of speech skills in expressive language development for minimally verbal autistic individuals. It demonstrates that aiming to improve speech skills via a parent-mediated app

intervention is feasible, and can be effectively evaluated with a randomized single case series design. Results indicate that future trials of the app will need to consider greater dosage (intensity, duration and frequency of speech practice).

6.1. Summary of Findings

In Chapter 1, I reviewed literature seeking to explain the huge variation in language trajectories and outcomes in autism. In a multi-factorial causal model factors combine in such a way to severely limit expressive language development for some autistic individuals, who remain minimally verbal. One potential causal factor that has received relatively little attention is speech skills, which I explore in the longitudinal study in Chapter 3.

In Chapter 2 I summarised results of intervention studies seeking to improve expressive language for minimally verbal autistic participants, concluding that high quality evidence using robust analysis methods was lacking, hence the pilot intervention developed and tested in Chapters 4 and 5.

Chapter 3 describes Study 1, in which I measured language development over 12-months in a cohort of minimally verbal 3-5 year olds ($n=27$), at four time-points. Expressive language growth was highly variable, with 65% remaining minimally verbal and two children rapidly gaining over 340 words. Children gained 17 words on average over the year ($sd=33$), after adjusting for these two outliers.

A planned composite comprised speech sounds observed in a semi-structured language sample (Observed Phoneme Inventory), those reported by parents (Reported Phonemes) and those elicited experimentally (Elicited Phonemes), in order to reflect multiple facets of speech production skill. These measures were moderately correlated with one another. The resulting Phonetic Repertoire measure correlated with each component $r>.60$ and also significantly with consonant inventory ($r=.87$).

Contrary to expectations, initial communicative intent, parent responsiveness and response to joint attention did not predict expressive language growth or outcome in this sample. In contrast, both consonant inventory and phonetic repertoire (which both measured speech skills) were significant predictors (adjusted R squared of 0.29 and 0.45 respectively). These results contrast with Yoder et al. (2015), who

found all four predictors to be significant in a younger and larger but otherwise similar cohort (n=87).

Chapter 4 describes the design and feasibility testing of a novel app-based intervention (BabbleBooster), designed to enable parents to practice preverbal speech skills regularly with their minimally verbal autistic children. Consultation with the autism community refined the design and culminated in a pilot trial (n=19) lasting 16 weeks. Acceptability scores were high (mean score 74%, sd=25%, range=55-89%) and qualitative feedback indicated that families liked the overall goals and features of the app but there were several key opportunities to improve the app in future. Usage and adherence figures were variable. Ten participants engaged meaningfully with the app and for these families, test schedule compliance (82% of weeks) and intervention schedule adherence (61% of weeks) were good. Mean engagement during intervention was lower than targeted (mean of 12 mins/week vs. 25 mins recommended, and a mean of only 15 trials/week). This reflected a mixture of technical difficulties and limitations in families' capacity to engage with therapy in general.

In Chapter 5 I described how and why I used a randomized single case series design to evaluate BabbleBooster's efficacy. Participants were allocated a randomized baseline length and undertook weekly probes using the app, which were then compared between baseline and intervention phases. Parents proved accurate raters of their child's productions in the weekly probes, a key component of the app-based evaluation (ICC = .95).

A randomization test was computed whereby actual mean difference scores for each participant's assigned intervention schedule were compared to the distribution of possible outcomes for non-assigned schedules, and resultant p-values were pooled together across participants. An effect-size quantified the standardized mean difference for the case series, taking into account between-participant variance. The randomization test and effect size demonstrated that BabbleBooster did not improve speech production scores during this 16-week period. Broader expressive language scores and consonant inventory measures

before and after the intervention also did not improve during the intervention period.

6.2. Implications

6.2.1. Expressive language trajectories for minimally verbal autistic children

6.2.1.1. Large variation but limited progress for many

A striking feature of the language trajectories in Study 1 was the large variation in outcome. One third of children surpassed minimally verbal criteria during the study, two of whom did so dramatically (+340 words). However, a large proportion of participants did not make meaningful progress in a 12-month window, by which time their mean age was 5;2 years. Over a quarter continued to have zero spoken words in their productive vocabularies. As a group (adjusted for outliers) mean progress over the year was just 17 words. This is slower progress than reported in Yoder et al. (2015), where 40% of participants remained minimally verbal after 16 months, and average progress was 75 words (sd=95). Although participants in Yoder et al.'s study were similar to this cohort on measures of developmental ratio, receptive and expressive language scores, this study's participants were recruited at an older age (3 to 5, mean 4;2) rather than age 1;8 to 3;11 (mean 2;11) in Yoder et al. (2015). It is possible that some of the younger children in Yoder et al. (2015) did not have such persistent expressive language impairment, but were experiencing a transient delay in language development which partially resolved during the study period. This suggests that this study's sample may include children with more severe symptoms and greater difficulty acquiring expressive language.

Few other similar longitudinal studies are directly comparable due to differences in design, definition of minimally verbal, or sample characteristics (e.g. Anderson et al., 2007; Bal et al., 2016; Norrelgen et al., 2014). Bak, Plavnick and Bryne (2019) measured vocalisations of minimally verbal autistic 6-10 year olds (n=9) using regular day-long recordings. They identified a similar pattern of very low growth in expressive language, with eight out of nine participants showing no increase in vocalisation rate over the course of a year. Against this backdrop of extremely slow

progress for many individuals, the lack of responsiveness to a brief 'light-touch' parent-mediated intervention described in Chapter 5 is not surprising and a need for a more intensive approach seems plausible (see below).

6.2.1.2. Stability

If individual rank order is maintained on a test measuring the same skill repeatedly over time it is considered that the measured skill is highly stable. Expressive language measures were stable over time, and particularly between adjacent time-points at Time 2, 3 and 4 (all $r \geq .90$). This stability is consistent with Bornstein et al. (2018), and further demonstrates that the high stability seen in neurotypical language profiles from an early stage is also the case in atypical populations. This raises interesting questions about the timing of interventions and the degree of impact needed to enable children to 'catch up' in view of such strong stability, and whether it would be more appropriate to focus intervention efforts on developing alternative forms of communication. In contrast to my findings, Pickles et al. (2014) observed greater variability and plasticity in language trajectories prior to age 6 and high stability thereafter, in a large autistic cohort. This pattern suggests that there could be a limited window for language interventions to maximise their impact. Further replications are needed to better understand stability in language development in autism and what this means for interventions. Bornstein et al. (2018) and Pickles et al. (2014) both measured language as a latent variable comprising expressive and receptive skill in contrast to the current study, where only expressive language is reported. A longer term follow-up of Study 1's participants would enhance our understanding of stability and variation in this particular cohort as they progress.

6.2.2. Importance of speech skills

6.2.2.1. Speech skills predicted later expressive language

The main finding that speech skills predict unique variation in expressive language skills over 12 months in autistic children who all began the study with minimal spoken language, adds to emerging evidence that speech production abilities are foundational to expressive language development in autistic preschoolers, as they are in typical populations (McDaniel et al., 2018, McGillion et al., 2017). Demonstrating longitudinal precedence is a first step in gathering evidence of a causal relationship, however a third variable (something which influences speech skills and expressive language) could be responsible for this finding. Devising an intervention to improve speech-motor skills and recording a subsequent improvement in expressive language skills, relative to a control group or control baseline period, would provide more convincing causal evidence, though such evidence is currently missing from the extant literature.

6.2.2.2. Why might speech skills matter?

This thesis was not designed to elucidate mechanisms through which speech skills contribute to expressive language, although this would be important for furthering our understanding of the nature of their influence. Speech attunement theory postulates that speech skill development is limited by social engagement factors in autistic children, and is based on findings that phonological skills seem to develop in line with general socio-communicative development in autism (Paul et al., 2013; Shriberg et al., 2011). However, Shriberg et al.'s study did not include minimally verbal participants, who may have a specific and different pattern of speech-motor abilities and challenges. Future studies of speech and phonological development could include children with minimal verbal language to determine whether evidence supporting speech attunement is applicable to all autistic children.

In order to disprove the speech attunement theory it would be necessary to properly assess and quantify the range of communicative competencies minimally verbal autistic children *do* have, for example receptive language competence or ability to express themselves via other symbolic means (e.g. sign or symbol use). However, great effort is required to accurately measure these abilities, and the reliability of parent's report of receptive skills is unclear. Belmonte et al. (2013) noted a disparity between receptive and expressive language skills in a motor-impaired subgroup of autistic preschoolers, which supports the hypothesis that motor difficulties contributed to their lack of speech progress. Evidence for minimally verbal participants' ability to use alternative forms of communication (sign, symbol, speech generating devices) would point to phonological development out of sync with other socio-communicative skills, and lend support to a speech-motor explanation.

The suggestion that for some minimally verbal autistic children, speech production difficulties reflect a motoric difficulty has been supported by the growing links between early motor skills and later language (Bhat et al., 2012; LeBarton & Landa, 2019; Mody et al., 2017). Imitation is a related but separate skill, which also relates to expressive language in concurrent and longitudinal studies of autistic children (Charman et al., 2000; Gernsbacher et al., 2008; Ingersoll & Meyer, 2011; Luyster et al., 2008; McDuffie & Yoder, 2005; Pecukonis et al., 2019; V. Smith et al., 2007; Stone et al., 1997; Van der Paelt, Warreyn & Roeyers, 2014). A variety of different protocols and types of imitation (manual, oral motor and vocal) have been associated with expressive language in standardized structured tests and in naturalistic social contexts. A larger scale survey of imitative and non-imitative vocal skills in this population would be helpful to further classify the range of oral-motor skills available to minimally verbal children and their developmental relationship with expressive language.

6.2.2.3. *Measuring speech in this population*

Multi-level statistical models are only as good as the variables contained within them, and accurately measuring facets of development in young, minimally verbal autistic participants is particularly challenging. A novel aspect of Study 1 was the attempt to measure different facets of speech skills in order to reduce measurement error and provide a more nuanced and complete picture of participants' abilities. When a skill is emerging, it is advantageous to incorporate various sources of reporting (observation, parent report, experimental measures, Broome et al., 2017). The enhanced predictive power of the phonetic repertoire composite (compared to consonant inventory alone), supports the idea that for minimally verbal autistic children, a broader measure of speech skills, incorporating information from multiple sources, better represents their true ability, a sentiment echoed more generally when considering language assessments for this group (Tager-Flusberg et al., 2009).

Future research could develop and test optimal measurement tools for speech skill development and oral-motor functioning in this population, for example, the novel 'articulatory features' assessment developed by Sullivan, Sharda, Greenon, Dawson and Singh (2013) or Chenausky et al.'s childhood apraxia of speech checklist (2019). When trying to separate what children can do (competence, reflecting underlying abilities) and what they actually do (performance, which can be affected by extraneous factors), it is important to obtain the best possible representation of a child's abilities. Use of technology should be explored. For example, gamified app-facilitated methods of eliciting speech productions remotely, would enable the child to be assessed at home by a familiar adult. This could yield more accurate information compared to a face-to-face assessment with an unfamiliar speech-language therapist conducting a probe in an unfamiliar clinic setting. Computer recognition could be deployed to analyse visual or auditory aspects of production, with a view to identify irregularities. Day-long recordings of natural language with automated analysis of atypical vocalisations present a further opportunity to characterise individual differences (Woynaroski et al., 2017,

Swanson et al., 2018). Looking ahead, these approaches have potential applications to early diagnosis of speech-motor issues if they prove valid.

6.2.2.4. *What predicts speech skills?*

Given the finding that speech skills are predictive of later expressive language, it is critical for future research to investigate what drives and predicts the development of speech skills in autistic children. Woynaroski et al. (2016) found that child intentional communication acts and parent linguistic responses (mediated by receptive language) predicted consonant inventory growth over 16 months in a cohort who were minimally verbal at intake (n=87). Attention to speech and motor imitation were not significant predictors in the regression model once these predictors and two background variables were controlled for, despite being originally identified as predictors in an earlier analysis of the same cohort (Patten, Watson & Yoder, 2012). Replication in a new cohort of minimally verbal children is required, in order to determine whether early motor abilities predict speech skill development or a socio-cognitive explanation is a better fit to the observed pattern of development. The socio-cognitive precedence would favour the speech attunement hypothesis, providing evidence that children tune up their speech production skills as their overall social communication skills develop. In the current study social variables were not linked to expressive language outcomes (see below). Future work could incorporate longitudinal measures of phonetic repertoire in order to build a more informed picture of trajectories of phonetic abilities in this population.

6.2.3. Social variables were not significant predictors

6.2.3.1. *Social skills not always protective for expressive language*

A surprising result of Study 1 was that the three predictors designed to assess social engagement (response to joint attention, child communicative intent and parent responsiveness) did not significantly relate to expressive language

outcomes. I am cautious not to over-interpret this null finding, and discuss in Chapter 3 the possibility that low power or floor effects could have contributed to this result. Statistical power is often problematic given the difficulty of recruiting this hard to reach population. Additionally, in longitudinal studies, attrition and missing data can be high which can skew results. A strength of my study is that I pre-registered my hypotheses and limited analysis to only 4 predictors given my sample size, in order to reduce the risk of type 1 error. Furthermore the dataset was almost complete at each time-point. Nevertheless, given the small sample and potential alternative explanations for the null finding, it will be necessary to replicate this finding with a larger sample.

The lack of predictive power of social engagement variables contrasts with Yoder et al. (2015) and may be linked to the current study's older age at intake discussed above, resulting in a less diverse group with greater overall language impairment. One interpretation of the lack of predictive relationship between social variables and expressive language is that some minimally verbal autistic children could have an additional disorder of speech-motor development. If this were the case, stronger socio-communicative skills would not act as protective factors for expressive language to the same extent that they do in younger and thus more diverse minimally verbal autistic cohorts. In the realm of joint attention, Toth, Munson, Melzoff and Dawson (2006) have put forward the notion of a 'starter set' or minimum required level of joint attention abilities needed to progress with language development. The same could be true of speech skills, which would explain why different cohorts have different patterns of predictors. Bottema-Beutel (2016) reiterated Toth et al.'s idea (again discussing joint attention rather than speech skills), claiming that there may be a threshold above which it is not so important how able you are, but if you are beneath it, it will have a significant cascading effect on language development.

Further investigating the role of 'starter sets' in language development will require carefully considered study designs. If the relationship between speech skills and expressive language were linear across the spectrum of language abilities (an increase of a in speech skills results in an increase of $a*b$ in expressive language

skills), we would expect a linear model with a heterogeneous sample to highlight speech skills as a predictor. However, if the relationship between speech and expressive language was only important when speech skills fell below a given threshold, this relationship may not be as apparent in the model, depending on the characteristics of the sample. One solution is to home in on the population of interest, i.e. examining only older minimally verbal participants, or to group participants by language age (e.g. Van der Paelt et al., 2014). Supporting and enabling the participation in research for families with minimally verbal autistic children should be a key goal in addressing previous sources of bias in studies that have led to limited generalisability of findings.

Study 1 provides evidence that the relevance of predictive factors may vary across the range of language ability, echoing recent findings by Pecukonis et al. (2019). They investigated the concurrent relationship between play, joint attention and manual imitation in minimally verbal autistic children and adolescents (n=37). Joint attention skills did not correlate significantly with expressive language after controlling for non-verbal IQ, and only manual imitation reached significance as a concurrent predictor in the regression model.

Van der Paelt et al. (2014) divided a sample of 83 autistic preschoolers into two groups based on whether expressive language age was under or over 2 years. They observed different patterns of concurrent relationships between play, imitation and joint attention skills and expressive language, noting that joint attention skills were only associated with the younger language age group, but imitation skills were significantly correlated with both groups. This finding reinforces the validity of imitation as a correlate of expressive language, and supports a 'starter set' notion for joint attention (Toth et al., 2006). Speech skills were not examined in this study but its conclusions lend support to the idea that distinct correlates of expressive language are evident at different stages of language development, and for groups with varying social, cognitive and language profiles.

Chenausky et al. (2019) grouped low and minimally verbal participants (n=54) into several groups depending on the presence or absence of characteristics

associated with childhood apraxia of speech. In participants with no suspicion of speech-motor difficulties, receptive language predicted concurrent expressive language (as measured by number of different words spoken). In both the group suspected of apraxia and the group with insufficient speech to determine apraxia status, speech skills were the only significant predictor, providing further support for the importance of a 'starter' set of speech skills, and the disconnect between receptive language skills (which could be acquired via social engagement) and expressive language outcome.

We need to ensure any findings from studies comprising diverse cohorts also apply to those most impacted by communication difficulties. Perhaps future work should centre on identifying which factors are the key barriers for specific individuals, rather than finding a suite of factors that are predictive in a general way to all autistic individuals.

Findings of Study 1 also demonstrate that it is important not to assume uniformly low communicative competence in minimally verbal autistic children. A few children in the current sample were frequent and productive users of alternative forms of communication (Makaton; speech generating application), despite their lack of verbal output. This is a further indicator of a specific additional difficulty with speech modality rather than a lack of motivation or symbolic representation, and demonstrates that some children are accessing language learning from social engagement, thus they may already have some joint attention abilities and benefit from parent responsiveness.

6.2.4. Novel contributions from intervention study

6.2.4.1. BabbleBooster has feasible potential for development

Feasibility results from Study 2 highlight that it is possible to target speech skills in a parent-mediated, app-based programme, and that parents can accurately assess speech accuracy using this app. There is merit in combining several novel

features to create a personalised app, including cued articulation, video modelling and frequent practice rewarded by personalised video stimuli. This delivery method has the potential to enhance salience of the speech stimuli, reduce attentional load to attend to speech models and to incorporate extrinsic (non-social) motivation for learning. Parents evaluated the features and aims of the app positively, highlighting that this approach has potential for further development.

These findings also contribute to evidence that co-design is both feasible and useful when devising interventions for special populations. In future I would go further with a participatory approach, and would also devise ways to involve autistic children (including those with minimal verbal skills) in the intervention development process as much as possible (as well as parents). Co-design with the autism community in this way may present additional challenges, but is vital for ensuring a user-centered approach that can be personalised for individual families.

Useful conclusions can nevertheless be drawn from the qualitative parent feedback for future development of BabbleBooster. Key improvements to the intervention suggested by parent feedback included: better training and support with customization features of the app; making the test module more visually engaging and rewarding for children; reducing the number of sound targets worked on at once; gamification of the intervention stimuli and inclusion of feedback to motivate and inform the parents on intervention progress. In future research, resolving technical difficulties and making the app available across all phone platforms (Android and iOS) should lead to an increase in parent engagement with the app and better treatment fidelity.

Although compliance with the intervention was highly variable, for over half of the participants, the testing regime was adhered to and parents engaged significantly with the app. There are not many direct comparators for this intervention since most parent-mediated interventions are delivered via face-to-face coaching and tend to be evaluated on course attendance or strategy deployment. Stockwell et al. (2019) trialed a parent-mediated app with a similar cohort (children with motor and communication disorders) and noted 44% attrition and median participation

rates of 26 out of 39 sessions. Parents reported that they found the intervention useful but cited time pressures and technical problems amongst reasons for lower engagement.

No statistical differences were observed between the high and low engagement groups in this study on measures of cognition, language, age, socioeconomic status, symptom severity or intervention receipt, however, the usage data contributes to the sparse research recording participation patterns for families with a minimally verbal child, which is important for maximising parent co-operation in future trials (Shire et al., 2014). Future studies could seek to identify other factors which could be important predictors of intervention compliance in this population, such as parent physical and mental health, employment, confidence in using technology or delivering therapy. Alternatives could include implementing the intervention in school rather than home settings or supplementing the app with face-to-face support. It should also be considered that not all children will be motivated by and enjoy app-based therapy. Three of the children's families reported that they were just not interested in technical devices. In these cases, future studies could consider whether the principles of the intervention design could be applied through different media.

6.2.4.2. *Methodological contribution*

The core aim of a feasibility trial is to determine 'can it work?', which was explored in Chapter 4 (Orsmond & Cohn, 2015). As the statistical analysis and testing mechanism were also novel in this context, Chapter 5 addresses the question 'can planned efficacy analyses be carried out using the app' as well as reporting the findings of these analyses. The current study shows that the chosen design (multiple baseline with random assignment) and statistical approaches (Randomization Test and BCES) are feasible with real-world data, as generated by a sample of 10 participants. They therefore represent a viable evaluation method applied to language interventions in autism for the first time.

Developing the use of single case designs in this field could build capacity to accumulate robust intervention evidence prior to a larger study, given the time and financial commitment required to execute an adequately powered RCT. Single case designs have historically been viewed as methodologically inferior but new analytical techniques, design standards, and open source analysis software are making them an accessible and robust alternative for evaluating small scale interventions in rare or heterogeneous groups.

The analytic technique used to calculate effect size in the intervention study can also be used for meta-analysis of several single case designs. This would enable a body of evidence to be generated incrementally, so that a subsequent large scale trial is built on a solid theoretical and empirical foundation.

Larger sample sizes are often called for in order to increase statistical power in intervention studies, but an equally valid approach to increasing statistical power is to reduce error variance by focusing on a well-defined cohort and assessing them using accurate and reliable assessments. This can be achieved by improving our measurement tools and using composites where possible, but also through repeated sampling of the variable of interest over time, as advocated by Smith and Little (2018). Such methodological improvements to study design can yield robust findings with smaller sample sizes.

In the feasibility trial, the baseline and intervention periods were randomly allocated and the maximum duration of treatment pre-determined. However it is clear from the longitudinal study that the pace of language change is not homogenous, and intervention parameters could be more usefully shaped by the needs of individual children and their families. Innovative intervention study designs such as SMART (sequential multiple assignment randomized trial, Kasari et al., 2014a) where the dosage and scope of the intervention is adjusted according to children's response, are able to provide slow/non-responders with additional input to maximise outcomes. Implementing these designs in future trials of BabbleBooster could elucidate the range of dosage parameters (intensity, frequency, duration) that are required to elicit meaningful change in speech production, and whether

personalising the approach in this way facilitates implementation of the intervention for families.

6.2.5. Lack of intervention effect

6.2.5.1. Dosage

The intervention efficacy data revealed no significant change in speech skills or broader expressive language measures during the 16-week intervention period. I postulate that there are numerous reasons why we saw little evidence of change during this period. The most obvious reason for minimal speech and language growth is the brevity and lack of intensity of the speech sound training. The speech interventions reviewed in Chapter 2 highlight that most studies to date have incorporated many, many hours of clinically supervised intervention to effect small, but meaningful change. Esch et al. (2009) employed an average of 50 10-minute sessions of Stimulus-Stimulus Pairing intervention to teach two specific words or syllables to three minimally verbal autistic children, but improvements were not generalised to non-target words. Chenausky et al. (2016) compared auditory-motor mapping therapy (AMMT) with traditional speech repetition therapy (SRT) in a matched group study (n=14). After 30 sessions of 45 minutes, participants were measured on change in % syllables approximated (29% improvement for AMMT vs. 3.6% for SRT). This study was limited by a 16-week timeframe which had to also include a baseline of a variable length, thus limiting the number of weeks of intervention. Future studies could use a longer case series design to determine optimal treatment intensity and duration, as well as including a suite of measures to monitor the functional impact of any speech improvement. This would assess whether skills gained have generalised to useful communication and the impact of this on broader quality of life.

6.2.5.2. *Inadequate power*

Despite adequate power at the start of the study, attrition depleted the sample size and resulted in a lower power to detect significant effects. The likelihood of participants dropping out or not meaningfully engaging with the app in this population must be borne in mind when recruiting for future trials, and a mechanism for replacement of participants needs to be determined in advance. Another practical consideration is that participants may not drop out of the study entirely but may delay starting the intervention or contribute incomplete data. This can result in an alteration to the pre-determined randomized schedule and reduction in weeks of intervention received. When ensuring a smooth implementation of future trials using randomized baselines, a clear protocol to deal with delays, missing data and attrition will be valuable. In the current study, given the flat expressive language trajectories of many participants prior to and during the intervention period, the null effect is unlikely to be merely due to low statistical power.

6.2.5.3. *Next steps*

As well as the improvements to the app identified in Chapter 4, the need for greater intensity over a longer period are key considerations for future development of BabbleBooster. Alternative and augmentive communication is also important to consider when deciding how best to meet overall communication goals within limits of current service delivery models. Parents may seek to deliver this type of speech practice via a low-cost and highly scalable app alongside other communication-focussed approaches. Future trials may consider comparing speech-based training with alternative and augmentative communication to determine which yields greatest improvements in communication, quality of life, and behaviour. Alternatively, there is no a priori reason to expect that a speech-based intervention cannot be delivered at the same time children are exposed to alternative and augmentative communication strategies. Determining the optimum time and combination of communication approaches is a priority for future research.

6.3. Limitations

Both the longitudinal and feasibility study have several shared limitations. In both cases the sample size is small, which impacts statistical power. Secondly, for practical and ethical reasons (the time to test a small child on a diagnostic test when diagnosis is known), no formal independent diagnostic verification process took place on the participants (e.g. ADOS assessment). However, each family reported that autism had been diagnosed by a qualified health professional, and children scored a mean of 41.3 on the Childhood Autism Rating Scale (CARS) autism symptom severity assessment (only one child scored less than the 30 cut-off score). Thirdly, several observed and elicited measures were derived from home visits. The data generated in such contexts is more vulnerable to measurement error and confounding factors, due to poorer control of the testing environment, e.g. presence of pets, siblings, television screens and other distractions. However, home visits are preferred by families of children with complex needs and facilitate their participation, thus creating a broader representation of families within the study. Therefore, greater ecological validity was judged to be worth the trade-off with experimental precision.

A specific limitation to the longitudinal work relates to the use of single estimates for most predictor measures and for the dependent variable, which was done to limit testing time for participants. Composite scores would have created more robust estimates and been preferable, however this is unlikely to substantively change the outcomes of this study.

The interventions and feasibility results are limited by the technological difficulties experienced during the BabbleBooster trial and the limited range of intervention weeks (6 to 13). Whilst we can speculate that a longer study with a more polished app would have had a greater impact on the targeted proximal and distal variables, this remains to be determined.

6.4. Concluding comments

In Chapter 1, I described how we can consider individual variation in language skills in autism to be the product of multiple interactive risk factors, some shared with other neurodevelopmental disorders, some autism-specific. I argued against the assumption that structural language difficulties are primarily a downstream effect of core autism features. The task of identifying causal factors for language impairment in autism is hampered by the huge diversity of language profiles, and the breadth of potential influencing factors (motor, memory, environment, attention, perception, consolidation). However, investigating specific causal factors is crucial to designing theoretically sound and practically useful interventions. Taken together, findings from my longitudinal study represent counter-evidence to the claim that all aspects of autistic language impairment relate to differences in core social skills. Further research is necessary to replicate these findings and to investigate whether predictive relationships remain the same over a longer time period in this cohort.

Establishing temporal precedence between speech and later language is the first step in demonstrating a causal role for speech in expressive language development, but does not however equate to causation. Intervention studies are ideally needed to provide direct evidence of a causal role for speech skills. If successful, this would suggest that early assessment and identification of speech-motor difficulties for autistic children is important and could foster a better understanding of individual barriers to communication amongst parents, clinicians and educators.

In Chapter 2, I reviewed the current intervention literature targeting expressive language and speech skills in autistic children. I concluded that very few intervention trials were designed for minimally verbal children, studies targeting speech production required a high degree of dosage and intensity, and that high quality evidence using robust analysis methods was lacking. The longitudinal findings in this thesis lend weight to the emerging drive to design interventions specifically targeting speech skills in this population (e.g. Chenausky et al. 2018).

The qualitative feasibility lessons learned in Chapter 4 can contribute to this process, and the quantitative methods employed to analyse the single cases, as reported in Chapter 5, may be applied to future speech interventions, in autism or any special population where samples are small, difficult to recruit and heterogenous.

Key questions for future research are 1) what causes speech skills to develop more slowly in some autistic children; 2) do speech skills have a causal influence on expressive language development or are they merely a by-product of attenuated social interest or motivation and 3) can speech or other predictor variables explain individual variation over the longer term in this cohort.

If we aim to help those autistic children most at risk of persistent expressive language difficulties, we need to understand the drivers of language growth more precisely and ensure that our conclusions are based on research evidence that includes this population, so that findings can be generalised and additional barriers to communication identified and addressed. We should also be mindful that even amongst those who meet minimally verbal criteria there is heterogeneity in skills, and thus a need to avoid one-size-fits all explanations or solutions.

In this thesis I advocate designing interventions to improve speech skills with the ultimate goal of aiding oral language development. This is not to impose a normative agenda on autistic individuals, but to give them sufficient oral language skills to participate, advocate for themselves, express needs and desires, and by doing so, improve quality of life. For some, alternative methods of communication may be preferred and/or necessary, but having an evidence-based speech-focused intervention increases individual choice in selecting the best supports available.

7. References

- Abrahams, B. S., Tentler, D., Perederiy, J. V., Oldham, M. C., Coppola, G., & Geschwind, D. H. (2007). Genome-wide analyses of human perisylvian cerebral cortical patterning. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(45), 17849–17854. <https://doi.org/10.1073/pnas.0706128104>
- Adamo, J. (1996). *A comparison of linguistic errors in autistic and DLD preschoolers*. Yeshiva University, New York.
- Adams, L. (1998). Oral-motor and motor-speech characteristics of children with autism. *Focus on Autism and Other Developmental Disabilities*, *13*(2), 108–112.
- Adi-Japha, E., Strulovich-Schwartz, O., & Julius, M. (2011). Delayed motor skill acquisition in kindergarten children with language impairment. *Research in Developmental Disabilities*, *32*(6), 2963–2971. <https://doi.org/10.1016/j.ridd.2011.05.005>
- Alarcon, M., Abrahams, B.S., Stone, J.L., Duvall, J.A., Perederiy, J.V., Bomar, J.M., ... Geschwind, D.H. (2008). Linkage, association, and gene-expression analyses identify CNTNAP2 as an autism-susceptibility gene. *American Journal of Human Genetics*, *82*, 150–159.
- Alfonsson, S., Englund, J., & Parling, T. (2019). Tailored Text Message Prompts to Increase Therapy Homework Adherence: A Single-Case Randomised Controlled Study. *Behaviour Change*, 180–91. <https://doi.org/10.1017/bec.2019.10>.
- Allison, D. B., & Gorman, B. S. (1993). Calculating Effect Sizes for Meta-Analysis: The Case of the Single Case. *Behaviour Research and Therapy*, *31*(6), 621–31. [https://doi.org/10.1016/0005-7967\(93\)90115-B](https://doi.org/10.1016/0005-7967(93)90115-B).
- Almirall, D., Distefano, C., Chang, Y., Shire, S., Lu, X., Nahum-shani, I., ... Kasari, C. (2016). Longitudinal Effects of Adaptive Interventions With a Speech-Generating Device in Minimally Verbal Children With ASD Longitudinal Effects of Adaptive Interventions With a Speech-Generating Device in Minimally Verbal Children With ASD. *Journal of Clinical Child & Adolescent Psychology*, *45*(4), 442–456. <https://doi.org/10.1080/15374416.2016.1138407>
- Altwater-Mackensen, N., & Grossmann, T. (2014). Learning to Match Auditory and Visual Speech Cues: Social Influences on Acquisition of Phonological Categories. *Child Development*, *00*(0), 1–17. <https://doi.org/10.1111/cdev.12320>
- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders (DSM-V)*. Arlington, VA: American Psychiatric Association.
- Anderson, D. K., Lord, C., Risi, S., DiLavore, P. S., Shulman, C., Thurm, A., ... Pickles, A. (2007). Patterns of growth in verbal abilities among children with autism spectrum disorder. *Journal of Consulting and Clinical Psychology*, *75*(4), 594–604. <https://doi.org/10.1037/0022-006X.75.4.594>
- Arunachalam, S., & Luyster, R. J. (2016). The integrity of lexical acquisition mechanisms in autism spectrum disorders: A research review. *Autism Research*, *9*(8), 810–828. <https://doi.org/10.1002/aur.1590>

- Bagaiolo, L. F., Mari, J. D. J., Bordini, D., Ribeiro, T. C., Martone, M. C. C., Caetano, S. C., ... Paula, C. S. (2017). Procedures and compliance of a video modeling applied behavior analysis intervention for Brazilian parents of children with autism spectrum disorders. <https://doi.org/10.1177/1362361316677718>
- Bailey, A., Palferman, S., Heavey, L., & Le Couteur, A. (1998). Autism: The phenotype in relatives. *Journal of Autism and Developmental Disorders*, 28(5), 369–392. <https://doi.org/10.1023/A:1026048320785>
- Bak, M. Y. S., Plavnick, J. B., & Byrne, S. M. (2019). Vocalizations of minimally verbal children with autism spectrum disorder across the school year. *Autism*, 23(2), 371–382. <https://doi.org/10.1177/1362361317747576>
- Baker-Ericzén, M. J., Brookman-Frazee, L., & Stahmer, A. (2005). Stress levels and adaptability in parents of toddlers with and without autism spectrum disorders. *Research and Practice for Persons with Severe Disabilities*, 30(4), 194–204.
- Baker, E., & McLeod, S. (2011). Evidence-Based Practice for Children with Speech Sound Disorders: Part 1 Narrative Review. *Language, Speech, and Hearing Services in Schools*, 42(2), 102–139. [https://doi.org/10.1044/0161-1461\(2010/09-0075\)b](https://doi.org/10.1044/0161-1461(2010/09-0075)b)
- Bal, V. H., Fok, M., Lord, C., Smith, I. M., Mirenda, P., Szatmari, P., ... Zaidman-Zait, A. (2019). Predictors of longer-term development of expressive language in two independent longitudinal cohorts of language-delayed preschoolers with Autism Spectrum Disorder. *Journal of Child Psychology and Psychiatry*. <https://doi.org/10.1111/jcpp.13117>
- Bal, V. H., Katz, T., Bishop, S. L., & Krasileva, K. (2016). Understanding definitions of minimally verbal across instruments: Evidence for subgroups within minimally verbal children and adolescents with autism spectrum disorder. *Journal of Child Psychology and Psychiatry and Allied Disciplines*. <https://doi.org/10.1111/jcpp.12609>
- Baldwin, D. A., & Moses, L. J. (2001). Links between Social Understanding and Early Word Learning: Challenges to Current Accounts. *Social Development*, 10(3), 309–329. <https://doi.org/10.1111/1467-9507.00168>
- Ballaban-Gil, K., & Tuchman, R. (2000). Epilepsy and epileptiform EEG: Association with autism and language disorders. *Mental Retardation and Developmental Disabilities Research Reviews*, 6(4), 300–308. [https://doi.org/10.1002/1098-2779\(2000\)6:4<300::AID-MRDD9>3.0.CO;2-R](https://doi.org/10.1002/1098-2779(2000)6:4<300::AID-MRDD9>3.0.CO;2-R)
- Baranek, G. T., Watson, L. R., Boyd, B. A., Poe, M. D., David, F. J., & McGuire, L. (2013). Hyporesponsiveness to social and nonsocial sensory stimuli in children with autism, children with developmental delays, and typically developing children. *Development and Psychopathology*, 25(2), 307–320. <https://doi.org/10.1017/S0954579412001071>
- Barger, B. D., Campbell, J. M., & McDonough, J. D. (2013). Prevalence and Onset of Regression within Autism Spectrum Disorders: A Meta-analytic Review. *Journal of Autism and Developmental Disorders*, 43, 817–828. <https://doi.org/10.1007/s10803-012-1621-x>
- Baril, E. M., & Humphreys, B. P. (2017). An Evaluation of the Research Evidence on the Early Start Denver Model. *Journal of Early Intervention*, 39(4), 321–338. <https://doi.org/10.1177/1053815117722618>

- Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, 68, 255–278.
- Barry, J. G., Yasin, I., & Bishop, D. V. (2007). Heritable risk factors associated with language impairments. *Genes Brain and Behavior*, 6, 66–76.
- Bartak, L., Rutter, M., & Cox, A. (1975). A comparative study of infantile autism and specific development receptive language disorder. I. The children. *British Journal of Psychiatry*, 126, 127–145.
- Bartolucci, G., Pierce, S. J., & Streiner, D. (1980). Cross-sectional studies of grammatical morphemes in autistic and mentally retarded children. *Journal of Autism and Developmental Disorders*, 10(1), 39–50. <https://doi.org/10.1007/BF02408431>
- Bartolucci, G., Pierce, S., Streiner, D., & Eppel, P. T. (1976). Phonological investigation of verbal autistic and mentally retarded subjects. *Journal of Autism and Childhood Schizophrenia*, 6(4), 303–316. <https://doi.org/10.1007/BF01537908>
- Barton, E. E., Pustejovsky, J. E., Maggin, D. M., & Reichow, B. (2017). Technology-Aided Instruction and Intervention for Students With ASD: A Meta-Analysis Using Novel Methods of Estimating Effect Sizes for Single-Case Research. *Remedial and Special Education*, 38 (6). 371–86. <https://doi.org/10.1177/0741932517729508>.
- Barton, S. (2006). Which clinical studies provide the best evidence? *British Medical Journal*, 321. 255–2256.
- Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67(1), 1-48.
- Bellini S & Akullian J. (2007) A meta-analysis of video modeling and video self-modeling interventions for children and adolescents with autism spectrum disorders. *Teaching Exceptional Children*, 73. 264-287.
- Belmonte, M. K., Saxena-Chandhok, T., Cherian, R., Muneer, R., George, L., & Karanth, P. (2013). Oral motor deficits in speech-impaired children with autism. *Frontiers in Integrative Neuroscience*, 7(July), 47. <https://doi.org/10.3389/fnint.2013.00047>
- Bernard-Opitz, V., Sriram, N., & Sapuan, S. (1999). Enhancing vocal imitations in children with autism using the IBM SpeechViewer. *Autism*, 3, 131–147.
- Betancur, C. (2011). Etiological heterogeneity in autism spectrum disorders: More than 100 genetic and genomic disorders and still counting. *Brain Research*, 1380, 42–77. <https://doi.org/10.1016/j.brainres.2010.11.078>
- Bhat, A. N., Galloway, J. C., & Landa, R. J. (2012). Infant Behavior and Development Relation between early motor delay and later communication delay in infants at risk for autism. *Infant Behavior and Development*, 35(4). 838–846. <https://doi.org/10.1016/j.infbeh.2012.07.019>
- Bishop, D. V. M. (2002). Motor immaturity and specific speech and language impairment: Evidence for a common genetic basis. *American Journal of Medical Genetics*, 114(1). 56–63. <https://doi.org/10.1002/ajmg.1630>
- Bishop, D. V. M. (2010). Which neurodevelopmental disorders get researched and why? *PLoS ONE*, 5. <https://doi.org/10.1371/journal.pone.0015112>

- Bishop, D. V. M. (2014). Problems with tense marking in children with specific language impairment: not how but when. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 369(1634), 20120401–20120401. <https://doi.org/10.1098/rstb.2012.0401>
- Bishop, D. V. M. (2017). Why is it so hard to reach agreement on terminology? The case of developmental language disorder (DLD). *International Journal of Language and Communication Disorders*.
- Bishop, D. V. M., Holt, G., Whitehouse, A. J. O., & Groen, M. (2014). No population bias to left-hemisphere language in 4-year-olds with language impairment. *PeerJ*, 2, e507. <https://doi.org/10.7717/peerj.507>
- Bishop, D. V. M., & Norbury, C. F. (2002). Exploring the borderlands of autistic disorder and specific language impairment: A study using standardized diagnostic instruments. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 43(7), 917–929. <https://doi.org/10.1111/1469-7610.00114>
- Bishop, D. V. M., & Norbury, C. F. (2005). Executive functions in children with communication impairments, in relation to autistic symptomatology. 2: Response inhibition. *Autism: The International Journal of Research and Practice*, 9(1), 29–43. <https://doi.org/10.1177/1362361305049028>
- Bloom, L., Rocissano, L., & Hood, L. (1976). Adult–child discourse: developmental interaction between information processing and linguistic knowledge. *Cognitive Psychology*, 8, 521–52.
- Bogdashina, O. (2011) Sensory perceptual issues in autism: why we should listen to those who experience them. *Annales Universitatis Paedagogicae Cracoviensis. Studia Psychologica*, 1, 145-160.
- Bondy, A., & Frost, L. (2001). The picture exchange communication system. *Behav Mod*, 25, 725–44.
- Bornstein, M. H., & Tamis-LeMonda, C. (1997). Maternal Responsiveness and Infant language abilities: specific predictive relations. *Infant Behaviour and Development*, 20(3), 283–296.
- Bornstein, M. H., & Putnick, D. L. (2012). Stability of language in childhood: A multiage, multidomain, multimeasure, and multisource study. *Developmental Psychology*, 48(2), 477–491. <https://doi.org/10.1037/a0025889>
- Bornstein, M. H., Hahn, C. S., Putnick, D. L., & Suwalsky, J. T. D. (2014). Stability of core language skill from early childhood to adolescence: A latent variable approach. *Child Development*, 85(4), 1346–1356. <https://doi.org/10.1111/cdev.12192>
- Bornstein, M. H., Hahn, C.-S., Putnick, D. L., & Pearson, R. M. (2018). Stability of core language skill from infancy to adolescence in typical and atypical development. *Science Advances*, 4(11), eaat7422. <https://doi.org/10.1126/sciadv.aat7422>
- Bottema-Beutel, K. (2016). Associations between joint attention and language in autism spectrum disorder and typical development: A systematic review and meta-regression analysis. *Autism Research*, 9(10), 1021–1035. <https://doi.org/10.1002/aur.1624>

- Botting, N., & Conti-Ramsden, G. (2003). Autism, primary pragmatic difficulties, and specific language impairment: can we distinguish them using psycholinguistic markers? *Developmental Medicine and Child Neurology*, 45(8), 515–526.
- Boucher, J. (2012). Research review: Structural language in autistic spectrum disorder - Characteristics and causes. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 53(3), 219–233. <https://doi.org/10.1111/j.1469-7610.2011.02508.x>
- Boucher, J., & Anns, S. (2018). Memory, learning and language in autism spectrum disorder. *Autism & Developmental Language Impairments*, 3, 239694151774207. <https://doi.org/10.1177/2396941517742078>
- Boucher, J., Bigham, S., Mayes, A., & Muskett, T. (2008). Recognition and language in low functioning autism. *Journal of Autism and Developmental Disorders*, 38(7), 1259–1269. <https://doi.org/10.1007/s10803-007-0508-8>
- Boucher, J., Mayes, A., & Bigham, S. (2012). Memory in autistic spectrum disorder. *Psychological Bulletin*, 138(3), 458–496. <https://doi.org/10.1037/a0026869>
- Boyd, T. K., Hart Barnett, J. E., & More, C. M. (2015). Evaluating iPad Technology for Enhancing Communication Skills of Children With Autism Spectrum Disorders. *Intervention in School and Clinic*, 51(1), 19–27. <https://doi.org/10.1177/1053451215577476>
- Bradshaw, J., Koegel, L. K., & Koegel, R. L. (2017). Improving Functional Language and Social Motivation with a Parent-Mediated Intervention for Toddlers with Autism Spectrum Disorder. *Journal of Autism and Developmental Disorders*, 47(8), 2443–2458. <https://doi.org/10.1007/s10803-017-3155-8>
- Braun, V., & Clark, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3, 77–101.
- Brehme, D. (2014). Perspectives on Personal Pronoun Reversal in Children with ASD: A Critical Review. *Journal of European Psychology Students*. Retrieved from <http://jeps.efpsa.org/article/view/jeps.br/120>
- Briganti, A. M., & Cohen, L. B. (2011). Examining the role of social cues in early word learning. *Infant Behavior & Development*, 34(1), 211–214. <https://doi.org/10.1016/j.infbeh.2010.12.012>
- Brignell, A., Chenausky, K. V., Song, H., Zhu, J., Suo, C., & Morgan, A.T. (2018). Communication interventions for autism spectrum disorder in minimally verbal children. *Cochrane Database of Systematic Reviews* 2018, Issue 11. Art. No.: CD012324. DOI: 10.1002/14651858.CD012324.pub2.
- Broome, K., McCabe, P., Docking, K., & Doble, M. (2017). A systematic review of speech assessments for children with autism spectrum disorder: Recommendations for best practice. *American Journal of Speech-Language Pathology*, 26(3). 1011–1029. https://doi.org/10.1044/2017_AJSLP-16-0014
- Brown, J., Aczel, B., Jiménez, L., Kaufman, S. B., & Grant, K. P. (2010). Intact implicit learning in autism spectrum conditions. *Quarterly Journal of Experimental Psychology (2006)*, 63(February), 1789–1812. <https://doi.org/10.1080/17470210903536910>

- Brosnan, M., Parsons, S., Good, J., & Yuill, N. (2016). How can participatory design inform the design and development of innovative technologies for autistic communities? *Journal of Assistive Technologies*, 10(2), 115–120.
- Brunsdon, V. E. A., & Happé, F. (2014). Exploring the “fractionation” of autism at the cognitive level. *Autism*, 18(1), 17–30. <https://doi.org/10.1177/1362361313499456>
- Bulté, I., & Onghena, P. (2019) Single-case Randomization Tests SCRT-package. 1–18. <https://doi.org/10.22237/jmasm/1383280020>>.License.
- Bulté, I., & Onghena, P. (2009). Randomization Tests for Multiple-Baseline Designs: An Extension of the SCRT-R Package. *Behavior Research Methods*, 41(2), 477–85. <https://doi.org/10.3758/BRM.41.2.477>.
- Cafiero, J. (2001). The Effect of an Augmentative Communication Intervention on the Communication, Behavior, and Academic Program of an Adolescent with Autism. *Focus on Autism and Other Developmental Disabilities*, 16(3), 179–189.
- Calet, N., Pérez-Morenilla, M. C., & De los Santos-Roig, M. (2019). Overcoming Reading Comprehension Difficulties through a Prosodic Reading Intervention: A Single-Case Study. *Child Language Teaching and Therapy*, 35(1), 75–88. <https://doi.org/10.1177/0265659019826252>.
- Carr, E. G., & Durand, V. M. (1985). Reducing behavior problems through functional communication training. *Journal of Applied Behavior Analysis*, 18, 111–126.
- Carter, A. S., Messinger, D. S., Stone, W. L., Celimli, S., Nahmias, A. S., & Yoder, P. (2011). A randomized controlled trial of Hanen’s “More Than Words” in toddlers with early autism symptoms. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 52(7), 741–52. <https://doi.org/10.1111/j.1469-7610.2011.02395.x>
- Charlop-Christy, M., Carpenter, M., Le, L., LeBlanc, L., & Kellet, K. (2002). Using the picture exchange communication system (pecs) with children with autism: assessment of pecs acquisition, speech, social-communicative behavior, and problem behavior. *Journal of Applied Behavior Analysis*, 35(3), 213–231.
- Charman, T., Baron-Cohen, S., Swettenham, J., Baird, G., Cox, A., & Drew, A. (2003). Predicting language outcome in infants with autism and pervasive developmental disorder. *International Journal of Language and Communication Disorders*, 38(3), 265–285
- Chenausky, K., Brignell, A., Morgan, A., & Tager-Flusberg, H. (2019). Motor speech impairment predicts expressive language in minimally verbal, but not low verbal, individuals with autism spectrum disorder. *Autism & Developmental Language Impairments*, 4, 239694151985633. <https://doi.org/10.1177/2396941519856333>
- Chenausky, K., Nelson III, C., & Tager-Flusberg, H. (2017). Vocalization Rate and Consonant Production in Toddlers at High and Low Risk for Autism, 60(April), 865–876.
- Chenausky, K. V., Norton, A. C., & Schlaug, G. (2017). Auditory-Motor Mapping Training in a More Verbal Child with Autism. *Frontiers in Human Neuroscience*, 11, 1–9. <https://doi.org/10.3389/fnhum.2017.00426>

- Chenausky, K., Norton, A., Tager-Flusberg, H., & Schlaug, G. (2016). Auditory-Motor Mapping Training: Comparing the Effects of a Novel Speech Treatment to a Control Treatment for Minimally Verbal Children with Autism. *PLoS ONE*, 1–22. <https://doi.org/10.1371/journal.pone.0164930>
- Chenausky, K., Norton, A., Tager-Flusberg, H., & Schlaug, G. (2018). Behavioral predictors of improved speech output in minimally verbal children with autism. *International Society for Autism Research*, 1–10. <https://doi.org/10.1002/aur.2006>
- Cleland, J., Gibbon, F. E., Peppé, S. J. E., O'Hare, A., & Rutherford, M. (2010). Phonetic and phonological errors in children with high functioning autism and Asperger syndrome. *International Journal of Speech-Language Pathology*, 12(1), 69–76. <https://doi.org/10.3109/17549500903469980>
- Coady, J. A., & Evans, J. L. (2008). Uses and interpretations of non-word repetition tasks in children with and without specific language impairments (SLI). *International Journal of Language and Communication Disorders*, 43(1), 1–40. <https://doi.org/10.1080/13682820601116485>
- Cohen, J. (1977). *Statistical power analysis for the behavioral sciences (Revised ed.)*. New York; London: Academic Press.
- Cohen, J., & Cohen, P. (1984). *Applied multiple regression / correlation analysis for the behavioral sciences (2nd ed.)*. Hillsdale, NJ: Erlbaum
- Concato, J., Shah, N., & Horwitz, R. (2000). Randomized, Controlled Trials, Observational Studies, and the Hierarchy of Research Designs. *The New England Journal of Medicine*, 342(25), 1887–93.
- Conti-Ramsden, G., Crutchley, A., & Botting, N. (1997). The extent to which psychometric tests differentiate subgroups of children with SLI. *Journal of Speech, Language, and Hearing Research*, 40, 765–777. <https://doi.org/10.1044/jslhr.4004.765>
- Conti-Ramsden, G., Botting, N., Simkin, Z., & Knox, E. (2001). Follow-up of children attending infant language units: outcomes at 11 years of age. *International Journal of Language & Communication Disorders / Royal College of Speech & Language Therapists*. <https://doi.org/10.1080/1368282001001988>
- Conti-Ramsden, G., Simkin, Z., & Botting, N. (2006). The prevalence of autistic spectrum disorders in adolescents with a history of specific language impairment (SLI). *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 47(6), 621–628. <https://doi.org/10.1111/j.1469-7610.2005.01584.x>
- Copple, K., Koul, R., Banda, D., & Frye, E. (2015). An examination of the effectiveness of video modelling intervention using a speech-generating device in preschool children at risk for autism. *Developmental Neurorehabilitation*, 18(2), 104–112. <https://doi.org/10.3109/17518423.2014.880079>
- Craig, P., Dieppe, P., Macintyre, S., Michie, S., Nazareth, I., & Petticrew, M. *Developing and Evaluating Complex Interventions*. (2006). <https://mrc.ukri.org/documents/pdf/complex-interventions-guidance/>

- Curtin, S., & Vouloumanos, A. (2013). Speech preference is associated with autistic-like behavior in 18-month-olds at risk for autism spectrum disorder. *Journal of Autism and Developmental Disorders*, 43(9), 2114–2120. <https://doi.org/10.1007/s10803-013-1759-1>
- D’Mello, A. M., Moore, D. M., Crocetti, D., Mostofsky, S. H., & Stoodley, C. J. (2016). Cerebellar gray matter differentiates children with early language delay in autism. *Autism Research*, (March), 1191–1204. <https://doi.org/10.1002/aur.1622>
- Dababnah, S., & Parish, S. L. (2016). Feasibility of an empirically based program for parents of preschoolers with autism spectrum disorder. *Autism*, 20(1), 85–95. <https://doi.org/10.1177/1362361314568900>
- Dada, S., & Alant, E. (2009). The Effect of Aided Language Stimulation on Vocabulary Acquisition in Children With Little or No Functional Speech. *American Journal of Speech-Language Pathology*, 18(1), 50–64.
- Dawson, G., Meltzoff, A. N., Osterling, J., Rinaldi, J., & Brown, E. (1998). Children with autism fail to orient to naturally occurring social stimuli. *Journal of Autism and Developmental Disorders*, 28(6), 479–485. <https://doi.org/10.1023/A:1026043926488>
- Dawson, G., Toth, K., Abbott, R., Osterling, J., Munson, J., Estes, A., & Liaw, J. (2004). Early Social Attention Impairments in Autism: Social Orienting, Joint Attention, and Attention to Distress. *Developmental Psychology*, 40(2), 271–283.
- de Barbaro, K., Johnson, C. M., Forster, D., & Deák, G. O. (2013). Methodological Considerations For Investigating the Microdynamics of Social Interaction Development. *IEEE Transactions on Autonomous Mental Development*, 5(3), 258–270.
- De Fossé, L., Hodge, S. M., Makris, N., Kennedy, D. N., Caviness, V. S., McGrath, L., ... Harris, G. J. (2004). Language-association cortex asymmetry in autism and specific language impairment. *Annals of Neurology*, 56(6), 757–766. <https://doi.org/10.1002/ana.20275>
- de Guibert, C., Maumet, C., Jannin, P., Ferré, J. C., Tréguier, C., Barillot, C., ... Biraben, A. (2011). Abnormal functional lateralization and activity of language brain areas in typical specific language impairment (developmental dysphasia). *Brain*, 134(10), 3044–3058. <https://doi.org/10.1093/brain/awr141>
- de Vries, R. M., Hartogs, B. M. A., & Morey, R. D. (2015). A Tutorial on Computing Bayes Factors for Single-Subject Designs. *Behavior Therapy*, 46 (6), 809–23. <https://doi.org/10.1016/j.beth.2014.09.013>.
- Delinicolás, E. K., & Young, R. L. (2007). Joint attention, language, social relating, and stereotypical behaviours in children with autistic disorder. *Autism: The International Journal of Research and Practice*, 11(5), 425–436. <https://doi.org/10.1177/1362361307079595>
- Desmottes, L., Meulemans, T., & Maillart, C. (2016). Implicit Spoken Words and Motor Sequences Learning Are Impaired in Children with Specific Language Impairment. *Journal of the International Neuropsychological Society: JINS*, 22(5), 520–529. <https://doi.org/10.1017/S135561771600028X>
- Dewey, M. (2019). *Metap: meta-analysis of significance values*. R package version 1.1.

- Dispaldro, M., Leonard, L. B., Corradi, N., Ruffino, M., Bronte, T., & Facchetti, A. (2013). Visual attentional engagement deficits in children with Specific Language Impairment and their role in real-time language processing. *Cortex*, 49(8), 2126–2139. <https://doi.org/10.1016/j.cortex.2012.09.012>
- Distefano, C., Shih, W., Kaiser, A., Landa, R., & Kasari, C. (2016). Communication growth in minimally verbal children with ASD: The importance of interaction. *Autism Research*, (January), 1093–1102. <https://doi.org/10.1002/aur.1594>
- Dongliang, G. E. (2009). *MetaP: A program to combine p values*. Available from http://igm.cumc.columbia.edu/MetaP/metap_.php.
- Drager, K. D. R., Postal, V. J., Carrolus, L., Castellano, M., Gagliano, C., & Glynne, J. (2006). The Effect of Aided Language Modeling on Symbol Comprehension and Production in 2 Preschoolers With Autism. *American Journal of Speech-Language Pathology*, 15, 112–125.
- Droucker, D., Curtin, S., & Vouloumanos, A. (2013). Linking infant-directed-speech and face preferences to language outcomes in infants at risk for autism spectrum disorder. *Journal of Speech, Language, and Hearing Research*, 56(April 2013), 567–576. [https://doi.org/10.1044/1092-4388\(2012/11-0266\)speech](https://doi.org/10.1044/1092-4388(2012/11-0266)speech)
- Dziuk, M., Larson, J.C., Apostu, A., Mahone, E.M., Denckla, M.B., & Mostofsky, S.H. (2007). Dyspraxia in autism: Association with motor, social, and communicative deficits. *Developmental Medicine & Child Neurology*, 49, 734–739.
- Eadie, P. A., Fey, M. E., Douglas, J. M., & Parsons, C. L. (2002). Profiles of grammatical morphology and sentence imitation in children with specific language impairment and Down syndrome. *Journal of Speech, Language, and Hearing Research*, 45, 720–732.
- Ebbels, S. H. (2017). Intervention research: Appraising study designs, interpreting findings and creating research in clinical practice. *International Journal of Speech-Language Pathology*, 19(3), 218–231. <https://doi.org/10.1080/17549507.2016.1276215>
- Edmunds, S. R., Kover, S. T., & Stone, W. L. (2019). The Relation between Parent Verbal Responsiveness and Child Communication in Young Children with or at Risk for Autism Spectrum Disorder: A Systematic Review and Meta-Analysis. *Autism Research*, 12, 715–731. <https://doi.org/10.1002/aur.2100>
- Eicher, J. D., & Gruen, J. R. (2015). Language impairment and dyslexia genes influence language skills in children with autism spectrum disorders. *Autism Research*, 8(2), 229–234. <https://doi.org/10.1002/aur.1436>
- Eigsti, I. M., & Bennetto, L. (2009). Grammaticality judgments in autism: deviance or delay. *Journal of Child Language*, 36, 999–1021. <https://doi.org/10.1017/S0305000909009362>
- Eigsti, I. M., Bennetto, L., & Dadlani, M. B. (2007). Beyond pragmatics: Morphosyntactic development in autism. *Journal of Autism and Developmental Disorders*, 37, 1007–1023. <https://doi.org/10.1007/s10803-006-0239-2>
- ELAN (Version 5.4) Computer software. Nijmegen: Max Planck Institute for Psycholinguistics. Retrieved from <https://tla.mpi.nl/tools/tla-tools/elan/> (on 12th December 2018).

- Elliott, S. N., & Treuting, M. V. B. (1997). The Behavior Intervention Rating Scale: Development and validation of a pretreatment acceptability and effectiveness measure. *Journal of School Psychology, 29*, 43–51.
- Elsabbagh, M., Hohenberger, A., Campos, R., Van Herwegen, J., Serres, J., De Schonen, S., ... Karmiloff-Smith, A. (2013). Narrowing perceptual sensitivity to the native language in infancy: Exogenous influences on developmental timing. *Behavioral Sciences, 3*(1), 120–132. <https://doi.org/10.3390/bs3010120>
- Enders, C. K. (2010). Applied missing data analysis. New York, NY: The Guilford Press.
- Esch, B. E., Carr, J. E., & Grow, L. L. (2009). Evaluation of an enhanced stimulus-stimulus pairing procedure to increase early vocalizations of children with autism. *Journal of Applied Behavior Analysis, 42*(2), 225–241. <https://doi.org/10.1901/jaba.2009.42-225>
- Esch, B. E., Carr, J. E., & Michael, J. (2005). Evaluating Stimulus-Stimulus Pairing and Direct Reinforcement in the Establishment of an Echoic Repertoire of Children Diagnosed with Autism. *The Analysis of Verbal Behavior, 21*, 43–58.
- Estes, A., Olson, E., Sullivan, K., Greenon, J., Winter, J., Dawson, G., & Munson, J. (2013). Parenting-related stress and psychological distress in mothers of toddlers with autism spectrum disorders. *Brain and Development, 35*(2), 133–138. <https://doi.org/10.1016/j.braindev.2012.10.004>
- Evans, J., Saffran, J., & Robe-Torres, K. (2009). Statistical learning in Children With Specific Language Impairment. *Journal of Speech, Language & Hearing Research, 52*, 321–335. [https://doi.org/10.1044/1092-4388\(2011/10-0298\)](https://doi.org/10.1044/1092-4388(2011/10-0298))
- Evans, J., Selinger, C., & Pollak, S. D. (2011). P300 as a measure of processing capacity in auditory and visual domains in specific language impairment. *Brain Research, 1389*, 93–102. <https://doi.org/10.1016/j.brainres.2011.02.010>
- Fatemi, S. H., Aldinger, K. a., Ashwood, P., Bauman, M. L., Blaha, C. D., Blatt, G. J., ... Welsh, J. P. (2012). Consensus paper: Pathological role of the cerebellum in Autism. *Cerebellum, 11*(3), 777–807. <https://doi.org/10.1007/s12311-012-0355-9>
- Fenson, L., Marchman, V. A., Thal, D. J., Dale, P. S., Reznick, J. S., & Bates, E. (2007). MacArthur-Bates Communicative Development Inventories: User's guide and technical manual (2nd ed.). Baltimore, MD: Brookes.
- Ferron, J., & Sentovich, C. (2002). Statistical Power of Randomization Tests Used with Multiple-Baseline Designs. *Journal of Experimental Education, 70*(2), 165–78. <https://doi.org/10.1080/00220970209599504>.
- Field, A. (2013). *Discovering statistics using IBM SPSS (4th Edition)*. London: Sage.
- Field, C., Allen, M. L., & Lewis, C. (2016). Attentional Learning Helps Language Acquisition Take Shape for Atypically Developing Children, Not Just Children with Autism Spectrum Disorders. *Journal of Autism and Developmental Disorders, 46*, 3195–3206. <https://doi.org/10.1007/s10803-015-2401-1>
- Fisher, S. E., & Scharff, C. (2009). FOXP2 as a molecular window into speech and language. *Trends in Genetics, 25*(4), 166–177. <https://doi.org/10.1016/j.tig.2009.03.002>

- Flagg, E. J., Cardy, J. E., Roberts, W., & Roberts, T.P. (2005) Language lateralization development in children with autism: insights from the late field magnetoencephalogram. *Neuroscience Letters*, 386, 82–7.
- Fletcher, F. E., Knowland, V., Walker, S., Gaskell, M. G., Norbury, C., & Henderson, L. M. (2019). Atypicalities in sleep and semantic consolidation in autism. *Developmental Science*, (March), 1–16. <https://doi.org/10.1111/desc.12906>
- Fletcher-Watson, S., & Happé, F. (2019). *Autism: a new introduction to psychological theory and current debates*. Routledge, Oxon.
- Fletcher-Watson, S., Pain, H., Hammond, S., Humphry, A., & McConachie, H. (2016). Designing for young children with autism spectrum disorder: A case study of an iPad app. *International Journal of Child-Computer Interaction*. <https://doi.org/10.1016/j.ijcci.2016.03.002>
- Fletcher-Watson, S., Petrou, A., Scott-Barrett, J., Dicks, P., Graham, C., O'Hare, A., ... McConachie, H. (2015). A trial of an iPad™ intervention targeting social communication skills in children with autism. *Autism: The International Journal of Research and Practice*, 1362361315605624-. <https://doi.org/10.1177/1362361315605624>
- Fletcher-Watson, S., Leekam, S. R., Benson, V., Frank, M. C., & Findlay, J. M. (2009). Eye-movements reveal attention to social information in autism spectrum disorder. *Neuropsychologia*, 47, 248–257.
- Floris, D. L., Lai, M., Auer, T., Lombardo, M. V., Ecker, C., Chakrabarti, B., ... Murphy, D. G. M. (2016). Atypically Rightward Cerebral Asymmetry in Male Adults With Autism Stratifies Individuals With and Without Language Delay, 253, 230–253. <https://doi.org/10.1002/hbm.23023>
- Fortunato-Tavares, T., Andrade, C. R. F., Befi-Lopes, D., Limongi, S. O., Fernandes, F. D. M., & Schwartz, R. G. (2015). Syntactic comprehension and working memory in children with specific language impairment, autism or Down syndrome. *Clinical Linguistics & Phonetics*, 29(7), 499–522. <https://doi.org/10.3109/02699206.2015.1027831>
- Frazier, T. W., Strauss, M., Klingemier, E. W., Zetzer, E. E., Hardan, A. Y., Eng, C., & Youngstrom, E. A. (2017). A Meta-Analysis of Gaze Differences to Social and Nonsocial Information Between Individuals With and Without Autism. *Journal of the American Academy of Child and Adolescent Psychiatry*, 56(7), 546–555. <https://doi.org/10.1016/j.jaac.2017.05.005>
- Gabriel, A., Maillart, C., Stefaniak, N., Lejeune, C., Desmottes, L., & Meulemans, T. (2013). Procedural learning in specific language impairment: effects of sequence complexity. *Journal of the International Neuropsychological Society: JINS*, 19(2013), 264–271. <https://doi.org/10.1017/S1355617712001270>
- Gage, N. M., Juraneck, J., Filipek, P. A., Osann, K., Flodman, P., Isenberg, A. L., & Spence, M. A. (2009). Rightward hemispheric asymmetries in auditory language cortex in children with autistic disorder: An MRI investigation. *Journal of Neurodevelopmental Disorders*, 1(3), 205–214. <https://doi.org/10.1007/s11689-009-9010-2>

- Gamliel, I., Yirmiya, N., Jaffe, D., Manor, O., & Sigman, M. (2009). Developmental trajectories in siblings of children with autism: Cognition and language from 4 months to 7 years. *Journal of Autism and Developmental Disorders*, 39, 1131–1144
- Ganz, J. B., Earles-Vollrath, T. L., Heath, A. K., Parker, R. I., Rispoli, M. J., & Duran, J. B. (2012). A meta-analysis of single case research studies on aided augmentative and alternative communication systems with individuals with autism spectrum disorders. *Journal of Autism and Developmental Disorders*, 42, 60–74. <https://doi.org/10.1007/s10803-011-1212-2>
- Gathercole, S. E. (2006). Nonword repetition and word learning the nature of the relationship. *Applied Psycholinguistics*, 27, 513–543.
- Gernsbacher, M. A., Geye, H. and Ellis Wesimer, S. (2005) The role of language and communication impairments within autism. In: P. Fletcher, & J. C. Miller. (Eds.) *Language Disorders and Developmental Theory* (pp.73–93). Philadelphia, PA: John Benjamins.
- Gernsbacher, M. A., Sauer, E. A., Geye, H. M., Schweigert, E. K., & Hill Goldsmith, H. (2008). Infant and toddler oral- and manual-motor skills predict later speech fluency in autism. *Journal of Child Psychology and Psychiatry*, 49(1), 43–50. <https://doi.org/10.1111/j.1469-7610.2007.01820.x>
- Gerow, S., Hagan-Burke, S., Rispoli, M., Gregori, E., Mason, R., & Ninci, J. (2018). A Systematic Review of Parent-Implemented Functional Communication Training for Children With ASD. *Behavior Modification*, 42(3), 335–363. <https://doi.org/10.1177/0145445517740872>
- Gerrard, S., & Rugg, G. (2009). Sensory impairments and autism: A re-examination of causal modelling. *Journal of Autism and Developmental Disorders*, 39(10), 1449–1463. <https://doi.org/10.1007/s10803-009-0773-9>
- Gervais, H., Belin, P., Boddaert, N., Leboyer, M., Coez, A., Sfaello, I., ... Zilbovicius, M. (2004). Abnormal cortical voice processing in autism. *Nature Neuroscience* 7, 801–802.
- Geurts, H., & Embrechts, M. (2008). Language profiles in ASD, SLI, and ADHD. *Journal of Autism and Developmental Disorders*, 38, 1931–1943.
- Gevarter, C., O'Reilly, M. F., Kuhn, M., Mills, K., Ferguson, R., & Watkins, L. (2016). Increasing the vocalizations of individuals with autism during intervention with a speech-generating device. *Journal of Applied Behavior Analysis*, 49, 17–33. <https://doi.org/10.1002/jaba.270>
- Girolametto, L., Weitzman, E., Wiigs, M., & Pearce, P. S. (1999). The Relationship Between Maternal Language Measures and Language Development in Toddlers With Expressive Vocabulary Delays. *American Journal of Speech-Language Pathology*, 8(November 1999), 364–374.
- Goin-Kochel, R. P., Esler, A. N., Kanne, S. M., & Hus, V. (2014). Developmental regression among children with autism spectrum disorder : Onset, duration, and effects on functional outcomes. *Research in Autism Spectrum Disorders*, 8(7), 890–898. <https://doi.org/10.1016/j.rasd.2014.04.002>
- Goldstein, H. (2002). Communication intervention for children with autism: A review of treatment efficacy. *Journal of Autism and Developmental Disorders*, 32(5), 373–396. <https://doi.org/10.1023/A:1020589821992>

- Gönen, M. (2004). Sample Size and Power for McNemar's Test with Clustered Data. *Statistics in Medicine*, 23, 2283–94. <https://doi.org/10.1002/sim.1768>.
- Gooch, D., Hulme, C., Nash, H. M., & Snowling, M. J. (2014). Comorbidities in preschool children at family risk of dyslexia. *Journal of Child Psychology and Psychiatry*, 55(3), 237–246. <https://doi.org/10.1111/jcpp.12139>
- Gooch, D., Thompson, P., Nash, H. M., Snowling, M. J., & Hulme, C. (2016). The development of executive function and language skills in the early school years. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 57(2), 180–187. <https://doi.org/10.1111/jcpp.12458>
- Goods, K. S., Ishijima, E., Chang, Y. C., & Kasari, C. (2013). Preschool based JASPER intervention in minimally verbal children with Autism: Pilot RCT. *Journal of Autism and Developmental Disorders*, 43(5), 1050–1056. <https://doi.org/10.1007/s10803-012-1644-3>
- Gordon, B., & Stark, S. (2007). Procedural learning of a visional sequence in individuals with Autism. *Focus on Autism and Other Developmental Disabilities*, 22(1), 14–22. <https://doi.org/10.1177/10883576070220010201>
- Graham, J. W. (2009). Missing data analysis: Making it work in the real world. *Annual Review of Psychology*, 60, 549–576.
- Grandin, T. (1996). *Thinking in Pictures and Other Reports from My Life with Autism*. New York: Vintage Books.
- Green, J., Charman, T., McConachie, H., Aldred, C., Slonims, V., Howlin, P., ... Pickles, A. (2010). Parent-mediated communication-focused treatment in children with autism (PACT): a randomised controlled trial. *The Lancet*, 375(9732), 2152–2160. [https://doi.org/10.1016/S0140-6736\(10\)60587-9](https://doi.org/10.1016/S0140-6736(10)60587-9)
- Gros-Louis, J. (2014). Maternal responsiveness and the development of directed vocalizing in social interactions. *Infancy*, 19(4), 385–408.
- Grynszpan, O., Weiss, P. L., Perez-Diaz, F., & Gal, E. (2014). Innovative technology-based interventions for autism spectrum disorders: A meta-analysis. *Autism*, 18(4), 346–361. <https://doi.org/10.1177/1362361313476767>
- Guillon, Q., Hadjikhani, N., Baduel, S., & Rogé, B. (2014). Visual social attention in autism spectrum disorder: Insights from eye tracking studies. *Neuroscience and Biobehavioral Reviews*, 42, 279–297. <https://doi.org/10.1016/j.neubiorev.2014.03.013>
- Haardörfer, R., & Gagné, P. (2010). The Use of Randomization Tests in Single-Subject Research. *Focus on Autism and Other Developmental Disabilities*, 25(1), 47–54. <https://doi.org/10.1177/1088357609350745>.
- Haebig, E., Kaushanskaya, M., & Ellis Weismer, S. (2015). Lexical Processing in School-Age Children with Autism Spectrum Disorder and Children with Specific Language Impairment: The Role of Semantics. *Journal of Autism and Developmental Disorders*, 45(12), 4109–4123. <https://doi.org/10.1007/s10803-015-2534-2>
- Haesen, B., Boets, B., & Wagemans, J. (2011). A review of behavioural and electrophysiological studies on auditory processing and speech perception in autism spectrum disorders. *Research in Autism Spectrum Disorders*, 5(2), 701–714. <https://doi.org/10.1016/j.rasd.2010.11.006>

- Hamilton, A., Plunkett, K., & Schafer, G., (2000). Infant vocabulary development assessed with a British Communicative Development Inventory: Lower scores in the UK than the USA. *Journal of Child Language*, 27, 689-705.
- Hampton, L. H., & Kaiser, A. P. (2016). Intervention effects on spoken-language outcomes for children with autism: A systematic review and meta-analysis. *Journal of Intellectual Disability Research*, 60(5), 444–463. <https://doi.org/10.1111/jir.12283>
- Hancock, T. B., & Kaiser, A. P. (2002). The effects of trainer- implemented enhanced milieu teaching on the social communication of children with autism. *Topics in Early Childhood Special Education*, 22, 39–54.
- Hanley, M., Riby, D. M., McCormack, T., Carty, C., Coyle, L., Crozier, N., ... McPhillips, M. (2014). Attention during social interaction in children with autism: Comparison to specific language impairment, typical development, and links to social cognition. *Research in Autism Spectrum Disorders*, 8(7), 908–924. <https://doi.org/10.1016/j.rasd.2014.03.020>
- Hardan, A. Y., Gengoux, G. W., Berquist, K. L., Libove, R. A., Ardel, M., Phillips, J., ... Minjarez, M. B. (2015). A randomized controlled trial of Pivotal Response Treatment Group for parents of children with autism. *Journal of Child Psychology and Psychiatry*, 56(8), 884–892. <https://doi.org/10.1111/jcpp.12354>
- Hari, R., & Renvall, H. (2001). Impaired processing of rapid stimulus sequences in dyslexia. *Trends in Cognitive Science*, 5(12), 525-532.
- Hart, K. I., Fujiki, M., Brinton, B., & Hart, C. H. (2004). The relationship between social and severity of language impairment. *Journal of Speech, Language, and Hearing Research*, 47(3), 647–662. [https://doi.org/10.1044/1092-4388\(2004/050\)](https://doi.org/10.1044/1092-4388(2004/050))
- Hayes, S. A., & Watson, S. L. (2013). The impact of parenting stress: A meta-analysis of studies comparing the experience of parenting stress in parents of children with and without autism spectrum disorder. *Journal of Autism and Developmental Disorders*, 43(3), 629–642. <https://doi.org/10.1007/s10803-012-1604-y>
- Heaton, P., Hudry, K., Ludlow, A., & Hill, E. L. (2008). Superior discrimination of speech pitch and its relationship to verbal ability in autism spectrum disorders. *Cognitive Neuropsychology*, 25(6), 771–782. <https://doi.org/10.1080/02643290802336277>
- Hedges, L. V. (1981). Distribution theory for Glass's estimator of effect size and related estimators. *Journal of Educational Statistics*, 6, 107–128.
- Hedges, L. V., Pustejovsky, J. E., & Shadish, W. R. (2012). A standardized mean difference effect size for single-case designs. *Research Synthesis Methods*, 3, 224–239.
- Hedges, L. V., Pustejovsky, J. E., & Shadish, W. R. (2013). A Standardized Mean Difference Effect Size for Multiple Baseline Designs across Individuals. *Research Synthesis Methods* 4, 324–41. <https://doi.org/10.1002/jrsm.1086>.
- Henderson, L., Powell, A., Gaskell, G. M., & Norbury, C. (2014). Learning and consolidation of new spoken words in autism spectrum disorder. *Developmental Science*, 17(6), 858–871. <https://doi.org/10.1111/desc.12169>

- Henry, L. A., Messer, D. J., & Nash, G. (2012). Executive functioning in children with specific language impairment. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 53(1), 37–45. <https://doi.org/10.1111/j.1469-7610.2011.02430.x>
- Herbert, M. R., Ziegler, D. A., Deutsch, C. K., O'Brien, L. M., Kennedy, D. N., Filipek, P. A., ... Caviness, V. S. (2005). Brain asymmetries in autism and developmental language disorder: A nested whole-brain analysis. *Brain*, 128(1), 213–226. <https://doi.org/10.1093/brain/awh330>
- Herringshaw, A. J., Ammons, C. J., DeRamus, T. P., & Kana, R. K. (2016). Hemispheric differences in language processing in autism spectrum disorders: A meta-analysis of neuroimaging studies. *Autism Research : Official Journal of the International Society for Autism Research*, (January), 1–12. <https://doi.org/10.1002/aur.1599>
- Heymann, P., Northrup, J. B., West, K. L., Parad, M. V., Leezenbaum, N. B., & Iverson, J. M. (2018). Coordination is key : Joint attention and vocalisation in infant siblings of children with Autism Spectrum Disorder, 1007–1020. <https://doi.org/10.1111/1460-6984.12418>
- Heyvaert, M., Wendt, O., Van den Noortgate, W., & Onghena, P. (2015). Randomization and Data-Analysis Items in Quality Standards for Single-Case Experimental Studies. *Journal of Special Education*, 49 (3), 146–56. <https://doi.org/10.1177/0022466914525239>.
- Hill, A. P., van Santen, J., Gorman, K., Langhorst, B. H., & Fombonne, E. (2015). Memory in language-impaired children with and without autism. *Journal of Neurodevelopmental Disorders*, 7(1), 19. <https://doi.org/10.1186/s11689-015-9111-z>
- Hill, E. L. (2001). Non-specific nature of specific language impairment: a review of the literature with regard to concomitant motor impairments, 36(2), 149–171. <https://doi.org/10.1080/1368282001001987>
- Hodge, S. M., Makris, N., Kennedy, D. N., Caviness, V. S., Howard, J., McGrath, L., ... Harris, G. J. (2010). Cerebellum, language, and cognition in autism and specific language impairment. *Journal of Autism and Developmental Disorders*, 40(3), 300–316. <https://doi.org/10.1007/s10803-009-0872-7>
- Hoff, E. (2003). The specificity of environmental influence: socioeconomic status affects early vocabulary development via maternal speech. *Child Development*, 74(5), 1368–1378.
- Honaker, J., King, G., & Blackwell, M. (2011). Amelia II: A Program for Missing Data. *Journal of Statistical Software* 45(7), 1-47. URL <http://www.jstatsoft.org/v45/i07/>.
- Hoogeboom, T. J., Kwakkenbos, L., Rietveld, L., Den Broeder, A. A., De Bie, R. A., & Van Den Ende, C. H. M. (2012). Feasibility and Potential Effectiveness of a Non-Pharmacological Multidisciplinary Care Programme for Persons with Generalised Osteoarthritis: A Randomised, Multiple-Baseline Single-Case Study. *BMJ Open* 2 (4). <https://doi.org/10.1136/bmjopen-2012-001161>.
- Hooton, J. W. L. (1991). Randomization Tests: Statistics for Experimenters. *Computer Methods and Programs in Biomedicine*, 35 (1), 43–51. [https://doi.org/10.1016/0169-2607\(91\)90103-Z](https://doi.org/10.1016/0169-2607(91)90103-Z).
- Horner, R. H., Carr, E. G., Halle, J., Mcgee, G., Odom, S., & Wolery, M. (2005). The Use of Single-Subject Research to Identify Evidence-Based Practice in Special Education. *Exceptional Children*, 71 (2), 165–79. <https://doi.org/10.1177/001440290507100203>.

- Hosozawa, A. M., & Tanaka, K. (2012). How Children With Specific Language Impairment View Social Situations : An Eye Tracking Study, *129*(6). <https://doi.org/10.1542/peds.2011-2278>
- Howlin, P. (2003). Outcome in high-functioning adults with autism with and without early language delays: implications for the differentiation between autism and Asperger syndrome. *Autism and Developmental Disorders*, *33*(1), 3–13.
- Howlin, P. (2005). Outcomes in autism spectrum disorders. In F. R. Volkmar, A. Klin, R. Paul, & D. J. Cohen (Eds.), *Handbook of autism and pervasive developmental disorders* (pp. 201–222). Hoboken, NJ: Wiley.
- Howlin, P., Gordon, R. K., Pasco, G., Wade, A., & Charman, T. (2007). The effectiveness of Picture Exchange Communication System (PECS) training for teachers of children with autism : a pragmatic , group randomised controlled trial. *Journal of Child Psychology and Psychiatry*, *48*(5), 473–481. <https://doi.org/10.1111/j.1469-7610.2006.01707.x>
- Howlin, P., Mawhood, L., & Rutter, M. (2000). Autism and developmental receptive language disorder - A comparative follow-up in early adult life. I: Cognitive and language outcomes. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *41*(5), 547–559. <https://doi.org/10.1017/S002196309900579X>
- Hsu, H. J., & Bishop, D. V. M. (2014). Training understanding of reversible sentences: a study comparing language-impaired children with age-matched and grammar-matched controls. *PeerJ*, *2*, e656. <https://doi.org/10.7717/peerj.656>
- Hsu, H. J., Tomblin, J. B., & Christiansen, M. H. (2014). Impaired statistical learning of non-adjacent dependencies in adolescents with specific language impairment. *Frontiers in Psychology*, *5*(MAR), 1–10. <https://doi.org/10.3389/fpsyg.2014.00175>
- Hudry, K., Chandler, S., Bedford, R., Pasco, G., Gliga, T., Elsabbagh, M., ... Charman, T. (2014). Early language profiles in infants at high-risk for autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *44*(1), 154–167. <https://doi.org/10.1007/s10803-013-1861-4>
- Hudson, S., Levickis, P., Down, K., Nicholls, R., & Wake, M. (2014). Maternal responsiveness predicts child language at ages 3 and 4 in a community-based sample of slow-to-talk toddlers. *International Journal of Language & Communication Disorders / Royal College of Speech & Language Therapists*, *0*(0), 1–7. <https://doi.org/10.1111/1460-6984.12129>
- Hwang, Y., Levin, J. R., & Johnson, E.W. (2018). Pictorial Mnemonic-Strategy Interventions for Children with Special Needs: Illustration of a Multiply Randomized Single-Case Crossover Design. *Developmental Neurorehabilitation*, *21* (4), 223–37. <https://doi.org/10.3109/17518423.2015.1100689>.
- Ingersoll, B., & Gergans, S. (2007). The effect of a parent-implemented imitation intervention on spontaneous imitation skills in young children with autism. *Research in Developmental Disabilities*, *28*(2), 163–75. <https://doi.org/10.1016/j.ridd.2006.02.004>
- Ingersoll, B., & Schreibman, L. (2006). Teaching reciprocal imitation skills to young children with autism using a naturalistic behavioral approach: Effects on language, pretend play, and joint attention. *Journal of Autism and Developmental Disorders*, *36*(4), 487–505. <https://doi.org/10.1007/s10803-006-0089-y>

- Ingram, D. G., Takahashi, T. N., & Miles, J. H. (2008). Defining Autism Subgroups: A Taxometric Solution. *Journal of Autism and Developmental Disorders*, 38(5), 950–960. <https://doi.org/10.1007/s10803-007-0469-y>
- James, E., Gaskell, M. G., Weighall, A., & Henderson, L. (2017). Consolidation of vocabulary during sleep: The rich get richer? *Neuroscience and Biobehavioral Reviews*, 77, 1–13. <https://doi.org/10.1016/j.neubiorev.2017.01.054>
- Jamshidi, L., Heyvaert, M., Declercq, L., Fernández-Castilla, B., Ferron, J. M., Moeyaert, M., ... Van den Noortgate, W. (2018). Methodological quality of meta-analyses of single-case experimental studies. *Research in Developmental Disabilities*, 79(October 2017), 97–115. <https://doi.org/10.1016/j.ridd.2017.12.016>
- Järvinen-Pasley, A., Pasley, J., & Heaton, P. (2008). Is the linguistic content of speech less salient than its perceptual features in autism? *Journal of Autism and Developmental Disorders*, 38(2), 239–248. <https://doi.org/10.1007/s10803-007-0386-0>
- Jaswal, V. K., & Akhtar, N. (2018). Being vs. Appearing Socially Uninterested: Challenging Assumptions about Social Motivation in Autism. *Behavioral and Brain Sciences*, 1–84. <https://doi.org/10.1017/S0140525X18001826>
- Johnson, M. H. (2012). Executive function and developmental disorders: The flip side of the coin. *Trends in Cognitive Sciences*, 16(9), 454–457. <https://doi.org/10.1016/j.tics.2012.07.001>
- Johnson, J. S., & Newport, E. L. (1989). Critical Period Effects in Second Language Learning: The influence of Maturational State on the Acquisition of English as a Second Language. *Cognitive Psychology*, 21, 60–99.
- Jones, C. R. G., Happé, F., Baird, G., Simonoff, E., Marsden, A. J. S., Tregay, J., ... Charman, T. (2009). Auditory discrimination and auditory sensory behaviours in autism spectrum disorders. *Neuropsychologia*, 47(13), 2850–2858. <https://doi.org/10.1016/j.neuropsychologia.2009.06.015>
- Jones, W. P. (2003). Single-Case Time Series with Bayesian Analysis: A Practitioner's Guide. *Measurement and Evaluation in Counseling and Development*, 36 (1), 28–39. <https://doi.org/10.1080/07481756.2003.12069078>.
- Jones, W., Carr, K., & Klin, A. (2008). Absence of preferential looking to the eyes of approaching adults predicts level of social disability in 2-year-old toddlers with autism spectrum disorder. *Arch. Gen. Psychiatry* 65, 946–954. doi: 10. 1001/archpsyc.65.8.946
- Jones, W., & Klin, A. (2013). Attention to eyes is present but in decline in 2-6-month-old infants later diagnosed with autism. *Nature*, 504(7480), 427–431. <https://doi.org/10.1038/nature12715>
- Just, M. A., Cherkassky, V. L., Keller, T. A., & Minshew, N. J. (2004). Cortical activation and synchronization during sentence comprehension in high-functioning autism: Evidence of under-connectivity. *Brain*, 127, 1811–1821.
- Kaganovich, N. (2017). Sensitivity to Audiovisual Temporal Asynchrony in Children With a History of Specific Language Impairment and Their Peers With Typical Development: A Replication and Follow-Up Study. *Journal of Speech, Language, and Hearing Research*, 60, 2259–2270.
- Kaganovich, N., Shumaker, J., Leonard, L., Gustafson, D., & Macias, D. (2014). Children with a history of SLI show reduced sensitivity to audiovisual temporal asynchrony: an ERP study.

- Journal of Speech, Language, and Hearing Research*, 57(June), 1679–1691.
<https://doi.org/10.1044/2014>
- Kagohara, D. M., Meer, L. Van Der, Ramdoss, S., Reilly, M. F. O., Lancioni, G. E., Davis, T. N., ... Sigafoos, J. (2013). Research in Developmental Disabilities Review article Using iPods 1 and iPads 1 in teaching programs for individuals with developmental disabilities: A systematic review. *Research in Developmental Disabilities*, 34(1), 147–156.
<https://doi.org/10.1016/j.ridd.2012.07.027>
- Kane, M., Connell, J. E., & Pellecchia, M. (2010). A quantitative analysis of language interventions for children with autism. *Behavior Analyst Today*, 11(2), 128–144.
- Kasari, C., Brady, N., Lord, C., & Tager-flusberg, H. (2013). Assessing the Minimally Verbal School-Aged Child With Autism Spectrum Disorder. *Autism Research*, 6, 479–493.
<https://doi.org/10.1002/aur.1334>
- Kasari, C., Gulsrud, A., Freeman, S., Paparella, T., & Hellemann, G. (2012). Longitudinal follow-up of children with autism receiving targeted interventions on joint attention and play. *Journal of the American Academy of Child and Adolescent Psychiatry*, 51(5), 487–495.
<https://doi.org/10.1016/j.jaac.2012.02.019>
- Kasari, C., Kaiser, A., Goods, K., Nietfeld, J., Mathy, P., Landa, R., ... Almirall, D. (2014a). Communication interventions for minimally verbal children with autism: A sequential multiple assignment randomized trial. *Journal of the American Academy of Child and Adolescent Psychiatry*, 53(6), 635–646. <https://doi.org/10.1016/j.jaac.2014.01.019>
- Kasari, C., Lawton, K., Shih, W., Barker, T. V., Landa, R., Lord, C., ... Senturk, D. (2014b). Caregiver-mediated intervention for low-resourced preschoolers with autism: an RCT. *Pediatrics*, 134(1), e72-9. <https://doi.org/10.1542/peds.2013-3229>
- Kaufman, N.R. (1995). *The Kaufman Speech Praxis Test for Children*. Detroit: Wayne State University Press.
- Kazdin, A. E. (1980). Obstacles in using randomization tests in single-case experimentation". *Journal of Educational Statistics*, 5, 253-260.
- Kazdin, A. E. (2019). Single-Case Experimental Designs. Evaluating Interventions in Research and Clinical Practice. *Behaviour Research and Therapy*, 117, 3–17.
<https://doi.org/10.1016/j.brat.2018.11.015>.
- Kemeny, F., & Lukacs, A. (2010). Impaired procedural learning in language impairment: results from probabilistic categorization. *Journal of Clinical and Experimental Neuropsychology*, 32(June 2015), 249–258. <https://doi.org/10.1080/13803390902971131>
- Kendall, J. M. (2003). Designing a Research Project: Randomised Controlled Trials and Their Principles. *Emergency Medicine Journal*, 20 (2), 164–68. <https://doi.org/10.1136/emj.20.2.164>.
- Kenny, L., Hattersley, C., Molins, B., Buckley, C., Povey, C., & Pellicano, E. (2016). Which terms should be used to describe autism? Perspectives from the UK autism community. *Autism*, 20(4), 442–462. <https://doi.org/10.1177/1362361315588200>

- Key, A. P., Yoder, P. J., & Stone, W. L. (2016). Consonant differentiation mediates the discrepancy between non-verbal and verbal abilities in children with ASD. *Journal of Intellectual Disability Research*, 60(5), 478–490. <https://doi.org/10.1111/jir.12286>
- Kjelgaard, M. M., & Tager-Flusberg, H. (2001). An investigation of language impairment in autism: Implications for genetic subgroups. *Language and Cognitive Processes*, 16(March 2015), 287–308. <https://doi.org/10.1080/01690960042000058>
- Kjellmer, L., Fernell, E., & Gillberg, C. (2018). Speech and language profiles in 4- to 6-year-old children with early diagnosis of autism spectrum disorder without intellectual disability. *Neuropsychiatric Disease and Treatment*, 14, 2415–2427.
- Klin, A. (1991). Young autistic children's listening preferences in regard to speech: A possible characterization of the symptom of social withdrawal. *Journal of Autism and Developmental Disorders*, 21(1), 29–42. <https://doi.org/10.1007/BF02206995>
- Klin, A. (1992). Listening Preferences in Regard to Speech in Four Children with Developmental Disabilities. *Journal of Child Psychology and Psychiatry*, 33(4), 763–769. <https://doi.org/10.1111/j.1469-7610.1992.tb00911.x>
- Klin, A., Jones, W., Schultz, R., Volkmar, F., & Cohen, D. (2002). Visual fixation patterns during viewing of naturalistic social situations as predictors of social competence in individuals with autism. *Archives of General Psychiatry*, 59(9), 809–816. <https://doi.org/10.1001/archpsyc.59.9.809>
- Koegel, L. K., Koegel, R. L., Harrower, J. K., & Carter, C. M. (1999). Pivotal response intervention I: Overview of approach. *Research and Practice for Persons with Severe Disabilities*, 24(3), 174–185. doi:10.2511/rpsd.24.3.174
- Koegel, R. L., Camarata, S., Koegel, L. K., Ben-Tall, A., & Smith, A. E. (1998). Increasing speech intelligibility in children with autism. *Journal of Autism and Developmental Disorders*, 28(3), 241–251. <https://doi.org/10.1023/A:1026073522897>
- Koegel, R. L., O'Dell, M., & Dunlap, G. (1988). Producing speech use in nonverbal autistic children by reinforcing attempts. *Journal of Autism and Developmental Disorders*, 18(4), 525–538. <https://doi.org/10.1007/BF02211871>
- Koo, T. K., & Li, M. Y. (2016). A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *Journal of Chiropractic Medicine*, 15(2), 155–163. <https://doi.org/10.1016/j.jcm.2016.02.012>
- Kourea, L., & Lo, Y. Y. (2016). The educational validity and utility of single-case design research in building evidence-based practices in education. *International Journal of Research and Method in Education*, 39(4), 349–364. <https://doi.org/10.1080/1743727X.2016.1160278>
- Kratochwill, T. R., Hitchcock, J., Horner, R. H., Levin, J. R., Odom, S. L., Rindskopf, D. M., & Shadish, W. R. (2010). Single-case designs technical documentation. Retrieved from <https://ies.ed.gov/ncee/wwc/Document/229>
- Kratochwill, T. R., & Levin, J. R. (2010). Enhancing the Scientific Credibility of Single-Case Intervention Research: Randomization to the Rescue. *Psychological Methods*, 15 (2), 124–44. <https://doi.org/10.1037/a0017736>.

- Kuhl, P. K. (2007). Is speech learning “gated” by the social brain? *Developmental Science*, *10*(1), 110–120. <https://doi.org/10.1111/j.1467-7687.2007.00572.x>
- Kuhl, P. K., Coffey-Corina, S., Padden, D., & Dawson, G. (2005). Links between social and linguistic processing of speech in preschool children with autism: behavioral and electrophysiological measures. *Developmental Science*, *8*(1), F1–F12. <https://doi.org/10.1111/j.1467-7687.2004.00384.x>
- Kuhl, P. K., Conboy, B. T., Coffey-Corina, S., Padden, D., Rivera-Gaxiola, M., & Nelson, T. (2008). Phonetic learning as a pathway to language: new data and native language magnet theory expanded (NLM-e). *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *363*(1493), 979–1000. <https://doi.org/10.1098/rstb.2007.2154>
- Kuhl, P. K., Ramírez, R. R., Bosseler, A., Lin, J.-F. L., & Imada, T. (2014). Infants’ brain responses to speech suggest analysis by synthesis. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(31), 11238–11245. <https://doi.org/10.1073/pnas.1410963111>
- Kuhl, P. K., Tsao, F.-M., & Liu, H.-M. (2003). Foreign-language experience in infancy: effects of short-term exposure and social interaction on phonetic learning. *Proceedings of the National Academy of Sciences of the United States of America*, *100*(15), 9096–9101. <https://doi.org/10.1073/pnas.1532872100>
- Kujala, T., Lepistö, T., & Näätänen, R. (2013). The neural basis of aberrant speech and audition in autism spectrum disorders. *Neuroscience and Biobehavioral Reviews*, *37*(4), 697–704. <https://doi.org/10.1016/j.neubiorev.2013.01.006>
- Kutner, M. H., Nachtsheim, C. J., & Neter, J. (2004). *Applied linear regression models* (4th ed.). New York: McGraw-Hill.
- Landry, R. & Bryson, S.E. (2004). Impaired disengagement of attention in young children with autism. *Journal of Child Psychology and Psychiatry* *45*, 1115–1122.
- Lane, J. D., Lieberman-Betz, R., & Gast, D. L. (2016). An Analysis of Naturalistic Interventions for Increasing Spontaneous Expressive Language in Children With Autism Spectrum Disorder. *Journal of Special Education*, *50*(1), 49–61. <https://doi.org/10.1177/0022466915614837>
- Lanter, E., Russell, S. D., Kuriakose, A., & Blevins, K. E. (2016). Incorporating AAC and General Instructional Strategies in Requesting Interventions: A Case Study in Down Syndrome. *Communication Disorders Quarterly*, *38*(1), 52–63. <https://doi.org/10.1177/1525740115602093>
- LeBarton, E. S., & Landa, R. J. (2019). Infant motor skill predicts later expressive language and autism spectrum disorder diagnosis. *Infant Behavior and Development*, *54*(March 2018), 37–47. <https://doi.org/10.1016/j.infbeh.2018.11.003>
- Ledbetter-Cho, K., O’Reilly, M., Lang, R., Watkins, L., & Lim, N. (2018). Meta-analysis of Tablet-Mediated Interventions for Teaching Academic Skills to Individuals with Autism. *Journal of Autism and Developmental Disorders*, *48*(9), 3021–3036. <https://doi.org/10.1007/s10803-018-3573-2>
- Lee, J. C., Mueller, K. L., & Tomblin, J. B. (2015). Examining procedural learning and corticostriatal pathways for individual differences in language: testing endophenotypes of DRD2/ANKK1. *Language, Cognition and Neuroscience*, *37*98(October), 1–17.

- Lenroot, R. K., & Yeung, P. K. (2013). Heterogeneity within Autism Spectrum Disorders: What have We Learned from Neuroimaging Studies? *Frontiers in Human Neuroscience*, 7(October), 733. <https://doi.org/10.3389/fnhum.2013.00733>
- Leonard, H. C., & Hill, E. L. (2014). Review: The impact of motor development on typical and atypical social cognition and language: a systematic review. *Child and Adolescent Mental Health*, n/a-n/a. <https://doi.org/10.1111/camh.12055>
- Leonard, H. C., Bedford, R., Pickles, A., & Hill, E. L. (2015). Research in Autism Spectrum Disorders Predicting the rate of language development from early motor skills in at-risk infants who develop autism spectrum disorder, 14, 15–24. <https://doi.org/10.1016/j.rasd.2014.12.012>
- Lepper, T. L., & Petursdottir, A. I. (2017). Effects of response-contingent stimulus pairing on vocalizations of nonverbal children with autism. *Journal of Applied Behavior Analysis*, 50(4), 756–774. <https://doi.org/10.1002/jaba.415>
- Lepper, T. L., Petursdottir, A. I., & Esch, B. E. (2013). Effects of operant discrimination training on the vocalizations of nonverbal children with autism. *Journal of Applied Behavior Analysis*, 46(3), 656–661. <https://doi.org/10.1002/jaba.55>
- Levin, J. R., Ferron, J. M., & Gafurov, B. S. (2017). Additional Comparisons of Randomization-Test Procedures for Single-Case Multiple-Baseline Designs: Alternative Effect Types. *Journal of School Psychology*, 63, 13–34. <https://doi.org/10.1016/j.jsp.2017.02.003>.
- Levin, J. R., Ferron, J. M., & Gafurov, B. S. (2018). Comparison of Randomization-Test Procedures for Single-Case Multiple-Baseline Designs. *Developmental Neurorehabilitation*, 21 (5), 290–311. <https://doi.org/10.1080/17518423.2016.1197708>.
- Levy, Y., & Ebstein, R. P. (2009). Research review: Crossing syndrome boundaries in the search for brain endophenotypes. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 50(6), 657–668. <https://doi.org/10.1111/j.1469-7610.2008.01986.x>
- Leyfer, O. T., Tager-Flusberg, H., Dowd, M., Tomblin, B. J., & Folstein, S. E. (2008). Overlap between autism and specific language impairment: Comparison of autism diagnostic interview and autism diagnostic observation schedule scores. *Autism Research*, 1(5), 284–296. <https://doi.org/10.1002/aur.43>
- Lim, H. A., & Draper, E., (2011) The effects of music therapy incorporated with applied behavior analysis verbal behavior approach for children with autism spectrum disorders. *Journal of Music Therapy*, 48, 532–550.
- Lindblad, T. L. (2012). Clinical corner: how do you increase speech intelligibility (Articulation skills) or the variability in the sounds produced by children with autism spectrum disorders? *Science in Autism Treatment*, 9(3), 3–6.
- Lindgren, K. A., Folstein, S. E., Tomblin, J. B., & Tager-Flusberg, H. (2009). Language and reading abilities of children with autism spectrum disorders and specific language impairment and their first-degree relatives. *Autism Research*, 2(1), 22–38. <https://doi.org/10.1002/aur.63>
- Lord C., Risi S., & Pickles A. (2004) Trajectory of language development in autistic spectrum disorders. In: M. L. Rice, & S. F. Warren (Eds.) *Developmental Language Disorders: From Phenotypes to Etiologies* (pp.7–29). New Jersey: Lawrence Erlbaum Associates, Inc.

- Lord, C., & Paul, R. (1997). Language and communication in autism. In D. J. Cohen, & F. R. Volkmar (Eds.), *Handbook of autism and pervasive developmental disorders* (2nd ed., pp. 195–225). New York: Wiley.
- Loucas, T., Charman, T., Pickles, A., Simonoff, E., Chandler, S., Meldrum, D., & Baird, G. (2008). Autistic symptomatology and language ability in autism spectrum disorder and specific language impairment, *11*, 1184–1192. <https://doi.org/10.1111/j.1469-7610.2008.01951.x>
- Loucas, T., Riches, N., Baird, G., & Pickles, A. (2013). Spoken word recognition in adolescents with autism spectrum disorders and specific language impairment. *Applied*, *34*, 301–322. Retrieved from http://journals.cambridge.org/abstract_S0142716411000701
- Lum, J. A. G., Conti-Ramsden, G., Morgan, A. T., & Ullman, M. T. (2014). Procedural learning deficits in specific language impairment (SLI): A meta-analysis of serial reaction time task performance. *Cortex*, *51*(1), 1–10. <https://doi.org/10.1016/j.cortex.2013.10.011>
- Lum, J. A. G., Conti-Ramsden, G., Page, D., & Ullman, M. T. (2012). Working, declarative and procedural memory in specific language impairment. *Cortex*, *48*(9), 1138–1154. <https://doi.org/10.1016/j.cortex.2011.06.001>
- Luyster, R. J., Kadlec, M. B., Carter, A., & Tager-Flusberg, H. (2008). Language Assessment and Development in Toddlers with Autism Spectrum Disorders. *Journal of Autism and Developmental Disorders*, *38*(8), 1426–1438. <https://doi.org/10.1007/s10803-007-0510-1>
- Luyster, R., & Lord, C. (2009). Word learning in children with autism spectrum disorders. *Developmental Psychology*, *45*(6), 1774–1786. <https://doi.org/10.1037/a0016223>
- MacNeil, L. K., & Mostofsky, S. H. (2012). Specificity of dyspraxia in children with autism. *Neuropsychology*, *26*(2), 165–171. <https://doi.org/10.1037/a0026955>
- Magal, A. (2013). A hypothetical universal model of cerebellar function: reconsideration of the current dogma. *Cerebellum (London, England)*, *12*(5), 758–772. <https://doi.org/10.1007/s12311-013-0477-8>
- Mallet, K. H., Shamloul, R. M., Corbett, D., Finestone, H. M., Hatcher, S., Lumsden, J., ... Dowlatshahi, D. (2016). RecoverNow: Feasibility of a mobile tablet-based rehabilitation intervention to treat post-stroke communication deficits in the acute care setting. *PLoS ONE*, *11*(12), 1–13. <https://doi.org/10.1371/journal.pone.0167950>
- Manolov, R., & Moeyaert, M. (2017). How Can Single-Case Data Be Analyzed? Software Resources, Tutorial, and Reflections on Analysis. *Behavior Modification*, *41* (2), 179–228. <https://doi.org/10.1177/0145445516664307>.
- Markram, H., Rinaldi, T., & Markram, K. (2007). The intense world syndrome--an alternative hypothesis for autism. *Frontiers in Neuroscience*, *1*(1), 77–96. <https://doi.org/10.3389/neuro.01.1.1.006.2007>
- Marshall, C. R., Harcourt-Brown, S., Ramus, F., & van der Lely, H. (2009). The link between prosody and language skills in children with specific language impairment (SLI) and/or dyslexia. *International Journal of Language & Communication Disorders / Royal College of Speech & Language Therapists*, *44*(4), 466–488. <https://doi.org/10.1080/13682820802591643>

- Martin, M. K., Wright, L. E., Perry, S., Cornett, D., Schraeder, M., & Johnson, J. T. (2015). Children with Developmental Verbal Dyspraxia: Changes in articulation and perceived resilience with intensive multimodal intervention. *Child Language Teaching and Therapy*. <https://doi.org/10.1177/0265659015615924>
- Matsuzaki, J., Kuschner, E. S., Blaskey, L., Bloy, L., Kim, M., Ku, M., ... Roberts, T. P. L. (2019). Abnormal auditory mismatch fields are associated with communication impairment in both verbal and minimally verbal/nonverbal children who have autism spectrum disorder. *Autism Research*, (May), 1225–1235. <https://doi.org/10.1002/aur.2136>
- Matyas, T. A., & Greenwood, K. M. (1990). Visual Analysis of Single-Case Time Series: Effects of Variability, Serial Dependence, and Magnitude of Intervention Effects. *Journal of Applied Behavior Analysis*, 23 (3), 341–51. <https://doi.org/10.1901/jaba.1990.23-341>.
- Mawhood, L., Howlin, P., & Rutter, M. (2000). Autism and developmental receptive language disorder—A comparative follow-up in early adult life: I. Cognitive and language outcomes. *Journal of Child Psychology and Psychiatry*, 41, 547–559.
- Mayo, J., & Eigsti, I.-M. (2012). Brief report: a comparison of statistical learning in school-aged children with high functioning autism and typically developing peers. *Journal of Autism and Developmental Disorders*, 42(11), 2476–2485. <https://doi.org/10.1007/s10803-012-1493-0>
- Mayor-Dubois, C., Zesiger, P., Van der Linden, M., & Roulet-Perez, E. (2012). Nondeclarative learning in children with Specific Language Impairment: Predicting regularities in the visuomotor, phonological, and cognitive domains. *Child Neuropsychology*, 7049(June 2015), 1–9. <https://doi.org/10.1080/09297049.2012.734293>
- McCoy, K., & Hermansen, E. (2007). Video Modeling for Individuals with Autism: A Review of Model Types and Effects. *Education and Treatment of Children*, 30(4), 183–213.
- McDaniel, J., D'Ambrose Slaboch, K., & Yoder, P. (2018). A meta-analysis of the association between vocalizations and expressive language in children with autism spectrum disorder. *Research in Developmental Disabilities*, 72, 202–213. <https://doi.org/10.1016/j.ridd.2017.11.010>
- McDuffie, A., & Yoder, P. (2010). Types of Parent Verbal Responsiveness That Predict Language in Young Children With Autism Spectrum Disorder. *Journal of Speech, Language and Hearing Research*, 53, 1026–1039.
- McGillion, M., Herbert, J. S., Pine, J., Vihman, M., dePaolis, R., Keren-Portnoy, T., & Matthews, D. (2017). What Paves the Way to Conventional Language? The Predictive Value of Babble, Pointing, and Socioeconomic Status. *Child Development*, 88(1), 156–166. <https://doi.org/10.1111/cdev.12671>
- McGonigle-Chalmers, M., Alderson-Day, B., Fleming, J., & Mosen, K. (2013). Profound expressive language impairment in low functioning children with autism: An investigation of syntactic awareness using a computerised learning task. *Journal of Autism and Developmental Disorders*, 43, 2062–2081. <https://doi.org/10.1007/s10803-012-1753-z>
- McGregor, K. K., Berns, A. J., Owen, A. J., Michels, S. A., Duff, D., Bahnsen, A. J., & Lloyd, M. (2011). Associations between syntax and the lexicon among children with or without ASD and

- language impairment. *Journal of Autism and Developmental Disorders*, 42(1), 35–47. <http://doi.org/10.1007/s10803-011-1210-4>
- McLeod, S., & Baker, E. (2014). Speech-language pathologists' practices regarding assessment, analysis, target selection, intervention, and service delivery for children with speech sound disorders. *Clinical Linguistics & Phonetics*, 28(7–8), 508–531. <https://doi.org/10.3109/02699206.2014.926994>
- McLeod, S., Baker, E., McCormack, J., Wren, Y., Roulstone, S., Crowe, K., ... Howland, C. (2017). Cluster-Randomized Controlled Trial Evaluating the Effectiveness of Computer-Assisted Intervention Delivered by Educators for Children With Speech Sound Disorders. *Journal of Speech Language and Hearing Research*, 60(7), 1891. https://doi.org/10.1044/2017_JSLHR-S-16-0385
- Michiels, B., & Onghena, P. (2018). Randomized single-case AB phase designs: Prospects and pitfalls. *Behavior Research Methods*. <https://doi.org/10.3758/s13428-018-1084-x>
- Michiels, B., Heyvaert, M., Meulders, A., & Onghena, P. (2017). Confidence intervals for single-case effect size measures based on randomization test inversion. *Behavior Research Methods*, 49(1), 363–381. <https://doi.org/10.3758/s13428-016-0714-4>
- Miller, C. A., Kail, R., Leonard, L. B., & Tomblin, J. B. (2001). Speed of processing in children with specific language impairment. *Journal of Speech, Language, and Hearing Research: JSLHR*, 44(2), 416–433. [https://doi.org/10.1044/1092-4388\(2001/034\)](https://doi.org/10.1044/1092-4388(2001/034))
- Miller, M., Chukoskie, L., Zinni, M., Townsend, J., & Trauner, D. (2014). Dyspraxia, motor function and visual-motor integration in autism. *Behavioural Brain Research*. <https://doi.org/10.1016/j.bbr.2014.04.011>
- Mirenda, P. (2008). A Back Door Approach to Autism and AAC. *Augmentative and Alternative Communication*, 24(3), 220–234. <https://doi.org/10.1080/08990220802388263>
- Misyak, J. B., & Christiansen, M. H. (2012). Statistical Learning and Language: An Individual Differences Study. *Language Learning*, 62(1), 302–331. <https://doi.org/10.1111/j.1467-9922.2010.00626.x>
- Mody, M., Shui, A. M., Nowinski, L. A., Golas, S. B., Ferrone, C., O'Rourke, J. A., & McDougle, C. J. (2017). Communication Deficits and the Motor System: Exploring Patterns of Associations in Autism Spectrum Disorder (ASD). *Journal of Autism and Developmental Disorders*, 47(1), 155–162. <https://doi.org/10.1007/s10803-016-2934-y>
- Moeyaert, M., Ferron, J. M., Beretvas, S. N., & Van den Noortgate, W. (2014). From a single-level analysis to a multilevel analysis of single-case experimental designs. *Journal of School Psychology*, 52(2), 191–211. <https://doi.org/10.1016/j.jsp.2013.11.003>
- Morgan, A. T., & Vogel, A. P. (2009). Intervention for childhood apraxia of speech. *Cochrane Database of Systematic Reviews*. <https://doi.org/10.1002/14651858.CD006278.pub2.www.cochranelibrary.com>
- Morgan, A.T., Murray, E., & Liégeois, F.J. (2018) Interventions for childhood apraxia of speech. *Cochrane Database of Systematic Reviews*, Issue 5. Art. No.: CD006278.

- Moriuchi, J. M., Klin, A., & Jones, W. (2017). Mechanisms of Diminished Attention to Eyes in Autism. *American Journal of Psychiatry*, 174(1), 26–35. <https://doi.org/10.1176/appi.ajp.2016.15091222>
- Mostofsky, S. H., Goldberg, M. C., Landa, R. J., & Denckla, M. B. (2000). Evidence for a deficit in procedural learning in children and adolescents with autism: implications for cerebellar contribution. *Journal of the International Neuropsychological Society: JINS*, 6(7), 752–759. <https://doi.org/10.1017/S1355617700677020>
- Mouridsen, S. E., & Hauschild, K. M. (2009). A longitudinal study of autism spectrum disorders in individuals diagnosed with a developmental language disorder as children. *Child: Care, Health and Development*, 35(5), 691–697. <https://doi.org/10.1111/j.1365-2214.2009.00954.x>
- Mucchetti, C. A. (2013). Adapted shared reading at school for minimally verbal students with autism. *Autism: The International Journal of Research and Practice*, 17(3), 358–72. <https://doi.org/10.1177/1362361312470495>
- Mulhern, T., Lydon, S., Healy, O., Mollaghan, G., Ramey, D., & Leoni, M. (2017). A systematic review and evaluation of procedures for the induction of speech among persons with developmental disabilities. *Developmental Neurorehabilitation*, 20(4), 207–227. <https://doi.org/10.3109/17518423.2016.1150360>
- Mullen, E. M., (1995). Mullen Scales of Early Learning. Circle Pines, MN: American Guidance Service.
- Mundy, P., & Neal, A. (2001). Neural plasticity, joint attention, and a transactional social-orienting model of autism. *International Review of Research in Mental Retardation*, 23, 139–168.
- Mundy, P., Delgado, C., Block, J., Venezia, M., Hogan, A., & Seibert, J. (2003) *A Manual for the Abridged Early Social Communication Scales (ESCS)*. Coral Gables, FL: University of Miami at Florida
- 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. <https://doi.org/10.1186/1471-2431-12-112>
- Murray, E., McCabe, P., & Ballard, K. (2014). A systematic review of treatment outcomes for children with childhood apraxia of speech. *Journal of Speech, Language, and Hearing Research*, 57(June), 1679–1691. <https://doi.org/10.1044/2014>
- Murray, E., McCabe, P., & Ballard, K. (2015). A Randomized Controlled Trial for Children With Childhood Apraxia of Speech Comparing Rapid Syllable Transition Treatment and the Nuffield Dyspraxia Programme – Third Edition. *Journal of Speech Language and Hearing Research*, 58(June). <https://doi.org/10.1044/2015>
- Naigles, L. R., & Tek, S. (2017). ‘Form is easy, meaning is hard’ revisited: (re) characterizing the strengths and weaknesses of language in children with autism spectrum disorder. *Wiley Interdisciplinary Reviews: Cognitive Science*, 8(4), 1–12. <https://doi.org/10.1002/wcs.1438>
- Nakano, T., Tanaka, K., Endo, Y., Yamane, Y., Yamamoto, T., Nakano, Y., ... Kitazawa, S. (2010). Atypical gaze patterns in children and adults with autism spectrum disorders dissociated from developmental changes in gaze behaviour. *Proceedings of the Royal Society*, 277, 2935–2943. <https://doi.org/10.1098/rspb.2010.0587>

- Nation, K., & Hulme, C. (2011). Learning to read changes children's phonological skills: Evidence from a latent variable longitudinal study of reading and nonword repetition. *Developmental Science*, *14*(4), 649–659. <https://doi.org/10.1111/j.1467-7687.2010.01008.x>
- Nicolson, R. I., & Fawcett, A. J. (2007). Procedural learning difficulties: Reuniting the developmental disorders? *Trends in Neurosciences*, *30*(4), 135–141. <https://doi.org/10.1016/j.tins.2007.02.003>
- Ninci, J., Vannest, K. J., Willson, V., & Zhang, N. (2015). Interrater Agreement Between Visual Analysts of Single-Case Data: A Meta-Analysis. *Behavior Modification*, *39*(4), 510–541. <https://doi.org/10.1177/0145445515581327>
- Norbury, C. F. (2014). Sources of variation in developmental language disorders: evidence from eye-tracking studies of sentence production. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *369*(1634), 20120393. <https://doi.org/10.1098/rstb.2012.0393>
- Norbury, C. F., Brock, J., Cragg, L., Einav, S., Griffiths, H., & Nation, K. (2009). Eye-movement patterns are associated with communicative competence in autistic spectrum disorders. *Journal of Child Psychology and Psychiatry*, *50*(7), 834–842. <https://doi.org/10.1111/j.1469-7610.2009.02073.x>
- Norbury, C. F., Gooch, D., Wray, C., Baird, G., Charman, T., Simonoff, E., ... Pickles, A. (2016). The impact of nonverbal ability on prevalence and clinical presentation of language disorder: evidence from a population study. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *57*(11), 1247–1257. <https://doi.org/10.1111/jcpp.12573>
- Norbury, C. F., Vamvakas, G., Gooch, D., Baird, G., Charman, T., Simonoff, E., & Pickles, A. (2017). Language growth in children with heterogeneous language disorders: a population study. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *58*(10), 1092–1105. <https://doi.org/10.1111/jcpp.12793>
- Norrelgen, F., Fernell, E., Eriksson, M., Hedvall, A., Persson, C., Sjolín, M., ... Kjellmer, L. (2014). Children with autism spectrum disorders who do not develop phrase speech in the preschool years. *Autism*, 1–10.
- Norrix, L.W., Plante, E., Vance, R., & Boliek, C.A. (2007). Auditory-visual integration for speech by children with and without specific language impairment. *Journal of Speech, Language and Hearing Research*, *50*, 1639–1651.
- O'Connor, K. (2012). Auditory processing in autism spectrum disorder: A review. *Neuroscience and Biobehavioral Reviews*, *36*(2), 836–854. <https://doi.org/10.1016/j.neubiorev.2011.11.008>
- Obeid, R., Brooks, P. J., Powers, K. L., Gillespie-Lynch, K., & Lum, J. A. G. (2016). Statistical learning in specific language impairment and autism spectrum disorder: A meta-analysis. *Frontiers in Psychology*, *7*(August), 1–18. <https://doi.org/10.3389/fpsyg.2016.01245>
- Obuchowski, N. A. (1998). On the comparison of correlated proportions for clustered data. *Statistics in Medicine*, *17*, 1495–1507.
- Odom, S. L., Barton, E. E., Reichow, B., Swaminathan, H., & Pustejovsky, J. E. (2018). Between-Case Standardized Effect Size Analysis of Single Case Designs: Examination of the Two

- Methods. *Research in Developmental Disabilities*, 79 (May), 88–96. <https://doi.org/10.1016/j.ridd.2018.05.009>.
- Olive, M. L., Lang, R. B., & Davis, T. N. (2008). An analysis of the effects of functional communication and a Voice Output Communication Aid for a child with autism spectrum disorder. *Research in Autism Spectrum Disorders*. <https://doi.org/10.1016/j.rasd.2007.06.002>
- Oller, D. K., Niyogi, P., Gray, S., Richards, J. a, Gilkerson, J., Xu, D., ... Warren, S. F. (2010). Automated vocal analysis of naturalistic recordings from children with autism, language delay, and typical development. *Proceedings of the National Academy of Sciences of the United States of America*, 107(30), 13354–13359. <https://doi.org/10.1073/pnas.1003882107>
- Orsmond, G. I., & Cohn, E. S. (2015). The Distinctive Features of a Feasibility Study. *OTJR: Occupation, Participation and Health*, 35(3), 169–177. <https://doi.org/10.1177/1539449215578649>
- Ozonoff, S., & Iosif, A. M. (2019). Changing conceptualizations of regression: What prospective studies reveal about the onset of autism spectrum disorder. *Neuroscience and Biobehavioral Reviews*, 100(September 2018), 296–304. <https://doi.org/10.1016/j.neubiorev.2019.03.012>
- Ozonoff, S., Iosif, A., Baguio, F., Cook, I. C., Hill, M. M., Hutman, T., ... Young, G. S. (2010). A Prospective Study of the Emergence of Early Behavioral Signs of Autism. *Journal of the American Academy of Child & Adolescent Psychiatry*, 49(3), 256–266. <https://doi.org/10.1016/j.jaac.2009.11.009>
- Parker, R. I., Vannest, K. J., Davis, J. L., & Sauber, S. B. (2011). Combining Nonoverlap and Trend for Single-Case Research: Tau-U. *Behavior Therapy*, 42(2), 284–299. <https://doi.org/10.1016/j.beth.2010.08.006>
- Parsonson, B. S., & Baer, D. M. (1992). The visual analysis of data, and current research into the stimuli controlling it. In T. R. Kratochwill, & J. R. Levin (Eds.), *Single-case research design and analysis* (pp. 15–40). Hillsdale, NJ: Erlbaum.
- Passy J. A. (2003). *Handful of Sounds: cued articulation in practice*. Camberwell, Vic: ACER.
- Paterson, S. J., Brown, J. H., Gsödl, M. K., Johnson, M. H., & Kamiloff-Smith, A. (1999). Cognitive modularity and genetic disorders. *Science*, 286(5448), 2355–2358.
- Patten, E., Belardi, K., Baranek, G. T., Watson, L. R., Labban, J. D., & Oller, D. K. (2014). Vocal patterns in infants with autism spectrum disorder: Canonical babbling status and vocalization frequency. *Journal of Autism and Developmental Disorders*, 44(10), 2413–2428. <https://doi.org/10.1007/s10803-014-2047-4>
- Patten, E., Watson, L., & Yoder, P. (2012). *Motor behaviors and associations with later consonant inventory in nonverbal children with ASD*. Paper presented at the International Meeting for Autism Research, Toronto, CA
- Paul, R., Campbell, D., Gilbert, K., & Tsiouri, I. (2013). Comparing spoken language treatments for minimally verbal preschoolers with autism spectrum disorders. *Journal of Autism and Developmental Disorders*, 43(2), 418–431. <https://doi.org/10.1007/s10803-012-1583-z>

- Paul, R., Chawarska, K., Fowler, C., Cicchetti, D., & Volkmar, F. (2007). "Listen My Children and You Shall Hear": Auditory Preferences in Toddlers With Autism Spectrum Disorders. *October*, 50(October), 1350–1365.
- Paul, R., Fuerst, Y., Ramsay, G., Chawarska, K., & Klin, A. (2011). Out of the mouths of babes: vocal production in infant siblings of children with ASD, 5, 588–598. <https://doi.org/10.1111/j.1469-7610.2010.02332.x>
- Pecukonis, M., Plesa Skwerer, D., Eggleston, B., Meyer, S., & Tager-Flusberg, H. (2019). Concurrent Social Communication Predictors of Expressive Language in Minimally Verbal Children and Adolescents with Autism Spectrum Disorder. *Journal of Autism and Developmental Disorders*, (0123456789). <https://doi.org/10.1007/s10803-019-04089-8>
- Pellicano, E. (2010). Individual differences in executive function and central coherence predict developmental changes in theory of mind in autism. *Developmental Psychology*, 46(2), 530–544. <https://doi.org/10.1037/a0018287>
- Pennington, B. F. (2006). From single to multiple deficit models of developmental disorders. *Cognition*, 101(2), 385–413. <https://doi.org/10.1016/j.cognition.2006.04.008>
- Pennington, B. F., & Bishop, D. V. M. (2009). Relations Among Speech, Language, and Reading Disorders. *Annua Rev Psychol*, 283–306. <https://doi.org/10.1146/annurev.psych.60.110707.163548>
- Peppé, S., Cleland, J., Gibbon, F. E., O'Hare, A., & Castilla, P. M. (2011). Expressive prosody in children with autism spectrum conditions. *Journal of Neurolinguistics*, 24(1), 41–53. <https://doi.org/10.1016/j.jneuroling.2010.07.005>
- Pickard, K. E., & Ingersoll, B. R. (2014). High and Low Level Initiations of Joint Attention, and Response to Joint Attention: Differential Relationships with Language and Imitation. *Journal of Autism and Developmental Disorders*, 45, 262–268. <https://doi.org/10.1007/s10803-014-2193-8>
- Pickett, E., Pullara, O., O'Grady, J., & Gordon, B. (2009). Speech acquisition in older nonverbal individuals with autism: a review of features, methods, and prognosis. *Cognitive and Behavioral Neurology: Official Journal of the Society for Behavioral and Cognitive Neurology*, 22(1), 1–21.
- Pickles, A., Anderson, D. K., & Lord, C. (2014). Heterogeneity and plasticity in the development of language: a 17-year follow-up of children referred early for possible autism. *Journal of Child Psychology and Psychiatry*, 55(12), 1354–1362. <https://doi.org/10.1111/jcpp.12269>
- Pickles, A., Bolton, P., Macdonald, H., Bailey, A., Le Couteur, A., Sim, C. H., & Rutter, M. (1995). Latent-class analysis of recurrence risks for complex phenotypes with selection and measurement error: A twin and family history study of autism. *American Journal of Human Genetics*, 57(3), 717–726.
- Pickles, A., Le Couteur, A., Leadbitter, K., Salomone, E., Cole-Fletcher, R., Tobin, H., ... Green, J. (2016). Parent-mediated social communication therapy for young children with autism (PACT): long-term follow-up of a randomised controlled trial. *The Lancet*, 6736(16). [https://doi.org/10.1016/S0140-6736\(16\)31229-6](https://doi.org/10.1016/S0140-6736(16)31229-6)

- Pickles, A., Simonoff, E., Conti-Ramsden, G., Falcaro, M., Simkin, Z., Charman, T., ... Baird, G. (2009). Loss of language in early development of autism and specific language impairment. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *50*(7), 843–852. <https://doi.org/10.1111/j.1469-7610.2008.02032.x>
- Piven, J., Palmer, P., Landa, R., Santangelo, S., Jacobi, D., & Childress, D. (1997). Personality and language characteristics in parents from multiple-incidence autism families. *American Journal of Medical Genetics*, *74*, 398–411.
- Plante, E. (1991). MRI findings in the parents and siblings of specifically language-impaired boys. *Brain and Language*, *41*(1), 67–80. [https://doi.org/10.1016/0093-934X\(91\)90111-D](https://doi.org/10.1016/0093-934X(91)90111-D)
- Plesa Skwerer, D., Jordan, S. E., Brukilacchio, B. H., & Tager-Flusberg, H. (2015). Comparing methods for assessing receptive language skills in minimally verbal children and adolescents with autism spectrum disorders. *Autism*, *20*(5), 591–604. <https://doi.org/10.1177/1362361315600146>
- Plumb, A., & Wetherby, A. (2013). Vocalization development in toddlers with autism spectrum disorder. *Journal of Speech, Language, and Hearing*, *56*(April 2013), 721–734. [https://doi.org/10.1044/1092-4388\(2012/11-0104\)](https://doi.org/10.1044/1092-4388(2012/11-0104))
- Pons, F., Andreu, L., Sanztorrent, M., Buillegaz, L., & Lewkowicz, D. J. (2013). Erratum: Perception of audio-visual speech synchrony in Spanish-speaking children with and without specific language impairment Corrigendum (Journal of Child Language DOI: 10.1017/S0305000912000189). *Journal of Child Language*, *40*(3), 701. <https://doi.org/10.1017/S0305000913000202>
- Porter, M., & Coltheart, M. (2005). Cognitive Heterogeneity in Williams Syndrome. *Developmental Neuropsychology*, *27*(2), 75–306. <https://doi.org/10.1207/s15326942dn2702>
- Powell, G., Wass, S. V., Erichsen, J. T., Leekam, S. R. (2016). First evidence of the feasibility of gaze-contingent attention training for school children with autism. *Autism*, *20*(8), 927–37.
- Preston, J. L., Leece, M. C., & Maas, E. (2016). Intensive Treatment with Ultrasound Visual Feedback for Speech Sound Errors in Childhood Apraxia. *Frontiers in Human Neuroscience*, *10*(August), 1–9. <https://doi.org/10.3389/fnhum.2016.00440>
- Prizant, B., Wetherby, A., Rubin, E., Laurent, A., & Rydell, P. (2006). *The SCERTS model: A comprehensive educational approach for children with autism spectrum disorders*. Baltimore, MD: Brookes.
- Pustejovsky, J. E. (2016). *scdhlm: Estimating hierarchical linear models for single-case designs*. University of Texas at Austin, Austin, TX. R package version 0.3, <URL: <http://github.com/jepusto/scdhlm>>.
- Pustejovsky, J. E., Hedges, L. V., & Shadish, W. R. (2014). Design-Comparable Effect Sizes in Multiple Baseline Designs: A General Modeling Framework. *Journal of Educational and Behavioral Statistics*, *39* (5), 368–93. <https://doi.org/10.3102/1076998614547577>.
- R Core Team (2017). *R: A language and environment for statistical computing*. R Foundations for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>.

- Ramdoss, S., Lang, R., Mulloy, A., Franco, J., Reilly, M. O., Didden, R., & Lancioni, G. (2011). Use of Computer-Based Interventions to Teach Communication Skills to Children with Autism Spectrum Disorders : A Systematic Review, 55–76. <https://doi.org/10.1007/s10864-010-9112-7>
- Ramus, F., Marshall, C. R., Rosen, S., & Van Der Lely, H. K. J. (2013). Phonological deficits in specific language impairment and developmental dyslexia: Towards a multidimensional model. *Brain*, 136(2), 630–645. <https://doi.org/10.1093/brain/aws356>
- Rapin, I., & Dunn, M. (2003). Update on the language disorders of individuals on the autistic spectrum. *Brain and Development*, 25(3), 166–172. [https://doi.org/10.1016/S0387-7604\(02\)00191-2](https://doi.org/10.1016/S0387-7604(02)00191-2)
- Rapin, I., Dunn, M., Allen, D., Stevens, M. C., & Fein, D. (2009). Subtypes of language disorders in school-age children with autism. *Developmental Neuropsychology*, 34(1), 66–84. <https://doi.org/10.1080/87565640802564648>
- Redcay, E. (2008). The superior temporal sulcus performs a common function for social and speech perception: Implications for the emergence of autism. *Neuroscience and Biobehavioral Reviews*, 32(1), 123–142. <https://doi.org/10.1016/j.neubiorev.2007.06.004>
- Redcay, E., & Courchesne, E. (2008). Deviant Functional Magnetic Resonance Imaging Patterns of Brain Activity to Speech in 2-3-Year-Old Children with Autism Spectrum Disorder. *Biological Psychiatry*, 64(7), 589–598. <https://doi.org/10.1016/j.biopsych.2008.05.020>
- Revelle, W. (2018) psych: Procedures for Personality and Psychological Research, Northwestern University, Evanston, Illinois, USA, <https://CRAN.R-project.org/package=psych> Version = 1.8.12.
- Rice, K., Moriuchi, J. M., Klin, A., & Jones, W. (2012). Parsing Heterogeneity in Autism Spectrum Disorders : Visual Scanning of Dynamic Social Scenes in School-Aged Children. *JAAC*, 51(3), 238–248. <https://doi.org/10.1016/j.jaac.2011.12.017>
- Rice, M. L. (2012). Toward epigenetic and gene regulation models of specific language impairment: looking for links among growth, genes, and impairments. *Journal of Neurodevelopmental Disorders*, 4(1), 27. <https://doi.org/10.1186/1866-1955-4-27>
- Rice, M. L., Warren, S. F., & Betz, S. K. (2005). Language symptoms of developmental language disorders: An overview of autism, Down syndrome, fragile X, specific language impairment, and Williams syndrome. *Applied Psycholinguistics*, 26, 7–27. <https://doi.org/10.1017/S0142716405050034>
- Rice, M. L., Wexler, K., & Hershberger, S. (1998). Tense over time: The longitudinal course of tense acquisition in children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, 41(6), 1412–1431. <https://doi.org/10.1044/jslhr.4106.1412>
- Richardson, J. (2012). *Critical Review: Is PROMPT an effective treatment method for children with speech production disorders?* Retrieved from https://www.uwo.ca/fhs/lwm/teaching/EBP/2011_12/Richardson.pdf
- Riches, N. G., Loucas, T., Baird, G., Charman, T., & Simonoff, E. (2012). Interpretation of compound nouns by adolescents with specific language impairment and autism spectrum

- disorders: An investigation of phenotypic overlap. *International Journal of Speech-Language Pathology*, 14(4), 307–317. <https://doi.org/10.3109/17549507.2012.679313>
- Righi, G., Tenenbaum, E. J., McCormick, C., Sheinkopf, S. J., Blossom, M., & Amso, D. (2018). Sensitivity to Audio-Visual Synchrony and Its Relation to Language Abilities in Children with and without ASD, (January), 645–653. <https://doi.org/10.1002/aur.1918>
- Roberts, J. a., Rice, M. L., & Tager-Flusberg, H. (2004). Tense marking in children with autism. *Applied Psycholinguistics*, 25, 429–448. <https://doi.org/10.1017/S0142716404001201>
- Rogers, S. J., Hayden, D., Hepburn, S., Charlifue-Smith, R., Hall, T., & Hayes, A. (2006). Teaching young nonverbal children with autism useful speech: A pilot study of the Denver model and PROMPT interventions. *Journal of Autism and Developmental Disorders*, 36, 1007–1024. <https://doi.org/10.1007/s10803-006-0142-x>
- Rojas, D.C., Bawn, S.D., Benkers, T.L., Reite, M.L., & Rogers, S.J. (2002). Smaller left hemisphere planum temporale in adults with autistic disorder. *Neuroscience Letters*, 328, 237–240.
- Romero-Munguia, M. (2008). Mnesic imbalance: a cognitive theory about autism spectrum disorders. *Annals of General Psychiatry*, 7(1), 20. <https://doi.org/10.1186/1744-859X-7-20>
- Romski, M., Sevcik, R. a, Adamson, L. B., Cheslock, M., Smith, A., Barker, R. M., & Bakeman, R. (2010). Randomized Comparison of Augmented and Nonaugmented Language Interventions for Toddlers With Developmental Delays and Their Parents. *Journal of Speech, Language, and Hearing Research*, 53(April), 350–364. [https://doi.org/10.1044/1092-4388\(2009/08-0156\)](https://doi.org/10.1044/1092-4388(2009/08-0156))
- Ronald, A., Happé, F., Bolton, P., Butcher, L. M., Price, T. S., Wheelwright, S., ... Plomin, R. (2006). Genetic heterogeneity between the three components of the autism spectrum: A twin study. *Journal of the American Academy of Child and Adolescent Psychiatry*, 45(6), 691–699. <https://doi.org/10.1097/01.chi.0000215325.13058.9d>
- Rose, V., Trembath, D., Keen, D., & Paynter, J. (2016). The proportion of minimally verbal children with autism spectrum disorder in a community-based early intervention programme. *Journal of Intellectual Disability Research*, 60, 464–477. doi: 10.1111/jir.12284
- Roseberry, S., Hirsh-Pasek, K., & Golinkoff, R. M. (2014). Skype me! Socially contingent interactions help toddlers learn language. *Child Development*, 85(3), 956–970. <https://doi.org/10.1111/cdev.12166>
- Rowe, M. L. (2012). A longitudinal investigation of the role of quantity and quality of child-directed speech in vocabulary development. *Child Development*, 83(5), 1762–1774.
- Roy, P., & Chiat, S. (2014). Developmental pathways of language and social communication problems in 9-11 year olds: Unpicking the heterogeneity. *Research in Developmental Disabilities*, 35(10), 2534–2546. <https://doi.org/10.1016/j.ridd.2014.06.014>
- Ruser, T. F., Arin, D., Dowd, M., Putnam, S., Winklosky, B., Rosen-Sheidley, B., ... Folstein, S. (2007). Communicative competence in parents of children with autism and parents of children with specific language impairment. *Journal of Autism and Developmental Disorders*, 37(7), 1323–1336. <https://doi.org/10.1007/s10803-006-0274-z>

- Rutter, M., & Bartak, L. (1971). Causes of infantile autism: Some considerations from recent research. *Journal of Autism and Childhood Schizophrenia*, 1(1), 20–32. <https://doi.org/10.1007/BF01537740>
- Rutter, M., Andersen-Wood, L., Beckett, C., Bredenkamp, D., Castle, J., Groothues, C., ... O'Connor, T. G. (1999). Quasi-autistic patterns following severe early global privation. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 40(4), 537–549. <https://doi.org/10.1017/S0021963099003935>
- Rvachew, S., & Matthews, T. (2017). Demonstrating treatment efficacy using the single subject randomization design: A tutorial and demonstration. *Journal of Communication Disorders*, 67(August 2016), 1–13. <https://doi.org/10.1016/j.jcomdis.2017.04.003>
- Sacrey, L. A. R., Bryson, S. E., & Zwaigenbaum, L. (2013). Prospective examination of visual attention during play in infants at high-risk for autism spectrum disorder: A longitudinal study from 6 to 36 months of age. *Behavioural Brain Research*, 256(July 2015), 441–450. <https://doi.org/10.1016/j.bbr.2013.08.028>
- Samson, F., Mottron, L., Jemel, B., Belin, P., & Ciocca, V. (2006). Can spectro-temporal complexity explain the autistic pattern of performance on auditory tasks? *Journal of Autism and Developmental Disorders*, 36, 65–76.
- Samson, F., Mottron, L., Soulières, I., & Zeffiro, T. A. (2012). Enhanced visual functioning in autism: An ALE meta-analysis. *Human Brain Mapping*, 33(7), 1553–1581. <https://doi.org/10.1002/hbm.21307>
- Sandiford, G.A., Mainess, K.J., & Daher, N.S. (2013). A pilot study on the efficacy of melodic based communication therapy for eliciting speech in nonverbal children with autism. *Journal of Autism and Developmental Disorders* 43, 1298–1307.
- Saffran, J. R., Senghas, A., & Trueswell, J. C. (2001). The acquisition of language by children. *Proceedings of the National Academy of Sciences of the United States of America*, 98(23), 12874–12875. <https://doi.org/10.1073/pnas.231498898>
- Saul, J., Norbury, C. (2020) Does phonetic repertoire in minimally verbal autistic preschoolers predict the severity of later expressive language impairment. *Autism* (accepted for publication).
- Scheuffgen, K., Happé, F., Anderson, M., & Frith, U. (2000). High “intelligence,” low “IQ”? Speed of processing and measured IQ in children with autism. *Development and Psychopathology*, 12(1), 83–90. <https://doi.org/10.1017/S095457940000105X>
- Schopler, E., Reichler, R. J., & Renner, B. (1988). *The Childhood Autism Rating Scale (CARS)*. Los Angeles: Western Psychological Services.
- Schreibman, L., & Stahmer, A. C. (2014). A Randomized Trial Comparison of the Effects of Verbal and Pictorial Naturalistic Communication Strategies on Spoken Language for Young Children with Autism. *Journal of Autism and Developmental Disorders*, 44(5), 1244–1251. <https://doi.org/10.1007/s10803-013-1972-y>
- Schulte, D., & Walach, H. (2006). F.M. Alexander Technique in the Treatment of Stuttering - A Randomized Single-Case Intervention Study with Ambulatory Monitoring. *Psychotherapy and Psychosomatics*, 75 (3), 190–91. <https://doi.org/10.1159/000091779>.

- Seery, A. M., Vogel-Farley, V., Tager-Flusberg, H., & Nelson, C. A. (2013). Atypical lateralization of ERP response to native and non-native speech in infants at risk for autism spectrum disorder. *Developmental Cognitive Neuroscience*, *5*, 10–24. <https://doi.org/10.1016/j.dcn.2012.11.007>
- Shadish, W. R., Hedges, L. V., & Pustejovsky, J. E. (2014). Analysis and Meta-Analysis of Single-Case Designs with a Standardized Mean Difference Statistic: A Primer and Applications. *Journal of School Psychology* *52* (2): 123–47. <https://doi.org/10.1016/j.jsp.2013.11.005>.
- Shadish, W. R., Hedges, L. V., Pustejovsky, J. E., Boyajian, J. G., Sullivan, K. J., Andrade, A., & Barrientos, J. L. (2014). A d-statistic for single-case designs that is equivalent to the usual between-groups d-statistic. *Neuropsychological Rehabilitation*, *24*(3–4), 528–553. <https://doi.org/10.1080/09602011.2013.819021>
- Shadish, W. R., Zuur, A. F., & Sullivan, K. J. (2014). Using Generalized Additive (Mixed) Models to Analyze Single Case Designs. *Journal of School Psychology* *52* (2): 149–78. <https://doi.org/10.1016/j.jsp.2013.11.004>.
- Shadish, W. R., Hedges, L. V., Horner, R. H., & Odom, S. L. (2015). *The role of between-case effect size in conducting, interpreting, and summarizing single-case research (NCER 2015-002)*. Washington, DC: National Center for Education Research, Institute of Education Sciences, U.S. Department of Education. Retrieved from <https://files.eric.ed.gov/fulltext/ED562991.pdf> on 2/12/19
- Shadish, W. R., Rindskopf, D. M., & Boyajian, J. G. (2016). Single-case experimental design yielded an effect estimate corresponding to a randomized controlled trial. *Journal of Clinical Epidemiology*, *76*, 82–88. <https://doi.org/10.1016/j.jclinepi.2016.01.035>
- Shamseer, L., Sampson, M., Bukutu, C., Shmid, C. H., Nikles, J., Tate, R., ... the CENT Group. (2015). CONSORT extension for reporting N-of-1 trials (CENT) 2015: Explanation and elaboration. *British Medical Journal*, *1793*(May), 1–30. <https://doi.org/10.1136/bmj.h1793>
- Shire, S. Y., Goods, K., Shih, W., Distefano, C., Kaiser, A., Wright, C., ... Kasari, C. (2014). Parents' Adoption of Social Communication Intervention Strategies: Families Including Children with Autism Spectrum Disorder Who are Minimally Verbal. *Journal of Autism and Developmental Disorders*, *1712–1724*. <https://doi.org/10.1007/s10803-014-2329-x>
- Shire, S. Y., Shih, W., & Kasari, C. (2018). Brief Report: Caregiver Strategy Implementation — Advancing Spoken Communication in Children Who are Minimally Verbal. *Journal of Autism and Developmental Disorders*, *48*(4), 1228–1234. <https://doi.org/10.1007/s10803-017-3454-0>
- Shriberg, L. D., Paul, R., Black, L. M., & van Santen, J. P. (2011). The hypothesis of apraxia of speech in children with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, *41*(4), 405–426. <https://doi.org/10.1007/s10803-010-1117-5>
- Shriberg, L.D., Paul, R., McSweeney, J.L., Klin, A., Cohen, D.J., Volkmar, F. R. (2001). Speech and Prosody Characteristics of Adolescents and Adults with High-Functioning Autism and Asperger Syndrome. *Journal of Speech, Language and Hearing Research*, *44*(October), 1097–1115.
- Shriberg, L. D., Tomblin, J. B., & Mcsweeney, J. L. (1999). Prevalence of speech delay in 6-year-old children and comorbidity with language impairment. *Journal of Speech, Language, and Hearing Research*, *42*, 1461–1481. <https://doi.org/10.1044/jslhr.4206.1461>

- Shumway, S., Thurm, A., Swedo, S. E., Deprey, L., Barnett, L. A., Amaral, D. G., ... Ozonoff, S. (2011). Brief report: Symptom onset patterns and functional outcomes in young children with autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *41*(12), 1727–1732. <https://doi.org/10.1007/s10803-011-1203-3>
- Sibbald, B., & Roland, M. (1998). Understanding controlled trials: Why are randomised controlled trials important? *BMJ*, *316*, 201.
- Siegelman, N., & Frost, R. (2015). Statistical learning as an individual ability: Theoretical perspectives and empirical evidence. *Journal of Memory and Language*, *81*, 105–120. <https://doi.org/10.1016/j.jml.2015.02.001>
- Sierra, V., Solanas, A., & Quera, V. (2005). Randomization Tests for Systematic Single Case Designs Are Not Always Appropriate. *Journal of Experimental Education*, *73*(2), 140–160.
- Sigman, M., & Ruskin, E. (1999). Continuity and change in the social competence of children with autism, Down syndrome, and developmental delays. *Monographs of the Society for Research in Child Development*, *64*
- Siller, M., Hutman, T., & Sigman, M. (2013). A parent-mediated intervention to increase responsive parental behaviors and child communication in children with ASD: A randomized clinical trial. *Journal of Autism and Developmental Disorders*, *43*(3), 540–555. <https://doi.org/10.1007/s10803-012-1584-y>
- Smith, P. L., & Little, D. R. (2018). Small is beautiful: In defense of the small-N design. *Psychonomic Bulletin & Review*, *25*, 2083–2101. <https://doi.org/10.3758/s13423-018-1451-8>
- Smith, T., Scahill, L., Dawson, G., Guthrie, D., Lord, C., Odom, S., ... Wagner, A. (2007). Designing Research Studies on Psychosocial Interventions in Autism. *Journal of Autism and Developmental Disorders*, *37*, 354–366. <https://doi.org/10.1007/s10803-006-0173-3>
- Smith, V., Miranda, P., & Zaidman-Zait, A. (2007). Predictors of expressive vocabulary growth in children with autism. *Journal of Speech, Language, and Hearing Research : JSLHR*, *50*(1), 149–160. [https://doi.org/10.1044/1092-4388\(2007\)013](https://doi.org/10.1044/1092-4388(2007)013)
- Solomon, B. G. (2014). Violations of Assumptions in School-Based Single-Case Data: Implications for the Selection and Interpretation of Effect Sizes. *Behavior Modification*, *38*(4), 477–496. <https://doi.org/10.1177/0145445513510931>
- St Clair, M. C., Durkin, K., Conti-Ramsden, G., & Pickles, A. (2010). Growth of reading skills in children with a history of specific language impairment: the role of autistic symptomatology and language-related abilities. *The British Journal of Developmental Psychology*, *28*(Pt 1), 109–131. <https://doi.org/10.1348/026151009X480158>
- Sterponi, L., & Shankey, J. (2014). Rethinking echolalia: repetition as interactional resource in the communication of a child with autism. *Journal of Child Language*, *41*(2), 275–304. <https://doi.org/10.1017/S0305000912000682>
- Stevenson, R. A., Segers, M., Ncube, B. L., Black, K. R., Bebko, J. M., Ferber S., & Barense, M. D. (2017). The cascading influence of multisensory processing on speech perception in autism. *Autism*, *22*(5), 609–624.

- Stockwell, K., Alabdulqader, E., Jackson, D., Basu, A., Olivier, P., & Pennington, L. (2019). Feasibility of parent communication training with remote coaching using smartphone apps. *International Journal of Language and Communication Disorders*, *54*(2), 265–280. <https://doi.org/10.1111/1460-6984.12468>
- Stone, W. L., Ousley, O. Y., & Littleford, C. D. (1997). Motor imitation in young children with autism: What's the object? *Journal of Abnormal Child Psychology*, *25*(6), 475–485. <https://doi.org/10.1023/A:1022685731726>
- Stoodley, C. J., Sillitoe, R. V., & College, B. (2014). Distinct regions of the cerebellum show gray matter decreases in autism, ADHD, and developmental dyslexia, *8*(May), 1–17. <https://doi.org/10.3389/fnsys.2014.00092>
- Sugden, E., Baker, E., Munro, N., & Williams, A. L. (2016). Involvement of parents in intervention for childhood speech sound disorders: a review of the evidence. *International Journal of Language and Communication Disorders*, *51*(6), 597–625. <https://doi.org/10.1111/1460-6984.12247>
- Sullivan, K., Sharda, M., Greenson, J., Dawson, G., & Singh, N. C. (2013). A novel method for assessing the development of speech motor function in toddlers with autism spectrum disorders. *Frontiers in Integrative Neuroscience*, *7*(March), 1–11. <https://doi.org/10.3389/fnint.2013.00017>
- Sundberg, M. L., & Partington, J. W. (1998). *Teaching language to children with autism or other developmental disabilities*. Pleasant Hill, CA: Behavior Analysts, Inc
- Swaminathan, H., Rogers, H. J., & Horner, R. H. (2014). An effect size measure and Bayesian analysis of single-case designs. *Journal of School Psychology*, *52*(2), 213–230. <https://doi.org/10.1016/j.jsp.2013.12.002>
- Swanson, M. R., Shen, M. D., Wolff, J. J., Boyd, B., Clements, M., Rehg, J., ... Gu, H. (2018). Naturalistic Language Recordings Reveal “Hypervocal” Infants at High Familial Risk for Autism. *Child Development*, *89*(2), e60–e73. <https://doi.org/10.1111/cdev.12777>
- Sweeney, M., & Lebersfeld, J. B. (2016). First Words for Nonverbal Children and Adults With Autism. *Archives of Physical Medicine and Rehabilitation*, *97*(12), e31–e32. <https://doi.org/10.1016/j.apmr.2016.09.086>
- Szatmari, P., Bryson, S. E., Boyle, M. H., Streiner, D. L., & Duku, E. (2003). Predictors of outcome among high functioning children with autism and Asperger syndrome. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *44*(4), 520–528. <https://doi.org/10.1111/1469-7610.00141>
- Tabachnick, B. G., & Fidell, L. S. (2001). *Computer-assisted research design and analysis*. Boston, MA: Allyn and Bacon.
- Tager-Flusberg, H. (2006). Defining language phenotypes in autism. *Clinical Neuroscience Research*, *6*(3–4), 219–224.
- Tager-Flusberg, H. (2010). The origins of social impairments in autism spectrum disorder: Studies of infants at risk. *Neural Networks*, *23*(8–9), 1072–1076. <https://doi.org/10.1016/j.neunet.2010.07.008>

- Tager-Flusberg, H., & Joseph, R. M. (2003). Identifying neurocognitive phenotypes in autism. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 358(1430), 303–314. <https://doi.org/10.1098/rstb.2002.1198>
- Tager-Flusberg, H., & Kasari, C. (2013). Minimally verbal school-aged children with autism spectrum disorder: The neglected end of the spectrum. *Autism Research*, 6(6), 468–478. <https://doi.org/10.1002/aur.1329>
- Tager-Flusberg, H., Plesa Skwerer, D., Joseph, R. M., Brukilacchio, B., Decker, J., Eggleston, B., ... Yoder, A. (2016). Conducting research with minimally verbal participants with autism spectrum disorder. *Autism*, 1362361316654605. <https://doi.org/10.1177/1362361316654605>
- Tager-Flusberg, H., Rogers, S., Cooper, J., Landa, R., Lord, C., Paul, R., ... & Yoder, P. (2009). Defining spoken language benchmarks and selecting measures of expressive language development for young children with autism spectrum disorders. *Journal of Speech, Language, and Hearing Research*, 52, 643–652.
- Tallal, P., & Piercy, M. (1973). Defects of non-verbal auditory perception in children with developmental aphasia. *Nature*, 241(5390), 468–469. <https://doi.org/10.1038/241468a0>
- Tate, R. L., Perdices, M., Rosenkoetter, U., Shadish, W., Vohra, S., Barlow, D. H., ... Cen-, J. W. (2016). The Single-Case Reporting Guideline In BEhavioural Interventions (SCRIBE) 2016 Statement. *Journal of Clinical Epidemiology*, 96(7), 1–11. <https://doi.org/10.1016/j.jclinepi.2016.04.006>
- Taylor, L. J., Maybery, M. T., & Whitehouse, A. J. O. (2012). Do children with specific language impairment have a cognitive profile reminiscent of autism? A review of the literature. *Journal of Autism and Developmental Disorders*, 42, 2067–2083. <https://doi.org/10.1007/s10803-012-1456-5>
- Tek, S., Jaffery, G., Fein, D., & Naigles, L. R. (2008). Do children with autism spectrum disorders show a shape bias in word learning? *Autism Research*, 1(4), 208–222. <https://doi.org/10.1002/aur.38>
- Thurm, A., Lord, C., Lee, L., & Newschaffer, C. (2007). Predictors of Language Acquisition in Preschool Children with Autism Spectrum Disorders. *J Autism Dev Disord*, 37, 13. <https://doi.org/10.1007/s10803-006-0300-1>
- Thurm, A., Manwaring, S. S., Luckenbaugh, D. A., Lord, C., & Swedo, S. E. (2014). Patterns of skill attainment and loss in young children with autism. *Development and Psychopathology*, 26, 203–214. <https://doi.org/10.1017/S0954579413000874>
- Thurm, A., Manwaring, S. S., Swineford, L., & Farmer, C. (2015). Longitudinal study of symptom severity and language in minimally verbal children with autism. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 56(1), 97–104. <https://doi.org/10.1111/jcpp.12285>
- Tierney, C., Mayes, S., Lohs, S. R., Black, A., Gisin, E., & Veglia, M. (2015). How Valid Is the Checklist for Autism Spectrum Disorder When a Child Has Apraxia of Speech? *Journal of Developmental and Behavioral Pediatrics : JDBP*, 36(8), 569–574.
- Tiger, J. H., Hanley, G. P., & Bruzek, J. (2008). Functional Communication Training: A Review and Practical Guide. *Behavior Analysis in Practice*, 1(1), 16–23. <https://doi.org/10.1007/bf03391716>

- Tomblin, B. (2011). Co-morbidity of autism and SLI: Kinds, kin and complexity. *International Journal of Language and Communication Disorders*, 46(2), 127–137. <https://doi.org/10.1111/j.1460-6984.2011.00017.x>
- Tomblin, J. B., Hafeman, L. L., & O'Brien, M. (2003). Autism and autism risk in siblings of children with specific language impairment. *International Journal of Language and Communication Disorders*, 38(3), 235–250. <https://doi.org/10.1080/1368282031000086363>
- Tomblin, J. B., Records, N. L., Buckwalter, P., Zhang, X., & O'Brien, M. (1997). Prevalence of Specific Language Impairment in. *Journal of Speech, Language, and Hearing Research*, 40(December 1997), 1245–1260.
- Toth, K., Munson, J., Meltzoff, A. N., & Dawson, G. (2006). Early predictors of communication development in young children with autism spectrum disorder: Joint attention, imitation, and toy play. *Journal of Autism and Developmental Disorders*, 36(8), 993–1005. <https://doi.org/10.1007/s10803-006-0137-7>
- Trauner, D., Wulfeck, B., Tallal, P., & Hesselink, J. (1995). Neurologic and MRI profiles of language impaired children (Technical Report CND-9513). University of California at San Diego, Center for Research in Language.
- Ullman, M. T. (2004). Contributions of memory circuits to language: The declarative/procedural model. *Cognition*, 92(1–2), 231–270. <https://doi.org/10.1016/j.cognition.2003.10.008>
- Ullman, M. T., & Pierpont, E. I. (2005). Specific language impairment is not specific to language: The procedural deficit hypothesis. *Cortex*, 41(3), 399–433. [https://doi.org/10.1016/S0010-9452\(08\)70276-4](https://doi.org/10.1016/S0010-9452(08)70276-4)
- Ullman, M. T., & Pullman, M. Y. (2015). A compensatory role for declarative memory in neurodevelopmental disorders. *Neuroscience & Biobehavioral Reviews*, 51, 205–222. <https://doi.org/10.1016/j.neubiorev.2015.01.008>
- Valentine, J. C., Tanner-Smith, E. E., Pustejovsky, J. E., & Lau, T. S. (2016). Between-case standardized mean difference effect sizes for single-case designs: a primer and tutorial using the scdhlms web application. *Campbell Systematic Reviews*, 12(1), 1–31. <https://doi.org/10.4073/cmdp.2016.1>
- Vallotton, C. D. (2009). Do infants influence their quality of care? Infants' communicative gestures predict caregivers' responsiveness. *Infant Behavior & Development*, 32(4), 351–365. <https://doi.org/10.1016/j.infbeh.2009.06.001>
- Van der Paelt, S., Warreyn, P., & Roeyers, H. (2014). Social-communicative abilities and language in preschoolers with autism spectrum disorders: Associations differ depending on language age. *Research in Autism Spectrum Disorders*, 8(5), 518–528. <https://doi.org/10.1016/j.rasd.2014.01.010>
- van Santen, J. P., Sproat, R. W., & Hill, A. P. (2013). Quantifying Repetitive Speech in Autism Spectrum Disorders and Language Impairment. *Autism Research*, 6(5), 1–19. <https://doi.org/10.1002/aur.1301>

- Van Severen, L., Van Den Berg, R., Molemans, I., & Gillis, S. (2012). Consonant inventories in the spontaneous speech of young children: A bootstrapping procedure. *Clinical Linguistics and Phonetics*, 26(2), 164–187. <https://doi.org/10.3109/02699206.2011.595527>
- van Weerdenburg, M., Verhoeven, L., & van Balkom, H. (2006). Towards a typology of specific language impairment. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 47(2), 176–189. <https://doi.org/10.1111/j.1469-7610.2005.01454.x>
- Vernay, F., Harma, K., Marrone, T., & Roussey, J. (2017). Self-paced segmentation of written words on a touchscreen tablet promotes the oral production of nonverbal and minimally verbal children with autism. *Journal of Research in Special Educational Needs*, 17(4), 265–273. <https://doi.org/10.1111/1471-3802.12384>
- Vernes, S. C., Newbury, D. F., Abrahams, B. S., Winchester, L., Nicod, J., Groszer, M., ... Fisher, S. E. (2008). A functional genetic link between distinct developmental language disorders. *The New England Journal of Medicine*, 359(22), 2337–2345. <https://doi.org/10.1056/NEJMoa0802828>
- Vohra, S., Shamseer, L., Sampson, M., Bukutu, C., Schmid, C. H., Tate, R., ... Shadish, W. R. (2016). CONSORT extension for reporting N-of-1 trials (CENT) 2015 Statement. *Journal of Clinical Epidemiology*, 76, 9–17. <https://doi.org/10.1016/j.jclinepi.2015.05.004>
- Vohra, S. (2016). N-of-1 trials to enhance patient outcomes: identifying effective therapies and reducing harms, one patient at a time. *Journal of Clinical Epidemiology*, 76, 6–8. <https://doi.org/10.1016/j.jclinepi.2016.03.028>
- Von Hippel, P. (2007). Regression with missing Ys: An improved strategy for analyzing multiply imputed data. *Sociological Methodology*, 37, 83–118.
- Von Hippel, P. (2009). How to impute interactions, squares, and other transformed variables. *Sociological Methodology*, 39, 265–291.
- Vugs, B., Hendriks, M., Cuperus, J., & Verhoeven, L. (2014). Working memory performance and executive function behaviors in young children with SLI. *Research in Developmental Disabilities*, 35(1), 62–74. <https://doi.org/10.1016/j.ridd.2013.10.022>
- Wallace, G. L., Anderson, M., & Happé, F. (2009). Brief report: Information processing speed is intact in autism but not correlated with measured intelligence. *Journal of Autism and Developmental Disorders*, 39(5), 809–814. <https://doi.org/10.1007/s10803-008-0684-1>
- Wan, C. Y., Bazen, L., Baars, R., Libenson, A., Zipse, L., Zuk, J., ... Schlaug, G. (2011). Auditory-motor mapping training as an intervention to facilitate speech output in non-verbal children with autism: A proof of concept study. *PLoS ONE*, 6(9), 1–7. <https://doi.org/10.1371/journal.pone.0025505>
- Wan, C. Y., Marchina, S., Norton, A., & Schlaug, G. (2012). Atypical hemispheric asymmetry in the arcuate fasciculus of completely nonverbal children with autism. *Annals of the New York Academy of Sciences*, 1252(1), 332–337. <https://doi.org/10.1111/j.1749-6632.2012.06446.x>
- Wang, H. T., & Koyama, T. (2014). An analysis and review of the literatures and a three-tier video modeling intervention model. *Research in Autism Spectrum Disorders*, 8(7), 746–758.

- Ward, R., Leitão, S., & Strauss, G. (2014). An evaluation of the effectiveness of PROMPT therapy in improving speech production accuracy in six children with cerebral palsy. *International Journal of Speech-Language Pathology*, *16*(4), 355–71. <https://doi.org/10.3109/17549507.2013.876662>
- Warlaumont, A. S., & Oller, D. K. (2014). A Social Feedback Loop for Speech development and its reduction in autism. *Psychological Science*, *25*(7), 1314–1324.
- Watson, L. R., Baranek, G. T., Roberts, J. E., David, F. J., & Twyla, Y. (2011). Behavioral and Physiological Responses to Child-Directed Speech as Predictors of Communication Outcomes in Children with Autism Spectrum Disorders. *Journal of Speech*, *53*(4), 1052–1064. [https://doi.org/10.1044/1092-4388\(2009/09-0096\)](https://doi.org/10.1044/1092-4388(2009/09-0096)).
- Weikum, W., Vouloumanos, A., Navarra, J., Soto-Faraco, S., Sebastian-Galles, N. & Werker, J.F. (2007). Visual language discrimination in infancy. *Science* *316*(1159). <https://doi:10.1126/science.1137686>.
- Wenman, R., Bowen, A., Tallis, R. C., Gardener, E., Cross, S., & Niven, D. (2003). Use of a randomised single case experimental design to evaluate therapy for unilateral neglect. *Neuropsychological Rehabilitation*, *13*(4), 441–459. <https://doi.org/10.1080/09602010343000048>
- West, G., Vadillo, M. A., Shanks, D. R., & Hulme, C. (2017). The procedural learning deficit hypothesis of language learning disorders: we see some problems. *Developmental Science*, *21*(2), 1–13. <https://doi.org/10.1111/desc.12552>
- West, G., Clayton, F. J., Shanks, D. R., & Hulme, C. (2019). Procedural and declarative learning in dyslexia. *Dyslexia*. <https://doi.org/10.1002/dys.1615>
- Wetherby, A., & Prizant, B. (2002). Communication and symbolic behavior scales developmental profile—First Normed edition. Baltimore, MD: Paul H. Brookes.
- Whitehouse, A. J. O., Bishop, D. V. M., Ang, Q. W., Pennell, C. E., & Fisher, S. E. (2011). CNTNAP2 variants affect early language development in the general population. *Genes, Brain and Behavior*, *10*(4), 451–456. <https://doi.org/10.1111/j.1601-183X.2011.00684.x>
- Whitehouse, A. J. O., Barry, J. G., & Bishop, D. V. M. (2008). Further defining the language impairment of autism: Is there a specific language impairment subtype? *Journal of Communication Disorders*, *41*, 319–336. <https://doi.org/10.1016/j.jcomdis.2008.01.002>
- Whitehouse, A. J. O., & Bishop, D. V. M. (2008a). Do children with autism “switch off” to speech sounds? An investigation using event-related potentials. *Developmental Science*, *11*(4), 516–524. <https://doi.org/10.1111/j.1467-7687.2008.00697>.
- Whitehouse, A. J. O., & Bishop, D. V. M. (2008b). Cerebral dominance for language function in adults with specific language impairment or autism. *Brain*, *131*(12), 3193–3200. <https://doi.org/10.1093/brain/awn266>
- Wilkinson, L., & the APA Task Force on Statistical Inference. (1999). Statistical methods in psychology journals: Guidelines and explanations. *American Psychologist*, *54*, 594–604.
- Williams, D. [Donna] (1999). *Nobody Nowhere*. London: Jessica Kingsley Publishers.
- Williams, D. [David], Botting, N., & Boucher, J. (2008). Language in autism and specific language impairment: where are the links? *Psychological Bulletin*, *134*(6), 944–963.

- Williams, M. E., Hastings, R. P., & Hutchings, J. (2020). The Incredible Years Autism Spectrum and Language Delays Parent Program: A Pragmatic, Feasibility Randomized Controlled Trial. *Autism Research*, 1–12.
- Wilson, K. P. (2011). Synthesis of single-case design research in communication sciences and disorders: Challenges, strategies, and future directions. *Evidence-Based Communication Assessment and Intervention*, 5(2), 104–115. <https://doi.org/10.1080/17489539.2011.623859>
- Wilson, A. C., & Bishop, D. V. M. (2018). Resounding failure to replicate links between developmental language disorder and cerebral lateralisation. <https://doi.org/10.7717/peerj.4217>
- Wodka, E. L., Mathy, P., & Kalb, L. (2013). Predictors of Phrase and Fluent Speech in Children With Autism and Severe Language Delay. *Pediatrics*, 131(4), e1128–e1134. <https://doi.org/10.1542/peds.2012-2221>
- Wolf, M., Risley, T., & Mees, H. (1964). Application of operant conditioning procedures to the behaviour problems of an autistic child. *Behaviour Research and Therapy*, 1, 305–312
- Wong, C., Odom, S. L., Hume, K., Cox, A. W., Fettig, A., Kucharszyk, S., . . . Schultz, T. R. (2014). *Evidence-based practices for children, youth, and young adults with Autism Spectrum Disorder*. Chapel Hill: The University of North Carolina, Frank Porter Graham Child Development Institute, Autism Evidence- Based Practice Review Group.
- Woynaroski, T., Watson, L., Gardner, E., Newsom, C. R., Keceli-Kaysili, B., & Yoder, P. J. (2016). Early Predictors of Growth in Diversity of Key Consonants Used in Communication in Initially Preverbal Children with Autism Spectrum Disorder. *Journal of Autism and Developmental Disorders*, 46(3), 1013–1024. <https://doi.org/10.1007/s10803-015-2647-7>
- Woynaroski, T., Oller, D. K., Keceli-Kaysili, B., Xu, D., Richards, J. A., Gilkerson, J., . . . Yoder, P. (2017). The stability and validity of automated vocal analysis in preverbal preschoolers with autism spectrum disorder. *Autism Research*, 10(3), 508–519. <https://doi.org/10.1002/aur.1667>
- Yang, Z., Sun, X., & Hardin, J. W. (2010). A Note on the Tests for Clustered Matched-Pair Binary Data. *Biometrical Journal*, 52 (5), 638–52. <https://doi.org/10.1002/bimj.201000035>.
- Yoder, P., & Stone, W. L. (2006). A Randomized Comparison of the Effect of Two Prelinguistic Communication Interventions on the Acquisition Preschoolers With ASD. *Journal of Speech, Language and Hearing Research*, 49, 698–711.
- Yoder, P., & Warren, S. F. (1999). Maternal Responsivity Mediates the Relationship Between Prelinguistic Intentional Communication and Later Language. *Journal of Early Intervention*, 22(2), 126–136. <https://doi.org/10.1177/105381519902200205>
- Yoder, P., Watson, L. R., & Lambert, W. (2015). Value-Added Predictors of Expressive and Receptive Language Growth in Initially Nonverbal Preschoolers with Autism Spectrum Disorders. *Journal of Autism and Developmental Disorders*, 45(5), 1254–1270.
- Young, G. S., Merin, N., Rogers, S. J., & Ozonoff, S. (2009). Gaze behavior and affect at 6 months: Predicting clinical outcomes and language development in typically developing infants and infants at risk for autism. *Developmental Science*, 12(5), 798–814. <https://doi.org/10.1111/j.1467-7687.2009.00833.x>

- Zhang, X., & Bruce Tomblin, J. (2000). The Association of Intervention Receipt with Speech-Language Profiles and Social-Demographic Variables. *American Journal of Speech-Language Pathology*, 9(4), 345–357. <https://doi.org/10.1044/1058-0360.0904.345>
- Zwaigenbaum, L., Bryson, S., & Garon, N. (2013). Early identification of autism spectrum disorders. *Behavioural Brain Research*, 251, 133–146. <https://doi.org/10.1016/j.bbr.2013.04.004>

8. Appendices

Appendix A: CSBS Coding information

Adapted from Wetherby, A., & Prizant, B. (2002). *Communication and symbolic behavior scales developmental profile—First Normed edition*. Baltimore, MD: Paul H. Brookes.

Step 1: Communicative Acts (CAs) must first be identified and the relevant section of video should be delineated using the ELAN programme.

To qualify as a Communicative Act, refer to Figure 4.1 (Wetherby & Prizant, p.34) summarised below.

Figure 4.1

For a behaviour to be considered a communicative act, the answer to all three of the following question must be yes; that is, the behaviour must be described by at least one of the options under each question.

1. Was the act a gesture, vocalisation or verbalization?
2. Was the act directed toward the adult?
3. Did the act serve a communicative function (see table 4.1 below)

Termination of a CA happens when:

- Exchange of turns (vocal/gestural)
- Pause greater than 3 seconds
- Child shifts topic/focus of attention

Incomplete CAs should not be counted. This is when child abandons the act or is interrupted before it is completed.

Step 2: identify the communicative function of the behaviour

Table 4.1 (Wetherby & Prizant, p.32)

<u>Behaviour regulation</u>	Acts used to regulate the behaviour of another person to obtain a specific result. Child's goal is to get the adult to do something or stop doing something.
Examples	Request object/action - acts used to direct another to give a desired object or to carry out an action
	Protest object/action – acts used to refuse an object that is not desired or to direct another to cease an action that is not desired
<u>Social Interaction</u>	Acts used to attract or maintain another's attention to oneself. Child's goal is to get adult to look at or notice him or her
Examples	Request social routine – act used to direct another to begin or continue carrying out a game-like social interaction
	Request comfort – acts used to seek another's attention to comfort from wariness, distress or frustration.
	Call – acts used to gain the attention of another to indicate that a communicative act is to follow
	Greet – acts used to indicate notice of a person or object's presence or to signal the initiation or termination of an interaction
	Show off – acts used to attract another's attention to oneself by displaying a performance

	Request permission – acts used to seek another’s consent to carry out an action; involves carrying out or wanting to carry out the action
<u>Joint attention</u>	Acts used to direct another’s attention to an object, event or topic of a communicative act. Child’s goal is to get adult to look at or notice an entity or event.
Examples	Comment on object/action – acts used to direct another’s attention to an entity or event
	Request information – acts used to seek information, explanations, or clarifications, about an entity, event, or previous utterance; includes wh-questions and other utterances with rising intonation contour
Unclear	Acts used for a communicative purpose but for which there is insufficient information to determine the category of function that it most appropriately fits. That is, it cannot be determined whether the child’s goal is to communicate for behaviour regulation, social interaction or joint attention.

Step 3: Identify Respondent Acts

- Yes: the CA is in response to an adult’s conventional gestures or speech. Topic of CA must be same as topic of adult act. Must occur 2-3 seconds after adult act to be considered. Does not have to display comprehension of the adult act.

- No: if not respondent, the act is 'initiated'

→ Use step 2 and 3 to insert IBR, RBR, IJA, RJA, ISI, RSI or unclear into the Communicative Function column in ELAN

Step 4: Identify the communicative means (Gesture alone; vocal act alone; gesture and vocal act together)

- **A gesture** is any non-vocal behaviour directed to another person that serves a communicative function. Directed eye-gaze without head or hand movement does not count as a gesture. Holding, touching or taking an object does not count as a gesture unless the child is directing the gesture towards the adult (e.g. showing, giving). Examples given:
 - Giving object to adult*
 - Touching adult's hand, arm, body or face
 - Moving adult's hand or face
 - Pushing object toward or away from adult*
 - Head shaking or nodding*
 - Hitting, biting or pinching self or adult
 - Throwing or dropping object
 - Showing off without object
 - Showing off with object near child's face
 - Making indicative gesture (pointing* or tapping with finger or fingers (not palm)); raising arms; open-hand reaching* with minimal body movement; showing object*
 - Making depictive gesture (i.e. pantomime-like action)
 - Waving*
 - Clapping

NB. Gestures marked * are considered CONVENTIONAL GESTURES (not required for coding scheme)

NB. Distal hand gestures are made when a child's hand does not touch a person or object when gesturing. Examples include reach, point, waving, some depictive gestures

- **Vocal act** is further broken down into:
 - o **Vocalisation** may be spontaneous or imitative vocal acts that do not contain recognisable words
 - Nontranscribable (e.g. laugh, cries, sighs, lip smacks, trills, single consonant without a vowel). Any vocalisation is non-transcribable if you cannot transcribe it within 3 attempts.
 - or transcribable (requires at least one vowel sound and may include a consonant, may be single syllable or multisyllabic)
 - o **Verbalisation** (spoken single or multiword utterance) NB if there is a non-spoken verbalisation (signed or AAC words), mark this as gestural and describe under the 'other' column

Consonants used for consonant inventory

(from table 4.2, Wetherby & Prizant, p.36)

Sound	Example
/m/	Mother
/n/	No
/ŋ/	Ring
/b/	Ball
/p/	Pop
/d/	Dog
/t/	Toy
/g/	Goat
/k/	Cookie
/w/	Wagon
/j/	Yellow
/l/	Little

/r/	<i>Read</i>
/v/	<i>Very</i>
/f/	<i>food</i>
/z/	<i>zoo</i>
/s/	<i>Soap</i>
/ð/	<i>The</i>
/θ/	<i>Thumb</i>
/ʃ/	<i>Shoe</i>
/tʃ/	<i>Church</i>
/ʒ/	<i>Measure</i>
/dʒ/	<i>Judge</i>

Observed Phoneme Inventory

One point per line

VOWELS

Sound	example
/ee/	<i>Keep</i>
/ai/	<i>Kind</i>
/oo/	<i>Soon</i>
/e/	<i>Get</i>
/u/	<i>Up</i>
/a/	<i>Cat</i>
/or/	<i>Door</i>
/i/	<i>Tip</i>
/oh/	<i>No</i>
/o/	<i>top</i>
/ah/	<i>Far</i>
/ay/ or /ou/ or /er/ or /oi/*	<i>Say, out, player, boy</i>

*using any of these sounds scores 1 point

CONSONANTS

Note these are slightly different to the consonant inventory above

Sound	Example
/m/	<i>Mother</i>
/n/	<i>No</i>
/ŋ/	<i>Ring</i>
/b/	<i>Ball</i>
/p/	<i>Pop</i>
/d/	<i>Dog</i>
/t/	<i>Toy</i>
/g/	<i>Goat</i>
/k/	<i>Cookie</i>
/w/	<i>Wagon</i>
/j/	<i>Yellow</i>
/l/	<i>Little</i>
/r/	<i>Read</i>
/v/ or /f/	<i>Very, food</i>
/h/	<i>hand</i>
/z/ or /s/	<i>Zoo, Soap</i>
/ʃ/	<i>Shoe</i>
/tʃ/ or /dʒ/	<i>Church, Judge</i>

Appendix B: Phoneme Questionnaire

A phoneme is a speech sound. A word like 'bat' contains three phonemes: 'b' 'a' and 't'. Tick the following sounds that your child regularly uses for communication.

m as in mouse	a as in cat
n as in no	e as in get
b/p as in ball/pot	i as in fit
d/t as in day/top	o as in pot
g/k as in good/can	u as in cup
y as in yes	ee as in see
w as in wait	oo as in do
l as in look	ie as in my
s as in sorry	ai as in say
sh as in ship	ou as in out

Appendix C: Alphabet Stimuli

Alphabet and phonics test stimuli

Part 1: Stimuli are presented for Q1 and then each question thereafter, until each question has been asked or the stop rules have ended the test. The experimenter asks the child to give them the underlined symbol using the letter name (e.g. 'give me B'). Stop rules: after three consecutive incorrect answers (to include not answering), the test ends.

Part 2: For each question in part 1 that was not correctly answered, present the stimuli again and ask for the phonics name (i.e. give me 'b'). Same stop rules as above apply.

Correct answer is underlined, other two stimuli are distractors

Question	Left	Centre	Right
1	a	s	<u>l</u>
2	<u>a</u>	q	t
4	o	<u>b</u>	z
3	e	p	<u>r</u>
5	r	p	<u>k</u>
6	<u>s</u>	e	t
8	a	<u>w</u>	t
9	<u>c</u>	n	i
10	<u>f</u>	o	m
12	x	<u>n</u>	a
14	i	p	<u>t</u>
15	p	j	e

16	<u>e</u>	t	j
18	<u>i</u>	n	g
19	h	f	<u>o</u>
24	h	<u>u</u>	c

NB: Abridged from original test of the full alphabet

Appendix D: Family Background Questionnaire

1. Parent's age: Mother _____ Father _____
2. Which adults does the child participating in this study live with?
 - Mother
 - Father
 - Other _____ caregiver(s) (please specify)
3. How many siblings does the child participating in this study have? _____
4. Please indicate gender, age, any developmental or major health concerns of any siblings below (please use space overleaf if there are more than 3 siblings)

	Gender (M/F)	Age	Developmental / Health concerns?	Does this sibling live with the child participant?
Sibling 1				
Sibling 2				
Sibling 3				

5. Is there any family history of autism, developmental delay, language disorder, hyperactivity or mental health problems? Please specify below.

6. Parent 1 (mother/father – delete as appropriate)
 - a. What is your highest qualification level?
 - High school
 - Undergraduate degree
 - Masters degree
 - PhD
 - Other (please specify below)
 - Prefer not to answer
 - b. What is your job title (e.g., primary school teacher, structural engineer)

-
- c. Are you working outside the home right now?
- Yes, Full time
- Yes, Part time
- Not working

7. Parent 2 (mother/father/other caregiver – delete as appropriate)

- a. What is your highest qualification level?

- High school
- Undergraduate degree
- Masters degree
- PhD
- Other (please specify below)
- Prefer not to answer

- b. What is your job title (e.g., primary school teacher, structural engineer)

-
- c. Are you working outside the home right now?
- Yes, Full time
- Yes, Part time
- Not working

8. Please confirm that English is the main language spoken at home

9. Please indicate any other languages your child is/has been exposed to:

Language (e.g. French, Urdu)	Source of exposure (e.g. mother, grandparent)	Duration and amount of exposure

10. This form was completed by:

- Mother
- Father
- Other caregiver specified above

Date _____

Appendix E: Therapy Report Questionnaire

Please describe all therapies currently experienced by your child or experienced by your child **in the last 4 months**. If you would like to comment separately on any previous therapies or add any further details, please do so in the 'any other comments' section.

My child has not undertaken any therapy in the last 4 months

My child has undertaken the following therapies in the last 4 months:

Therapy 1

Name of therapy: (e.g. Occupational Therapy, Portage, Speech and Language Therapy, Behavioral therapy, Music therapy, Physiotherapy, Social Skills groups, other therapies)

This therapy was organized:

privately by me

by my local authority/primary care trust

This therapy takes place at:

Home

Education setting (e.g. preschool)

Clinic

My child receives this therapy:

In a group

On one-to-one basis

On average my child takes part in _____ hours per week of this therapy.

Therapy 2

Name of therapy: _____

This therapy was organized:

privately by me

by my local authority/primary care trust

This therapy takes place at:

Home

Education setting (e.g. preschool)

Clinic

My child receives this therapy:

In a group

On one-to-one basis

On average my child takes part in _____ hours per week of this therapy.

Therapy 3

Name of therapy: _____

This therapy was organized:

privately by me

by my local authority/primary care trust

This therapy takes place at:

Home

Education setting (e.g. preschool)

Clinic

My child receives this therapy:

In a group

On one-to-one basis

On average my child takes part in _____ hours per week of this therapy.

Therapy 4

Name of therapy: _____

This therapy was organized:

privately by me

by my local authority/primary care trust

This therapy takes place at:

Home

Education setting (e.g. preschool)

Clinic

My child receives this therapy:

In a group

On one-to-one basis

On average my child takes part in _____ hours per week of this therapy.

Therapy 5

Name of therapy: _____

This therapy was organized:

privately by me

by my local authority/primary care trust

This therapy takes place at:

Home

Education setting (e.g. preschool)

Clinic

My child receives this therapy:

In a group

On one-to-one basis

On average my child takes part in _____ hours per week of this therapy.

Any other comments regarding therapies undertaken by your child?

Appendix F: Parent-Child Interaction Coding information

Adapted from “Partial Interval Time Sampling of Adaptive Strategies for the Useful Speech Project v 3/1/10” (Yoder, Fey, Thompson, McDuffie, Lieberman, Flippin, Watson, Firestine).

- Each video is 15 minutes long and divided into 180 intervals of 5-seconds
- The following toys are presented to each dyad:
 - brightly coloured 'Mr Tumble' hat
 - box of two-piece tactile animal puzzles
 - brightly coloured wig
 - glove with flashing lights at each finger tip
 - comb
 - plastic figurine of 'In The Night Garden' character Upsy Daisy
 - plastic figurine of 'Toy Story Character Buzz Lightyear
 - 2 x puppets (polar bear and multi-coloured monster)
 - 5 x brightly coloured partially transparent bricks
 - rainmaker
 - spikey ball
 - duplo bus
 - In The Night Garden spinning shapersorter
 - plastic wind up musical Jack in box

First Pass

- The goal of the first pass is to mark each interval as “N” or “C” – only those marked “C” (codeable) will be reviewed for the second pass
- At this point the only goal is to establish if N or C (no blanks allowed)
- Reasons to code an interval as N:
 - Distractions occur during the interval
 - Child is behaving in a way that requires the adult to intervene (e.g. trying to leave the room, hitting, biting, self-harm, crying uncontrollably)

- Part of session is interrupted (parent leaves to answer phone, child needs nappy change etc)
- Due to arrangement of referents/people/camera there is no clear view of either leads or responses that have occurred
 - Child or adult are off screen or video is out of focus
 - Can't see adult or child hands to determine whether they are doing a touch lead or physical play
 - Can't see child's face or head orientation to determine attentional lead
- Even if only part of interval is affected by above – code as N

Second Pass

- Only applies to intervals coded “C”
- The first goal is to identify whether there is a child lead in each interval and if so, whether there was a parent response – check both of these at the same time even though they are described sequentially below. Additionally, we code the referent to make it easier to identify adopted leads.

Child Lead

- **If no child lead is present**, leave the 'lead' column blank and no further coding is needed for that interval regardless of what the parent is doing
- A child lead is present when:
 - the child is **demonstrating attention** toward a **referent**
 - AND they '**own**' the lead:
- Attention is demonstrated by
 - looking toward it for at least 1 second
 - OR touching the referent (moving it with hand or moving hand on it)
 - For our coding purposes we are not making the distinction, however, you need to be aware of this because whether a lead is look or touch is sometimes important

- A referent can be toys, people, things in the environment, and also (less frequently) a communicative word, gesture or sign used by the child
- 'Owning the lead' occurs:
 - Immediately if the child initiated the lead (code as CI) e.g. child starts looking at a book without any verbal instruction (code CI, referent = book) adult NONVERBAL actions are fine
 - With a slight delay if the child adopts an adult lead (code as CA) e.g. adult introduces a book and the child attends to it for a further 3 intervals – only the 3rd interval is coded as CA
- A child adopted lead is one where the parent has introduced a referent but the child has sustained a focus on it for 2 consecutive intervals not including the one in which it was introduced

Codeable	Child Lead	Response	Child referent	Parent referent	Description
C			bus	ball	Referent introduced by parent
C			ball	ball	Child attends to referent [1]
C			ball	ball	Child attends to referent [2]
C	CA	U	ball	ball	This interval is classified as adopted lead for the referent

- What to do if the child displays more than one lead per interval?
 1. Check if any of the leads cause an adult response
 2. If so, code the first child lead to which an adult responded
 3. If not, just code the first one
- What if child looks at one referent whilst touching another?

1. Looking is credited before touching
 2. UNLESS there are multiple leads affected by the rule above (i.e. the first adult response corresponds with the touched item)
- If a child looks at and touches a referent, it is counted as a touch lead (no BOTH option)
 - The same lead can carry over into multiple intervals and should be credited each time if a child sustains attention

Parent responses

- All intervals with a child lead are considered for response
- Response must correspond with the lead in that interval (not prior intervals)
- Parent responses can be coded as physical (“P”), utterance (“U”) or “BOTH”
 - Physical response and BOTH can ONLY follow a touch lead (not attentional)
 - Utterance response can follow any lead
 - Only code an utterance response if you can establish what the parent said after 3 listens (otherwise -> no response or P depending on if it was U or BOTH)
- If no response displayed – leave cell blank
- Responses only get coded if they refer to the child referent
- Physical response examples
 - Imitate child action with same or similar referent
 - Aid the child’s action
 - Demonstrate new action on child’s referent
 - Demonstrate new action on different referent and relate this to child’s referent
- Follow-in utterance examples
 - Must be about the child’s referent (or a part of it, or an item related to it) **AND**
 - Must have specific semantic relationship to the child referent

- Follow-in utterances that are allowed (so long as based on child lead referent)
 - The ball rolled away
 - There's the ball. You like the big blue ball
 - Child plays with puppet "Oo so pretty"
 - Child plays with cow "moo"
 - "put the block in"
 - "turn the page"
 - "do you want the red one or the green one?"
- Follow-in utterances that are not allowed
 - non-specific affirmatives (ok, right, sure, yes, yay, good job, that's right, way to go, you got it or any UK equivalents e.g. well done, good boy)
 - Recitative utterances (singing songs, reading out text, counting in sequence)
 - No
 - Interjections (huh? Eh? Hey? Oh?) etc
 - Statements intended to keep the child from doing something in future (e.g. stop throwing the toy) or about past memories or general statements (we never play like this do we, do you remember what you had for breakfast)

Referent coding

- It is useful to code the child and adult referent in order to track how long since a referent has been introduced by the adult for purpose of CA coding
- Do not need there to be a valid lead or response in order to code referents
 - E.g. child is looking at ball but this was introduced in previous interval by adult (= not a CI or a CA, referent= ball)
 - e.g. child lead relates to shapersorter, adult is talking about the book (code CI, no response, child referent= shapersorter, adult referent= book)

Appendix G: Stage 1 Consultation

1. Aim

The aim of this consultation was to seek initial qualitative feedback from stakeholders regarding the appropriateness of the proposed BabbleBooster app.

2. Methods

Participants

Participants were recruited via a flyer posted on social media. It specifically asked for respondents who could answer any of the three questions positively:

- Did/does your child with autism struggle to learn to produce spoken words and sounds?
- Did/does your child with autism play with smartphone or tablet apps?
- Are you a therapist who works with children with autism who struggle to acquire speech?

Six parents made initial contact with the first author to attend the focus group but for logistical reasons only four attended the group, henceforth referred to as Participants L, E, R and A.

Participant L was the mother of a 7 year-old autistic boy currently educated in a specialist autism unit attached to a mainstream school. She describes him as verbal now, but having had delayed speech and communication development, with his articulation still a bit impacted. Participant E was the mother of an 8 year-old autistic boy currently educated in a specialist school, whom she describes as verbal but cognitively impaired. She describes him as having ongoing articulatory problems. Participant A was the mother of a 10 year-old autistic girl, currently educated in mainstream school with support. She describes her as a very late talker with extreme verbal dyspraxia, which has now resolved. Her current profile is of typical cognitive and verbal ability. Participant R is the mother of an 8 year-old autistic boy currently educated in a language unit. She describes a history of regression and language delay. He is now verbal with remaining language difficulties.

Procedure

The focus group was held in March 2017 and was chaired by the first author. Two members of the UCL Computer Science student team who were creating the app were also present.

The focus group took the form of a brief presentation of the proposed app's features, aims and mechanisms, followed by a series of open ended questions designed to prompt group discussion. Questions are summarised in Table A. The meeting was audio recorded and transcribed verbatim by the first author. The first

author analyzed the data using thematic analysis (Braun & Clarke, 2006) where the content of the focus group transcript was analyzed and themes identified.

3. Results

Output of the focus group pertaining to the development of the app is summarized below. Due to the limited time (90 minutes), not every question in Table 5 was asked, the discussion flowed in a conversational style.

Technology: all parents reported that mobile and tablet devices were inherently motivating for their children, with the most commonly used function being to access to video content online (e.g. via YouTube). Content was often esoteric, user-uploaded and specific to the child's special interests (e.g. people going on waterslides, opening toys).

Aim: all liked the idea of the app and the mirror function. Parents suggested having images that match the sounds would make it more functional. Parents would like to have input on the initial sound selection process.

Time commitment: all agreed five minutes per day is an achievable target.

Cued articulation aspect: only Participant A had heard of this approach, but when her daughter was minimally verbal she had found it very helpful in progressing speech skills.

Video modelling aspect: all agreed this would be a useful feature. Participant R remembered it being hard to get her son to look at her whilst she modelled language, and that is why she thinks he found PECs (a picture exchange communication system) easier than Makaton (a simplified form of sign language, requiring learners to copy manual signs).

Parent feedback on child productions: parents unanimously disliked the proposed red "no" button, reporting that their children were very sensitive to 'getting things wrong'. They suggested changing it to a 'try again' button and altering the colours.

Reinforcing videos: all agreed that customisable content is a must-have feature of the app. We discussed various ways of supporting parents to create content, for example providing a parent idea sheet or 'how to' videos.

Table A: Focus group discussion guide

Theme	Questions
Technology for young autistic children	<p><i>What kind of devices do you/have you in the past let your child play with?</i></p> <p><i>Have you encouraged app play? Or placed restrictions on it?</i></p> <p><i>When your child was younger did you let them play alone or support their app play?</i></p> <p><i>What do you see as the main benefit of the apps your child uses/used?</i></p> <p><i>Please share any positive and negative experiences using apps?</i></p> <p><i>Do you view the purpose of apps as primarily educational or recreational?</i></p> <p><i>Do apps have a limited 'shelf-life' with your child if they use them intensively?</i></p> <p><i>Which apps have had the most impact on your child?</i></p> <p><i>What makes a good app for this population?</i></p> <p><i>Which app features are particularly effective?</i></p> <p><i>Which app features are particularly annoying?</i></p> <p><i>What do you think app-designers need to consider for this population?</i></p> <p><i>Any other uses of technology worth mentioning (e.g. Siri, youtube etc)?</i></p> <p><i>Have you used any apps that are aimed at improving speech?</i></p> <p><i>If so at what point in speech development was your child?</i></p> <p><i>Do you think an app could support speech development?</i></p> <p><i>What features do you think would be necessary/desirable for such an app?</i></p>
Feedback on the proposed BabbleBooster app	<p><i>Do you think the proposed structure of 5-10 minutes per day, 5 days a week for 12 weeks is acceptable and achievable time commitment?</i></p> <p><i>Do you think parents will carry on the visual prompts in day to day?</i></p> <p><i>Do you think having the app is likely to create regular practicing routines for families?</i></p> <p><i>Would built-in text reminders be a good idea?</i></p> <p><i>Parent needs to play app alongside child at all times but should there be a 'viewing only mode' for a child who wants to watch video modelling on a loop by themselves?</i></p> <p><i>Do you think parents will find it easy enough to rate their child's speech productions?</i></p> <p><i>Is the playback option useful?</i></p> <p><i>How important is it for parents to also receive feedback on how the therapy is going?</i></p> <p><i>The idea is to take footage of child pre-intervention and determine which sounds are already in the repertoire, and which are not. Any thoughts on this?</i></p> <p><i>Should parents have any input into the sound selection process?</i></p>

	<p><i>In your experience were consonants and vowels equally important?</i></p> <p><i>How many sounds do you think should be targeted at once?</i></p> <p><i>After how long working on a sound in vain should an alternative sound be tried instead?</i></p> <p><i>Any comments on how sounds should be modelled?</i></p> <p><i>Is cued articulation aspect useful / neutral / distracting for your child?</i></p> <p><i>Should there be any choice in the video modelling or option to upload yourself doing it?</i></p> <p><i>How could videos be visually improved to make articulation part more salient (e.g. artificially slowed, zooming in, several faces at once, other video effects?)</i></p> <p><i>Design will be customisable so that parents can upload reinforcing videos and images - would you use this option? Will this add value for the child?</i></p> <p><i>Should there be some 'placeholder' videos for those who prefer not to record themselves?</i></p> <p><i>Will the child get bored of the reinforcement videos and need them to be changed frequently?</i></p> <p><i>Does the reinforcement aspect need to be more dynamic? E.g. you pop bubbles and photos of favourite things are revealed?</i></p> <p><i>Would child benefit from watching back their 'best bits' at the end of the session? E.g. a clip of a successful production and audio saying – "Yay! [name of child] said 'b'"</i></p> <p><i>Any other issues you see with the app or possible things to improve?</i></p> <p><i>Do you think help files might be useful in case of technical problems with the app?</i></p>
<p>Speech intervention experiences</p>	<p><i>How did you try to improve your child's speech production? Did you feel you understood the developmental sequence of speech production, what skills you should be targeting, and how to do it?</i></p> <p><i>What specific therapies did you try for this?</i></p> <p><i>Which professionals did you ask for advice on this?</i></p> <p><i>Did you agree with focus of therapy on offer?</i></p> <p><i>What seemed to be the main barriers to progress with speech?</i></p> <p><i>Do you feel it was something that could be helped by repetition or practice or was there a more fundamental difficulty that needed to be addressed first (e.g. attention, low muscle tone, challenging behaviours...)?</i></p> <p><i>Was the therapy that you tried able to be adapted to your child's specific needs/interests? In particular the materials (songs, visuals, videos) or the method of delivery</i></p>

4. Discussion

Given the positive feedback regarding aims and features of the app, very few changes were proposed to BabbleBooster. Two pieces of feedback were implemented by the app team. Firstly, rather than just providing the sound and a letter symbol in the sound modelling phase, this was changed to three images corresponding to the sound (e.g. for 'b': 'baby', 'ball', 'biscuit'). These images were able to be replaced or exchanged by parent customization. Secondly, rather than presenting parents with three choices for feedback buttons ('yes', 'good try' or 'no'), this was changed to 'yes', 'good try' or 'try again' and red and green colours were removed, in order to appease parents' concerns that the red 'no' would be interpreted by the child in a negative way.

One suggestion was not possible to incorporate into the intervention, which was the desire for parents to have input into the choice of sound targets. The experimenter systematically selected sound targets from a pre-determined list by following a written protocol ('Sound Selection protocol, Additional File X). This was done in order to retain experimental control over this aspect of the study and remove potential sources of bias. Nevertheless, it is best practice for speech-language therapists to liaise with those who know the child best when selecting intervention targets, and so for future studies this could be incorporated into the protocol.

A limitation of the focus group was that despite inviting therapists who are involved in language interventions for autistic children, there were none present. This may have been due to the time of day of the focus group, which was primarily scheduled to suit parents. The first author held several bilateral discussions with speech-language and behavioural therapists, to seek additional confirmation that the aims, mechanisms and scope of the intervention were appropriate. No changes to the proposed app arose from these discussions.

Appendix H: Stage 2 Consultation

1. Aim

To trial the prototype app to establish whether there were any technical difficulties with it, or if any features needed clarifying or amending prior to the pilot study.

2. Methods

Participants

The first author gathered a convenience sample using personal contacts. Participants needed to have access to an android mobile device. Three participants (V, S and J) had children who are or were preverbal with additional needs, one participant (A) had an autistic parent and sibling, three (L, D and E) were typically developing adults. The participants used a range of android devices.

Procedure

The first author sent each participant a copy of the app either in an email or via a file sharing service, with a pdf of the proposed instruction booklet, including details of how to download the app. Approximately one week later, the first author sought feedback via telephone and email regarding the ease of download, use, and any technical problems that arose.

3. Results

Feedback from each participant is summarised in Table B below.

Table B: Feedback following app testing exercise

Participant	Feedback
V (mother to R, 18 months old with hemiplegia)	<ul style="list-style-type: none">• <i>the 'end recording' button for video capture appears at the side instead of the bottom of the screen - it isn't clear which button to press to end the recording. It would be better either at the bottom centrally or with wording explaining what it was for.</i>• <i>R could not attempt any of the sounds, so we were in a loop of attempt -> try again -> attempt -> try again. There were no instructions about having a chance to stop this sound and try another to vary it [the use of the back button here could be made clearer in the instructions.]</i>• <i>it would be good if the camera could be activated when you are in the menu to customise photos, i.e. so you don't have to take a photo first and then select it from the gallery</i>• <i>one of the audio stimuli was a lot quieter than all the rest (m quieter than a and t)</i>

	<ul style="list-style-type: none"> • <i>R loved the pics and vids though and found them motivating, just wasn't able to make his own attempts.</i>
S (mother to D, 5 years old with verbal dyspraxia)	<ul style="list-style-type: none"> • <i>The app loaded without a problem. The app was easy to use. I went to settings and added some of my own pictures which was surprisingly easy to do. I take it when app is fully developed you would be able to taylor [sic] sounds to those that need to focus on/ choose our own.</i> • <i>When I used this with my son (aged 5) he really liked it. He engaged well and copied the cued articulation signs from the video while saying the word. He really liked that it videoed him and watched it back. This will really help as he is working on the shape of his mouth/lips when forming sounds.</i> • <i>My son has been having speech therapy for the past three years and we are always trying to find new ways to do things to keep him interested and motivated. This app will certainly do this. I would happily pay for this app.</i>
J (father of an autistic child with dyspraxia)	<ul style="list-style-type: none"> • <i>in test mode why are there numbers next to the letters?</i> • <i>it would look better if the letters were lower case to show that they are sounds and not capitals</i> • <i>I missed the playback on one of the attempts, it's a shame there is no 'replay' button for if the child does a really good attempt</i>
A (has two autistic relatives)	<ol style="list-style-type: none"> 5. <i>I think the starry background should instead be a plain one as it might be distracting</i> 6. <i>some autistic people find mirrors aversive or overly distracting, it would be great to have the option to turn that mirror view off whilst still recording</i>
L (student)	<ol style="list-style-type: none"> 7. <i>Managed to load it fine. On my first few log in attempts it said 'please wait until 0. Remaining: 2' under the password line and the numbers seem to increase and decrease and stop for a while. With a few attempts and refreshes it did let me in though. When I move off the app it logs me off straight away and it takes a few attempts to log back in due to the 'please wait' thingy. Having a go now and all the videos seem to be playing easily and the record function seems to work. Its great!</i>
D (student)	<ol style="list-style-type: none"> 8. <i>I could download the app but not log in – crash report attached</i>
E (retired)	<ol style="list-style-type: none"> 9. <i>managed to load the app but closed itself each time the password was entered (no crash report).</i>

4. Discussion

The testing process highlighted technical problems, which were forwarded to the developer to be addressed. Examples of these were: re-recording some of the sound stimuli so that they were all at the same volume, and understanding why the app was failing to log in on some devices.

There were two substantive improvements to the app generated by this process. Firstly, a replay button was added to the test and intervention trial screens, so that the attempt videos can be re-watched by parent-child dyads, providing a further opportunity for feedback and reinforcement. Secondly, an in-app camera function was added, so that parents could take photos of stimuli directly from the customisation menu, in order to facilitate customisation. Other changes were more cosmetic in nature, such as changing capital letters to lower case letters for describing the target sounds, altering the 'end recording' button to make it more prominent, removing a 'busy' background screen and making it plain. The instruction booklet emphasised how to use the 'back' button on android devices, for those only familiar with Apple ones.

One proposed change that could not be made for financial/logistical reasons was the proposal to have the option to turn off the mirror function. This was based on the fact that the participant's autistic relative has demonstrated an active dislike of mirrors. In hindsight, following the feedback from a couple of participants in the pilot study, this change may have improved the app considerably for the minority of children who did not appear to enjoy the mirror function.

It became apparent in this testing process that not all android devices could support the app reliably. This prompted us to ensure the participants in the pilot study all had access to the same android tablet, which had been tested extensively with the app. Despite this testing process, login problems still persisted in the pilot study, and are considered to be one reason for the variable rates of adherence and engagement observed.

Appendix I: Sound Target Protocol

This decision tree is designed to generate a list of 9 selected sounds for each child from which target sounds will be drawn. The first three sounds on the list will be the first three targets and from then on, any replacements will go down the list in order.

It is not expected that any child will already be able to produce more than 4 of the sounds mentioned below, therefore each child will have nine sounds on their list.

Decision tree: Stop when 9 unique sounds have been selected (see illustration below)

1. Can they say b?

YES - p (if not mastered, otherwise go to 2)

NO - b

2. Can they say n?

YES - go to 3

NO - n

3. Can they say d?

YES - t (if not mastered, otherwise go to 4)

NO - d

4. Can they say ee?

YES - oo (if not mastered, otherwise go to 5)

NO - ee

5. Can they say a?

YES - go to 6

No - a

6. Can they say p?

YES - m (if not mastered, otherwise go to 7)

No - p

7. Can they say t?

YES - go to 8.

NO - t

8. Can they say oo?

YES - ar (if not mastered, otherwise go to 9)

No - oo

9. Can they say m?

YES - go to 10

NO - m

10. Can they say ar?

YEs - go to 11

No - ar

11. can they say w?

YES - go to 12

No - w

12. Can they say h?

YES - go to 13

NO - h

13. Can they say i?

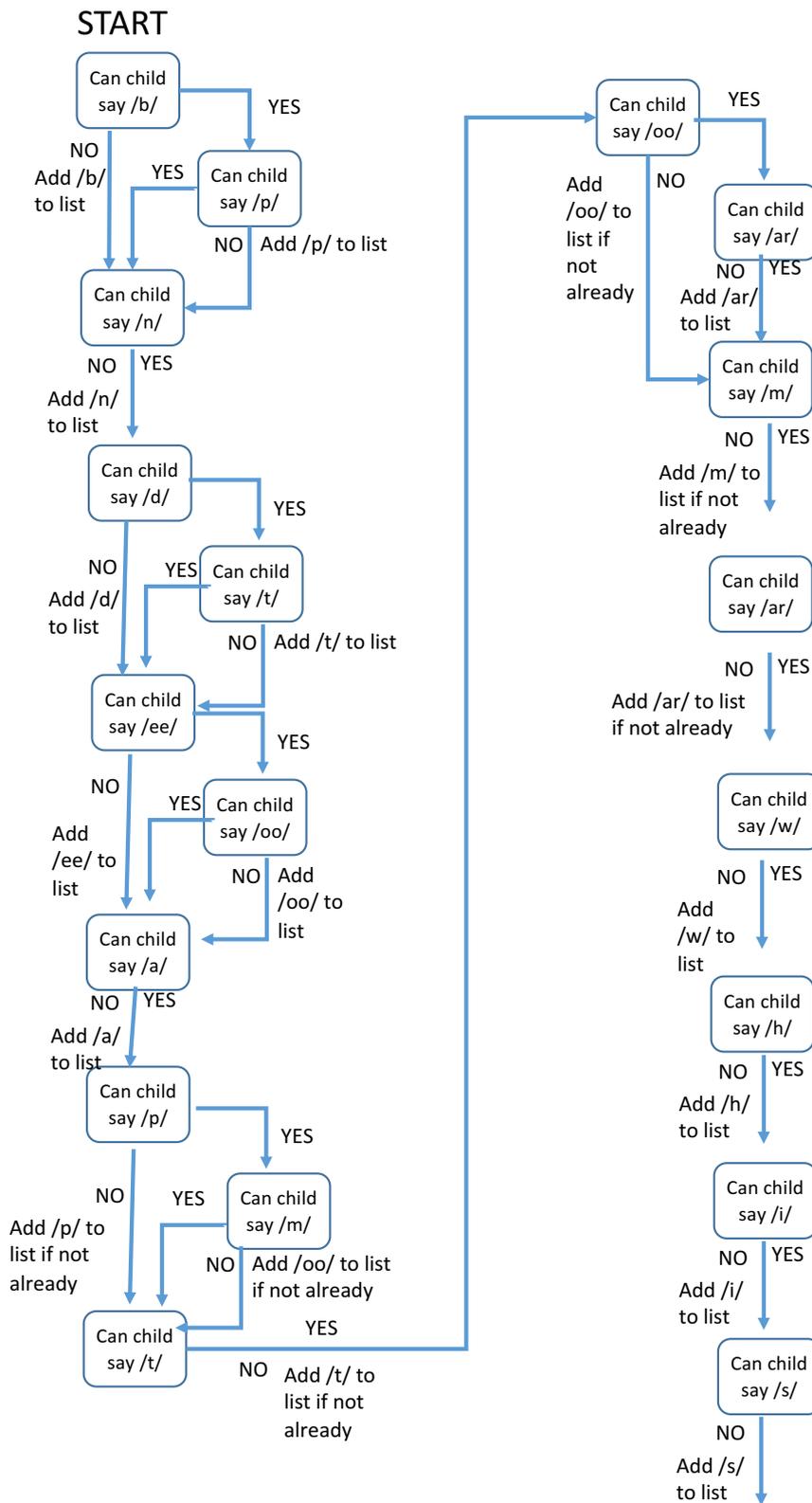
YES - got to 14

NO - i

14. can they say s?

YES - end

NO - s



Decision Tree for Speech Sound Protocol

Appendix J: App User Questionnaire

App name: BabbleBooster (Android Version)

Reviewer name: _____

Date: _____

	4	3	2	1
<u>Relevance:</u> Does the app target the skill it aims to target (=rehearsal of single speech sounds)?	The app's focus has a strong connection to the purpose for the app and was appropriate for the learner	The app's focus is related to the purpose for the app and mostly appropriate for the learner	Limited connection to the purpose for the app and may not be appropriate for the learner	Does not connect to the purpose for the app and not appropriate for the learner
<u>Engagement:</u> Does the app provide enough motivation to hold the interest of the learner?	Learner is highly motivated to use the app throughout the time	Learner is motivated to use the app most of the time	Learner somewhat engaged, but lost motivation after a short time	Learner avoids the use of the app
<u>Customization:</u> Was the customization able to meet the needs of the learner?	I was fully able to alter content to meet learner needs	I was able to alter content to meet learner needs but would like further flexibility	I was able to customize app but this was of limited value	Unable to access customization options
<u>In app feedback (reward videos):</u>	Feedback was clear and strongly aided motivation	Feedback somewhat aided motivation	Feedback was neutral / learner did not connect performance with feedback	Feedback was distracting or counterproductive

	4	3	2	1
<u>Layout/Design:</u>	Layout is simple, logical consistent and easy to navigate with no distracting features	Layout is fairly straightforward with a few minor problems	Layout seemed counterintuitive and unhelpful	I did not like the layout at all
<u>Installation:</u>	I could launch the app independently from first use	I needed limited initial support in launching the app	I needed frequent support with installation	I was unable to install this app
<u>Navigation:</u>	Very easy to learn how to use and directions are clear and simple to follow	Easy to learn and direction can be followed	Quite difficult to learn and follow	Very complex to learn – not user-friendly
<u>Support:</u> Evaluate quality of written instructions, video tutorial, in-person training and offline support	Support was very easy to access, helpful and clear	I found support mostly helpful, clear and accessible	Support was difficult to access and/or of limited use	I did not find the support options useful at all
<u>Reporting:</u>	I found the progress summary data easy to access, intuitive and encouraging	The progress summary data was quite helpful	The progress output was of limited use	I couldn't access and/or understand the progress output
<u>Technical performance:</u> Including video capture	Performs and loads quickly. No issues and very reliable	Performs and Loads quickly. Some minor technical issues.	Loads and performs slowly. Sometimes Crashes.	Crashes fairly often and takes multiple times to open.

OVERALL COMMENTS

1. What overall rating would you give the app?



2. Does this app allow you to do something you were unable to do with your child in the past? If so what?

3. What are the strengths of this app?

4. What are the weaknesses of this app?

5. Comments and recommendations:

Appendix K: Score Breakdown of Acceptability Questionnaire

Domain	Factor	N	Mean Score	Sd
1	Engagement	17	2.06	0.90
2	Technical performance	17	2.62	0.93
3	Reporting	11	2.82	1.47
4	Customization	14	2.86	0.77
5	In-app feedback	16	2.97	0.87
6	Relevance	17	3.35	1.06
7	Installation	15	3.53	0.74
8	Navigation	17	3.71	0.47
9	Layout / design	17	3.82	0.39
10	Support	17	3.88	0.33

Note: SD: Standard Deviation.

Appendix L: Simulation

Hypothetical datasets were simulated 100 times each in R (using code below), for a range of effect sizes (.5, .75, 1 and 1.25) with the following assumptions:

- 18 participants
- alpha of .05
- 16-week experimental period with a minimum of 3 weeks baseline (A) and 6 weeks intervention (B), resulting in 8 permutations of A and B weeks

Intra-class correlation between scores of each participant was either assumed to be .25 (low autocorrelation results) or .75 (high autocorrelation results), to model different assumptions about how similar within participant scores are to each other

As well as modelling the power of a randomization test to detect a real treatment effect under these parameters, I calculated a comparable power metric for the same dataset, assuming a group design. To do this I randomly assigned half the participants to a control group and half to an intervention group. Mean week 8 baseline score is subtracted from mean week 16 baseline score in the control group (no mean treatment effect expected) and mean week 8 baseline score is subtracted from mean week 16 intervention score in the intervention group. The two group mean differences are then compared using an independent t-test.

Simulation results are presented in Figure I.

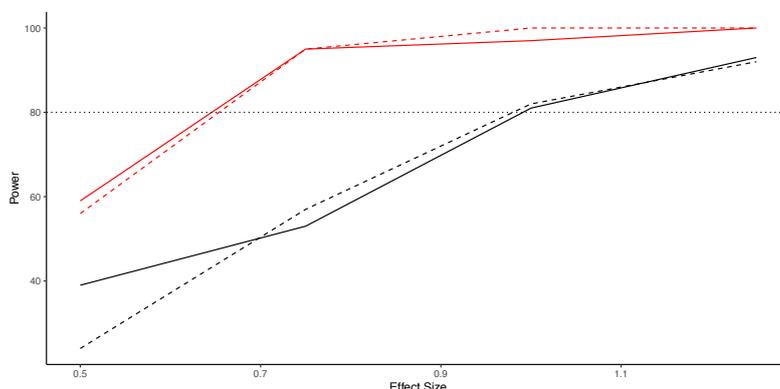


Figure I

Power to detect treatment effects of different sizes, assuming $n=18$, 100 simulations, 8 permutations of A and B weeks, $\alpha=.05$. Black lines represent power in group design, red lines represent randomized case series design. Solid line assumes $ICC=.25$ (low autocorrelation) and dashed line assumes $ICC=.75$ (high autocorrelation)

From the above graph, I conclude that neither analysis method (randomisation test or independent t-test) is particularly sensitive to autocorrelation parameters. Furthermore, the randomized case series design has adequate power to detect a real effect size of .65 or more, whereas for the group design an effect size of 1 is required to achieve adequate power.

I also illustrate an example dataset (from one simulation, effect size = 1) for each of the high and low autocorrelation scenarios in Figures II to V.

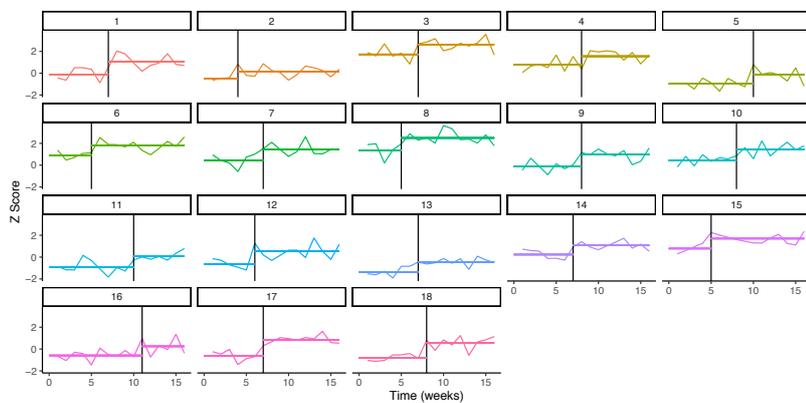


Figure II. Example of low autocorrelation simulation scores over time (ICC=.25)
 Z scores over time in weeks, each panel represents one participant, vertical line indicates intervention start, horizontal lines represent mean A week and B week scores

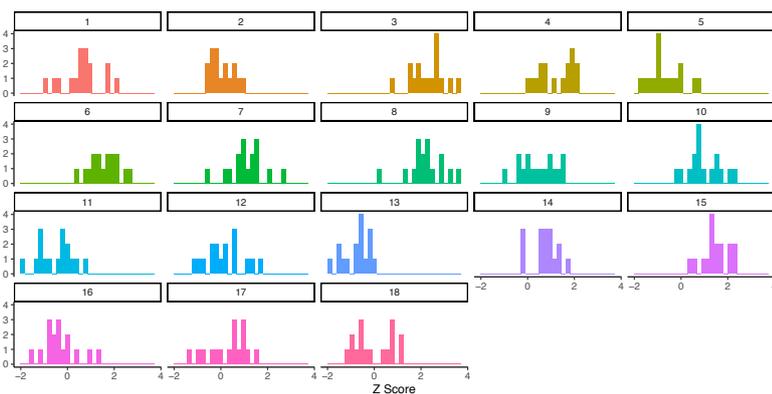


Figure III. Example of low autocorrelation simulation scores histogram (ICC=.25)
 Z scores distribution, each panel represents one participant

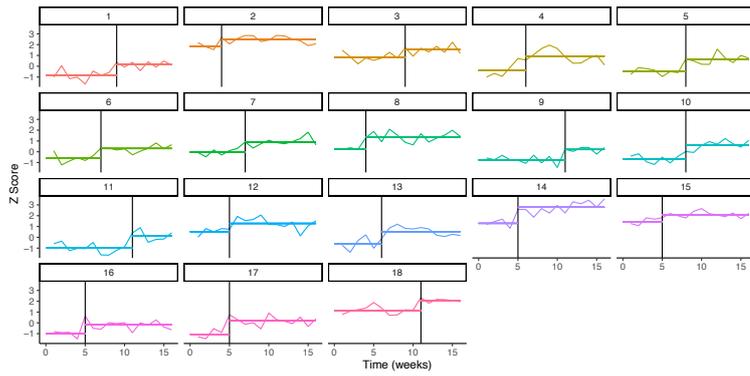


Figure IV. Example of high autocorrelation simulation scores over time (ICC=.75)
 Z scores over time in weeks, each panel represents one participant, vertical line indicates intervention start, horizontal lines represent mean A week and B week scores

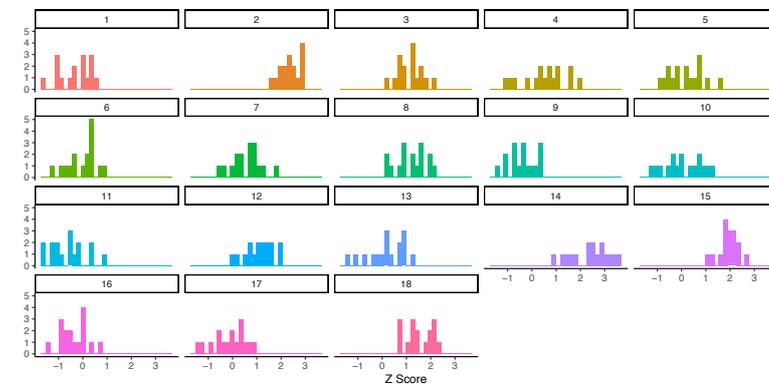


Figure V. Example of high autocorrelation simulation scores histogram (ICC=.75)
 Z scores distribution, each panel represents one participant

Simulation Code

```
#install key packages
install.packages("metap") #needed for sumz() function
library(metap)
install.packages("multilevel") #needed for sumz() function
library(multilevel)
install.packages("tidyverse") # enables use of tidyverse family of packages
library(tidyverse)
install.packages("scdhlm")
library(scdhlm)

# clear workspace
rm(list=ls())

#-----
# establish variables (this should be all you need to change)
#-----
myN <- 18 # number of cases in the series, i.e. your planned number of participants must be
even number
mySim <- 100 # number of simulations you want to run
myM2 <- 1 # intervention mean score (equates to effect size given that baseline has mean 0 and
SD of 1)
myW <- 16 # number of weeks in total that your experiment will last (alter this figure to change
the number of permutations, or alter the length of pre- and post phases below)
myICC <- 0.75 # correlation within participants
myCor <- 0 # correlation between participants
myMissingData <- 0 # % of missing data points (0=none, .2=20% missing) This only affects
weeks 2 to myW-1 in order to avoid having NAs in the mean difference results

#-----
# these variables can also be changed if needed
#-----
myPre <- 3 # min number of baseline weeks before intervention can begin. So if this is 3, the
first possible week for intervention start is week 4
myPost <- 6 # min number of weeks that intervention must last. So if this is 6, the last 6 weeks of
your experiment are intervention for everybody
myAlpha <- 0.05 # the alpha level

#-----
# derive variables
#-----
myInt <- myW - myPre - myPost #number of weeks with variable allocation
myPerm <- myInt + 1 #number of permutations possible based on the intervention weeks
rangeN <- 1:myN #range of rows for participants
rangeW <- 1:myW #range of weeks in total
rangeInt <- (myPre + 1):(myPre + myInt) #range of weeks where intervention is variable i.e. from
wk3 to wk10 out of 16
totalN <- myN*myW #number of scores to be generated in the df (gives sample size for
distribution)

#-----
# create empty df for simulation results to go in
#-----
dfSim <- data.frame(matrix(data=NA, nrow=mySim, ncol=5)) # makes a reference dataframe to
hold results in
colnames(dfSim) <- c("Sim", "p", "sig", "groupstudy", "groupsig") #group study and group sig refer
to the equivalent case if a group study had been run
```

```

#-----
#begin big loop (whole process repeated for each of 1:k simulations)
#-----
#-----
for (k in 1:mySim)

{

# 1. make the intervention allocation table for myN participants over myW using randomisation

rangeRand <- sample(1:(myPerm), myN, replace=TRUE) # assigns each participant a random
permutation
dfRef <-data.frame(matrix(data=NA,nrow=myN,ncol=(2+myInt))) # makes a reference dataframe
to hold results in and populates with participant number and random permuation allocation
colnames(dfRef) <-c("n","Rand",rangeInt)
dfRef$n <-rangeN
dfRef$Rand <- rangeRand

#-----
# smaller loop to populate 0s and 1s to denote which weeks of the intervention period are
allocated to each participant in dfRef
for(j in 1:myPerm)
{
for(i in 1:myN)
{
if(dfRef$Rand[i]==j) {dfRef[i,3:(myInt + 2)] <- c(rep(0,j-1),rep(1,(myPerm-j)))};
# for the dfRef table first 2 columns are taken up so range begins at 3 and is 2 greater than
myInt
# this creates a matrix of the weeks which are variable (rangeInt) for each person
}
}
}
#-----

# df incorporating the above plus the mandatory pre and post weeks before and after it
dfRand <- data.frame(matrix(data=NA,nrow=myN,ncol=myW)) #empty df
dfRand[1:myPre] <-0 #all pre weeks are 0
dfRand[(myPre+myInt+1):myW] <-1 #all post weeks are 1
dfRand[rangeInt] <- dfRef[3:(2+myInt)] #weeks in the intervention period are copied from dfRef

# 2.Create the simulated scores for A weeks and B weeks (accounts for intraclass correlation
within participants)

dfScore1 <- sim.icc(myW, 1, 0.75, nitems=myN,item.cor=0) #multilevel package
dfScore1$GRP <- NULL # don't need this column
dfScore1 <- as.data.frame(t(dfScore1)) %>% scale() %>% as.data.frame() # creates the df for
baseline

dfScore2 <- dfScore1 + myM2 # adjusts the baseline by effect size

# 3. combine matrices from (1) and (2) to make a matrix of the actual score (depending if it is a 0
or a 1 week)

dfScore <- (dfRand*dfScore2) + ((-dfRand+1)*dfScore1)

# 4. Model missing data here

# create a df of appropriate size

```

```

dfNA <-data.frame(matrix(data=NA,nrow=myN,ncol=myW)) #empty df
# calculate how many cells have option to be missing (A) - first and final week cannot
missA <- (myW-2)*myN

# calculate number of cells to change to indicate NA (rounding this to integer) (B)
missB <- round(missA*myMissingData,0)

# create a vector with (A-B) reps of 1 and B reps of NA, sample from this, and fill the space of A
with it
missAB <- sample(c(rep(1,missA-missB), rep(NA, missB)), (myN*(myW-2)), replace=FALSE)
dfNA[1] <-1
dfNA[myW] <-1
dfNA[2:(myW-1)] <- missAB

dfScore <- dfNA*dfScore # create the final scores you will use, including NAs

# 5. Create a table of pre, post and Mean Difference (MD) scores for each possible permutation
for each person

dfPre <-data.frame(matrix(data=NA,nrow=myN,ncol=myPerm)) #empty df
#-----
#smaller loops to calculate the baseline score for each permutation for each participant
for(j in 1:myPerm)
{
for(i in 1:myN)
{
{dfPre[i,j] <- sum(dfScore[i,1:(myPre+j-1)],na.rm=TRUE)/sum(dfNA[i,1:(myPre+j-1)],
na.rm=TRUE)}; #adjusts for NAs
}
}
}
#-----

dfPost <-data.frame(matrix(data=NA,nrow=myN,ncol=myPerm)) #empty df
#-----
#smaller loops to calculate the post-intervention score for each permutation for each participant
for(j in 1:myPerm)
{
for(i in 1:myN)
{
{dfPost[i,j] <- sum(dfScore[i,(myPre+j):myW], na.rm=TRUE)/sum(dfNA[i,(myPre+j):myW],
na.rm=TRUE)};#adjusts for NAs
}
}
}
#-----

dfMD <-data.frame(matrix(data=NA,nrow=myN,ncol=myPerm)) #empty df
dfMD <- dfPost-dfPre # calculates a difference score between the two means

# 6. Populate the dfRef table with the actual MD, pre and post score (needed for graph)

dfRef$MD <- NA #empty column
#-----
#smaller loops to select the right one
for (j in 1:myPerm)
{
for (i in 1:myN)
{
if(dfRef$Rand[i]=j) {dfRef$MD[i] <- dfMD[i,j]};
}
}
}

```

```

#-----

# used for the graph
dfRef$Pre <- NA #empty column
#-----
#smaller loops to select the right one
for (j in 1:myPerm)
{
for (i in 1:myN)
{
if(dfRef$Rand[i]==j) {dfRef$Pre[i] <- dfPre[i,j]};
}
}
#-----

# used for the graph
dfRef$Post<- NA #empty column
#-----
#smaller loops to select the right one
for (j in 1:myPerm)
{
for (i in 1:myN)
{
if(dfRef$Rand[i]==j) {dfRef$Post[i] <- dfPost[i,j]};
}
}
#-----

# 7. Create comparison table to count the permutatations which are greater than or equal to the
actual MD

dfComp <-data.frame(matrix(data=NA,nrow=myN,ncol=myPerm)) #empty df
#-----
#smaller loops to place a one for each permutation where the actual allocation MD is >= the
permutation MD
for (j in 1:myPerm)
{
for (i in 1:myN) {

if(dfRef$MD[i]>=dfMD[i,j]) {dfComp[i,j] <- 1};
}
}
#-----

# 8. populate the dfRef table by adding together the 1s in each row of dfComp to determine
relative position / p-value. e.g.all 4 1s means actual MD is higher or equal to all -> p value 1/4
#-----
#smaller loop
for (i in 1:myN)
dfRef$Comps[i] <- sum(dfComp[i,1:myPerm], na.rm = TRUE) #new col of dfRef to count the #of
permutations it is >=than
#-----

dfRef$pval <-((myPerm - dfRef$Comps) + 1)/(myPerm + 0.0000001) #derives the p value (NB
+0.000001 is so that no pval is 1.0)

# 9. Using the derived p-values, calculate Stouffer's Z / p-value for the sample (as per Rvachew &
Matthews, 2017)

mySumz <- sumz(dfRef$pval) # this is the p-value for the test statistic

```

```

# 10. equivalent group study power? this uses mean diff

controlgroup <- 1:(myN/2) # first half of N assigned to group 1 (NB this only works with even
numbers of N)
intgroup <- (myN/2+1):myN # second half of N assigned to group 2

controlPost <- dfScore1[controlgroup,myW] - dfScore1[controlgroup,(myW/2)] # final week score -
half point score from the dfScore1 distribution (this assumes no change in mean)

intgroupPost <- dfScore2[intgroup,myW] - dfScore1[intgroup,(myW/2)] # final week score from
dfScore2 distribution - half point score from dfScore1 distribution (assumes change in mean)

ttest <- t.test(controlPost, intgroupPost) #t test of both mean differences calculated above
(between groups)

ttest$p.value # this is the p-value assuming a group study

# 11. Populate dfSim with the results from this simulation (k=simulation number)

dfSim$Sim[k] <-k # Sim number
dfSim$groupstudy[k] <- ttest$p.value # group study p-value
dfSim$p[k] <- mySumz$p # case series p-value

# this code checks if either p-values are greater than the alpha value
if(dfSim$p[k]> myAlpha){dfSim$sig[k] <- 0} else dfSim$sig[k] <- 1;
if(dfSim$groupstudy[k]> myAlpha){dfSim$groupsig[k] <- 0} else dfSim$groupsig[k] <- 1;

} #end of bigger loop
#-----
#-----

#this line derives the power calc #number of Sims with significant p-value
print(paste0(
"Simulation details: Effect Size= ",
myM2,
", n= ",
myN,
", number of weeks= ",
myW,
", power estimation of case series= ",
mean(dfSim$sig)/1*100,
", power estimation of equiv group study= ",
mean(dfSim$groupsig)/1*100))

```

Appendix M: Educational Setting and Communication System Questionnaire

We would like to understand more about the type of educational setting your child has been in over the past 12 months. We also have questions regarding any alternative form of communication that your child might have been exposed to or use (like sign language or picture exchange cards).

1. Please tick the description that fits your child's educational setting over the past 12 months. If your child has been in two settings in the past 12 months please tick both and specify dates.

- at home with no formal nursery/school placement
 - at mainstream nursery/school without support, full/part time
 - at mainstream nursery/school with limited support, full/part time
 - at mainstream nursery/school with dedicated 1:1 support, full/part time
 - at specialist nursery (private), full/part time
 - at specialist nursery (state-funded) , full/part time
 - at a specialist unit in a mainstream school, full/part time
 - at special school, full/part time
 - other (please specify)
-

2. Please tick the description of your child's EHCP status

- My child has no EHCP and we have not applied for one
- We have begun the process and are waiting for our EHCP to be completed
- We received a draft EHCP but are currently appealing it
- We obtained a final EHCP on _____ date

3. How straightforward was it to obtain an EHCP and/or obtain an educational placement for your child? Please include any details you feel are relevant.

4. Which of the following augmentive alternative communication systems has your child been exposed to (either at home, in education setting or via speech and language therapy)

- PECS
- makaton
- AAC device (e.g. proloquo)
- other (please specify) _____

5. Does your child regularly request 10 or more items using any of the following augmentive alternative communication systems:

PECS

makaton

AAC device (e.g. proloquo)

other (please specify) _____

Appendix N: Code to Calculate Randomisation Test

```
# firstly make table of all MD possibilities and tag which ones are greater than the actual one

#load in data
dfScore <- read.csv("weekly.csv")

# creates df of 1s and 0s for missed weeks (so mean can be adjusted below)
dfNA <- is.na(dfScore)
dfNA[1][dfNA[1] %in% TRUE] <- 0
dfNA <- (dfNA - 1) * - 1

myPre = 4 # no intervention for first 4 weeks, including week 0
myW = 17 # 17 weeks in total including week 0
myPerm = 8 # 8 possibilities - starting week 4:11, means weeks 11:16 always B (6 weeks)
myN = 10

dfRef <-data.frame(matrix(data=NA,nrow=myN,ncol=3)) # makes a reference dataframe to hold
results in and populates with participant number and random permutation allocation
colnames(dfRef) <-c("n","Rand","MD")
dfRef$Rand <-realrange - myPre +1 # derive permutation number (1=starts in week 4, 6= startsin
week 11 )
dfRef$n <- rownames(weekly)
# calculates all the possible mean differences (just like in the simulation)

#correcting the 12/10 to an 8 to make the analysis work (only 8 permutations)
dfRef$Rand[7:8] <- 8

dfPre <-data.frame(matrix(data=NA,nrow=myN,ncol=myPerm)) #empty df
#smaller loops to calculate the baseline score for each permutation for each participant
for(j in 1:myPerm)
{
for(i in 1:myN)
{
{dfPre[i,j] <- sum(dfScore[i,1:(myPre+j-1)], na.rm=TRUE)/sum(dfNA[i,1:(myPre+j-1)],
na.rm=TRUE)};#
}
} #eg for permutation 1, pre score is that in the range 1:4 (weeks 0 to 3)
```

```

dfPost <-data.frame(matrix(data=NA,nrow=myN,ncol=myPerm)) #empty df
#smaller loops to calculate the post-intervention score for each permutation for each participant
for(j in 1:myPerm)
{
for(i in 1:myN)
{
{dfPost[i,j] <- sum(dfScore[i,(myPre+j):myW], na.rm=TRUE)/sum(dfNA[i,(myPre+j):myW])};
}
}
}#eg for permutation 1, post score is that in the range 5:17 (weeks 4 to 16)

```

```

dfMD <-data.frame(matrix(data=NA,nrow=myN,ncol=myPerm)) #empty df
dfMD <- dfPost-dfPre # calculates a difference score between the two means

```

```

#smaller loops to select the right one for refernce table
for (j in 1:myPerm)
{
for (i in 1:myN)
{
if(dfRef$Rand[i]==j) {dfRef$MD[i] <- dfMD[i,j]};
}
}
}#looks up value in the dfMD table for each permutation (Rand=permutation number)

```

```

dfComp <-data.frame(matrix(data=NA,nrow=myN,ncol=myPerm)) #empty df -
#smaller loops to place a one for each permutation where the actual allocation MD is >= the
permutation MD
for (j in 1:myPerm)
{
for (i in 1:myN) {
if(dfRef$MD[i]>=dfMD[i,j]) {dfComp[i,j] <- 1};
}
}
}

```

Secondly, add together the 1s in each row of dfComp to determine relative position / p-value.
e.g.three 1s means actual MD is higher or equal to all -> p value 1/3

```
for (i in 1:myN)
dfRef$Comps[i] <- sum(dfComp[i,1:myPerm], na.rm = TRUE) #new col of dfRef to count the #of
permutations it is >=than
dfRef$pval <-((myPerm - dfRef$Comps) + 1)/(myPerm + 0.00001) #derives the p value (NB +0.001
is so that no pval is 1.0)
```

Using the derived p-values, calculate Stouffer's Z / p-value for the sample (as per Rvachew & Matthews, 2017)

```
mySumz <- sumz(dfRef$pval) #need metap
mySumz$validp # to print out the p vals
mySumz$z #ns effect ***** z= .314 p.38
```