A DYNAMIC ASSESSMENT OF SINGLE WORD LEARNING IN TWO YEAR OLD LATE TALKERS

A thesis submitted in partial fulfilment of the requirements for the Degree of Master of Science in the Department of Communication Disorders

Victoria Singer

University of Canterbury

2014

TABLE OF CONTENTS

List of Tables

List of Abbreviations

AC – Auditory Comprehension

CDI – MacArthur-Bates Communicative Development Inventory: Words and Sentences

DA – Dynamic Assessment

EC – Expressive Communication

LB – Late Bloomer

LD – Language Delayed

LI – Language Impaired

LT – Late Talker

PLS – PreSchool Language Scale, Fourth Edition

SD – Standard Deviation

SA – Static Assessment

TD – Typically Developing

Acknowledgements

Firstly, I wish to thank all at the Department of Communication Disorders and specifically, my supervisors, Professor Thomas Klee and Associate Professor Catherine Moran for their guidance, encouragement and expertise throughout this process. Their commitment and patience has been very much appreciated.

I would also thank to thank my friends and colleagues Alison Fardey, Melissa Brazier, Sue Doris and Dr Rachel Beard for a huge amount of support, ranging from administrative assistance and the loan of test materials, to facilitating vital recruitment connections and keeping me motivated and on track.

Thank you to Rosemary Emery for her proof reading, assistance with scoring and invaluable ‘highly experienced’ speech-language therapy perspective.

Thank you also to my parents, Kingsley and Susanna Singer for their unerring support throughout all my study.

And lastly, this work is dedicated to the late Alan Smyth, an amazing uncle, who inspired creative problem-solving and practical solutions.

Abstract

Purpose: Between 13% and 20% of two year olds are late to talk; of those, up to 25% are at risk of persistent language impairment. This highly exploratory study examined whether a dynamic assessment (DA) of single word learning could be used to predict medium term development trajectory and thus provide better information regarding which late talkers were most at risk of continued language delay.

Method: Six novel non-words were taught within a scripted play activity, which controlled for number of exposures. Retention and recall of each novel word was tested following both 3 and10 exposures using a predetermined hierarchy of prompts. Participants were 20 typically developing children and 20 late talkers aged 24-29 months. The late talking group (mean age 26 months) was tracked for 8 to 9 months with re-administration of the DA task 3 months after the initial testing. The task employed a graduated prompting framework because it is highly scripted and can be completed within a single brief session; advantageous for screening purposes.

Results: Findings indicated that the DA scores for single word learning were associated with change over an 8 to 9 month period. The association between the task and standardised assessment (PLS) change scores was observed to increase over a 3 month period, when the average age of the late talking participants was 29 months. At this time, participants achieving DA scores more closely approximating those of typically developing children were operating within the normal range on standardised testing (PLS) 5 to 6 months later.

Conclusions: More accurate differentiation of children who were late blooming versus those likely to be language impaired was achieved closer to 2 ½ years of age. Implications for service provision in terms of directing input to where it is most needed and also in identifying most optimal timing for input are discussed.

Chapter 1: Literature Review

Late Talkers: Prevalence and Predictors

It is broadly estimated that between 10 and 20% of 24 month olds are Late Talkers (LTs) with no clear explanation for their language delay. Parent report, as measured by the Language Development Survey (LDS) (Rescorla, 1989), and a criterion of fewer than 50 words or no word combinations has put this figure at 15% (Klee et al., 1998), 13% (Rescorla & Achenbach, 2002), and more recently 19.1% (Zubrick, Taylor, Rice, & Slegers, 2007). With the same cohort, Zubrick, et al. (2007) obtained a figure of 13.4% using an alternative tool and cut-off of – 1 SD on the Communication scale of the Ages and Stages Questionnaire (ASQ) (Bricker & Squires, 1999). A third instrument, the MacArthur-Bates Communicative Development Inventory (CDI) (Fenson, 2007) has also been used. Similar to the LDS, this is a parent checklist with percentile scores based on age in months but with a larger pool of potential vocabulary items. Using the 10th percentile as the cut-off has elicited a figure as high as 19.7% (Reilly et al., 2009) for children aged 23.5 to 25.5 months. Of note, when considering use of percentile cut-off figures, it should be acknowledged that there may be a broad range in total number of words produced that correspond to the 10th percentile, depending on the age range of the sample. Also, differing figures will be obtained depending on whether the criterion is based on vocabulary alone or on vocabulary alongside a word combining measure.

Many of these 24 month old LTs catch up during the preschool years and are Late Bloomers (LB) operating within the normal range by age 3 years (Rescorla, Roberts, & Dahlsgaard, 1997). Paul and Roth's (2011) recent review of the literature suggested up to 75%, with only a minority of this number receiving language intervention; however, as noted by Paul (1996) up to 26% will have a persistent language impairment at 4 to 6 years. Rescorla

(2002) has also suggested that greater numbers form a subclinical group; a finding further supported by evidence of on-going difficulties with syntax, and more specifically, grammatical tense marking in a group of late talkers followed up at age 7 years (Rice, Taylor, & Zubrick, 2008). There is also evidence that although these Resolved Late Talkers (RLT) can obtain scores in the normal range, they are still likely to perform lower than their typically developing counterparts (Rescorla, 2009). Given the benefits of Early Intervention in altering developmental trajectories, there is a strong case for determining the degree to which this late start may be prognostic for longer lasting risk.

A number of child-centred predictors of change and/or “red flags” have been documented (Olswang, Long, & Fletcher, 1997). These include less diverse verb repertoires with a higher production of general-all-purpose (GAP) verbs and fewer intransitive and ditransitive items. Also, the presence of: delays greater than 6 months in comprehension, limited number of consonants in babble, few spontaneous imitations, lack of object or symbolic play, few communicative gestures or vocalisations, communicative intents limited to requesting, and difficulty gaining access to peer interactions (Olswang, Rodriguez, & Timler, 1998; Paul & Roth, 2011).

Within this, inventory of conventional gestures has been determined more predictive of later receptive ability, and communication for joint attention and inventory of consonants, more significantly related to expressive outcome (Watt, Wetherby, & Shumway, 2006). The compounding effect of comprehension and socio-pragmatic deficits has also been acknowledged by the delineation of three sub-types of LT (Desmarais, Sylvestre, Meyer, Bairati, & Rouleau, 2010). These are differentiated by the cumulative addition of engagement in communication and comprehension to the primary expressive vocabulary delay. The expressive delay also grew in severity with the addition of these other variables.

Olswang and Bain (1996) further specified that the discrepancy between the child's receptive and expressive skills had more predictive value than any gap between language and non-verbal cognition.

Given these varied developmental pathways, and the challenge in differentiating between those likely to resolve and those likely to develop a more persistent language disorder, there is a need to explore how these "red flags" and other potential contributors might combine to achieve greatest predictive effect. One method which has been shown to be useful in differentiating language difference from language disorder is dynamic assessment (DA). Although the use of DA has not been studied directly with LTs, its value in clinical-decision making has been shown in a number of other areas. An overview of DA procedures, and a case for applying a DA graduated prompting methodology to the assessment and monitoring of 2-year old LTs is outlined below.

Dynamic Assessment

Vygotsky (1978) proposed that all learning takes place within a "zone of proximal development". This zone represents the difference between the child's independent performance, and what they can achieve in collaboration with a more experienced other. Dynamic assessment is therefore an umbrella term (Hasson & Botting, 2010) for a range of approaches which focus on measuring the gap between current and assisted linguistic performance. DA differs from traditional static assessment across three parameters: 1) an emphasis on process rather than product, that is, on how the child approaches the task rather than pre-existing knowledge; 2) the inclusion of examiner feedback; and, 3) a shift from examiner neutrality towards an individualised, teaching and helping relationship (Grigorenko & Sternberg, 1998).

Dynamic assessment should also be considered a supplement to static assessment, beginning where the child fails on these measures (Lidz, 2002). Within a DA approach, ‘failures’ are an opportunity to uncover the underlying processing strategies which, once identified, may be modifiable. Furthermore, dynamic procedures increase the ecological validity of the assessment process through more fully capturing the communicative-interactive aspect of language behaviour (Jitendra & Kameenui, 1993).

Dynamic Assessment and Clinical Decision-Making

As the medium for instructional interaction, DA therefore fits well within the field of speech-language therapy and has a key role in clinical decision-making. Underpinning the teaching and helping component of all DA approaches are two main methodologies: child-driven test-teach-retest and more task-driven graduated prompting (Grigorenko & Sternberg, 1998). Each has a number of advantages, and ultimately the approach selected should depend on the primary clinical purpose for administration. As will be highlighted, the two are not necessarily mutually exclusive either with many studies incorporating elements of both.

The use of DA with culturally and linguistically diverse (CLD) populations is one of the more extensively investigated clinical areas. This population is somewhat similar to LTs in that the principal clinical question is establishing whether an observed delay is likely to be temporary or indicative of a disorder. Many children lack familiarity with English and the interaction pattern of pointing to pictures (Laing & Kamhi, 2003). With static assessments biased toward what the child has learned prior to testing, there is a risk of both under and over identification of language impairment. Errors may be incorrectly assumed to be dialect related, or uncontrolled adjustments for variations in life experience and socialisation practice may simply skew the norm so that children with CLD still perform below the mean but within normal age limits (Hasson & Joffe, 2007; Laing & Kamhi, 2003). Conversely, a

dynamic teaching phase has the potential to provide almost all the information necessary for mastery of the task, and to eliminate this confound of prior learning (Elleman, Compton, Fuchs, Fuchs, & Bouton, 2011).

The classification accuracy of dynamic assessment in differentiating language disorder from language difference has been evaluated across a number of language domains. These include narrative ability (Moore-Brown, Huerta, Uranga-Hernandez, & Pena, 2006; Pena et al., 2006), single word learning (Gutierrez-Clellen & Pena, 2001; Pena, Iglesias, & Lidz, 2001; Pena, Quinn, & Iglesias, 1992) and categorisation (Ukrainetz, Harpell, Walsh, & Coyle, 2000). These studies employed a test-teach-retest format, with a teach phase or "Mediated Learning Experience" (MLE) characterised by four principles (Lidz, 1991): 1) mediation of intentionality which is a conscious attempt to influence the behaviour of the child by clearly communicating the purpose of the interaction; 2) mediation of meaning which focuses the child's attention on what is important; 3) mediation of transcendence which promotes cognitive bridges with the child's own experiences; and, 4) mediation of competence which teaches the child to be self-regulated and an active participant in their own learning. Competence also includes the provision of reinforcement and motivational support. Combined, the aim is to promote the child's conscious awareness and control of their abilities rather than simply provide assistance toward completing the task.

Scoring was completed using a combination of retest scores and modifiability ratings, the latter based more subjectively on qualitative learning behaviours observed during the teaching phase. All results indicated that initial static scores did not reliably differentiate the two groups but combined modifiability and post-test scores did. Pena et al.'s (2006) narrative DA determined that pre-test narrative analysis with 6 and 7 year old European American, African American and Latino American children was not sensitive to language impairment; however, combined measures of modifiability and post-test scores for number of different

words, total number of words, and Story Components yielded 100% correct classification. By itself, the modifiability measure was also the most accurate diagnostic measure. The goal within Pena, et al.'s (2001) word learning study was to make the 3 and 4 year old participants more aware of labels as a way of identifying pictures of objects. Two mediation sessions were administered, and in line with the aforementioned principles of mediation, there was no predetermined format to the MLE; rather the study focused on promoting self-directed learning. Teaching was done through drawing explicit attention to single word labels and providing explanations of when to use "special names". The study again concluded that there was a lack of significant difference between pre-test scores for typically developing and language impaired groups; however, language impaired children were less responsive to mediation and required a more intense effort on the part of the examiner to induce change. In terms of clinical placement, the post-test score was more valid. Again, this has considerable diagnostic value: difficulty with progressing without a high level of investment is integral to the definition of a language disorder (Hasson, Camilleri, Jones, Smith, & Dodd, 2013). Ukrainetz et al. (2000) replicated Pena's (2001) earlier work with 3 year old Native American children using a categorization rather than a labelling task. Mediation was intended to introduce the idea of grouping, and the principle of providing a category name, rather than to teach specific vocabulary items. Results again supported the construct that a combination of post-test scores and modifiability was diagnostically the most sensitive measure.

In addition to its diagnostic usefulness, studies have looked at the efficacy of dynamic assessment in predicting likely response to a proposed intervention. This has been investigated across a number of specific speech-language domains including reading comprehension (Elleman, et al., 2011; Fuchs, Compton, Fuchs, Bouton, & Caffrey, 2011; Swanson, 2011), narrative (Camilleri, 2005), morphology (Larsen & Nippold, 2007; Ram, Marinellie, Benigno, & McCarthy, 2013), pragmatic skills (Donaldson & Olswang, 2007;

Lin, 2010), triadic gaze (Olswang, Feuerstein, Pinder, & Dowden, 2013), use of two word utterances (Bain & Olswang, 1995) and phonological development (Glaspey & Stoel-Gammon, 2007). In contrast to the test-teach-retest approach outlined above, readiness for progress in an intervention targeting a specific language domain may be more appropriately determined by a graduated prompting methodology (Gutierrez-Clellen & Pena, 2001). This alternative DA format identifies the skills essential to complete a particular task (Laing & Kamhi, 2003), then standardises the teaching phase by providing the child with a hierarchy of predetermined prompts. Modifiability is measured more quantitatively through recording the effect of the cues embedded into the test task (Grigorenko & Sternberg, 1998), and transfer distance or 'generalisation' of the task content to novel items (Campione & Brown, 1987).

Bain and Olswang (1995) devised a protocol to examine the immediate potential for children performing at a one word stage of development to produce a range of two term utterances. Following the examiner's ostensive labelling and manipulation of a range of toys, six cues were presented in a hierarchy from general statement (least supportive) to a direct model plus elicitation question (most supportive). Over the nine week period of the study, the resulting weighted scores presented as a valid predictive tool in determining which children were ready to produce two-word relations, when provided with this intensive environmental stimulation. Children were aged between 2 ½ and 3 years; the youngest participants in any of the dynamic assessment procedures published to date. From a pragmatic perspective, Donaldson and Olswang (2007) contrasted typically developing and "more able" 5 to 7 year old children with Autism Spectrum Disorder (ASD) on their ability to produce requests for information, both on static assessment and during a DA procedure. Three levels of contextual support were provided: naturally occurring opportunities, making available highly preferred objects/activities and changing the physical setting. Static assessment was acknowledged as a starting point for determining the need for further assessment and the study concluded that it

was a combination of both approaches that was most successful in demonstrating which children actually had a deficit and in elucidating intervention possibilities.

Use of the graduated prompting approach is likely, most familiar to speech-language therapists in assessment of phonology. In an attempt to balance the inclusion of individualised examiner feedback with a standardised and replicable procedure, Glaspey and Stoel-Gammon (2007) developed the Scaffolding Scale of Stimulability (SSS). This constitutes a hierarchy of cues and environmental manipulations which can be used to facilitate production of phonemes, and document transfer to the most complex linguistic environment of connected speech. The authors highlighted how accurate analysis of task requirements and the sequential steps to achievement are crucial to the success of this approach.

Finally, from a clinical efficiency perspective, a graduated prompting approach has considerable potential as a screening tool. It is highly scripted, can be completed in a single, brief session (Patterson, Rodriguez, & Dale, 2013), and may achieve a crucial balance between responding subjectively to individual need and administering a replicable procedure.

In addition to carefully considering the type and amount of adult assistance that may be required to induce change, management decisions also need to address timing of intervention for language impaired children. With resourcing constraints now impacting more than ever on the majority of clinicians, determining the optimal time for offering the right intervention is paramount. The provision of support will differentially affect learning depending on when it occurs in the acquisition process of the new skill. A child will only be able to make use of adult assistance in bringing a target skill to the surface if the underlying, pre-requisite skills have already been mastered. Ultimately, the child must be in a position of “readiness”, and it is this readiness which can be revealed within an appropriately structured

DA procedure (Olswang & Bain, 1991). Related to this, DA sits on a continuum with interactive intervention and diagnostic teaching (Laing & Kamhi, 2003). This means it provides an opportunity to further analyse behaviours related to and/or preceding the target; the first step in creating an individualised treatment approach with an optimal outcome. Children who are less responsive to a DA scaffolding procedure may benefit more from an intervention program targeting earlier precursor behaviours (Olswang, et al., 2013). Comparing the relative efficacy of prompt type also provides qualitative information which can be easily transferred into an intervention program.

Further, when considering language differences, modifying a static test to match the content, style and use of the child's own language provides little functional information about that child's ability to meet the demands of the classroom (Pena, et al., 1992). This is unlike a DA specifically designed to incorporate the type of cueing typical of mainstream educational settings.

The standalone value of dynamic assessment as assessment for the sake of more efficient instruction has also been emphasised by Alony and Kozulin (2007), who explored the applicability of a receptive vocabulary DA procedure for children with Down Syndrome using an adapted version of the Peabody Picture Vocabulary Test (PPVT-R) (Dunn & Dunn, 1981). Administration was made dynamic by incorporating two types of mediation: focusing and verbal mediation. Focusing was used if the child was overly impulsive, or if they failed to look at the pictures rather than to address an incorrect response. Three increasingly supportive levels of verbal mediation followed: general explanation, elaborate explanation, and concrete demonstration. General explanations were definition based, and elaborate were more individually tailored to connect with the child's existing knowledge. The final demonstration prompt used a more hands-on approach to provide specific concrete examples; fulfilling criteria for Lidz's (1991) earlier documented mediation of transcendence. Results

indicated that even minimal ‘focusing’ was enough to support with more efficient word retrieval. The authors again discussed the need for a target skill to be emergent in a child’s system in order for them to be able to make use of adult assistance in bringing it to the surface. The ongoing benefits of this DA approach in predicting and altering developmental trends would need to be determined through a longitudinal study.

DA can also be more sensitive to smaller changes, and therefore clinically useful in establishing whether clients with more significant impairments are making gains within an intervention program. For example, using a case study design with a small sample of three boys aged 11 to 12 years, Hasson and Botting (2010) devised an MLE, based on the Sentence Assembly sub-test of the Clinical Evaluation of Language Fundamentals. (CELF-3) (Semel, 2000). In this instance, the dynamic measure was proposed as a more sensitive tool for monitoring progress. The functioning of the participants on some static subtests was so low in relation to their chronological age that even clear improvements in the raw scores was insufficient to raise the standard score. Again, this adds to the case for further development of DA procedures which have been designed to tap into and give credit for mastery of steps underlying emergence of the target.

Drawing together the above, there is a clear role for DA with LTs at a number of stages in the assessment and intervention process: 1) screening immediate level of risk for language impairment; 2) establishing readiness for a proposed intervention and efficacy of support; and, 3) recording medium to long term progress for clinical outcome measurement and diagnostic decision-making.

Dynamic Assessment and Single Word Learning

As previously discussed, one of the hallmarks of the 2 year old LT is a low vocabulary count. Although DA has not been widely used with 2 year olds, it has been used

in the area of word learning. As word learning ability is the most logical measure when considering potential for language improvement in a 2 year old LT, the following sections explore the use of DA principles within this language domain and consider factors which may be manipulated to create a clinical tool appropriate for use with 2 year old LTs.

Relevant to the current discussion, three levels of word learning analysis are described in the research literature. Fast mapping is the ability to determine the referent of a novel word on the basis of a single exposure (Carey & Bartlett, 1978). Extending beyond this is quick incidental learning (QUIL), an evaluation of the impact of the first few exposures in an incidental context. An advantage of QUIL over fast mapping is that it provides a measure of word learning rather than a static vocabulary count (Burton & Watkins, 2007), in itself suggestive of a natural fit with DA procedures. Slow mapping is the formation of a long term memory representation, such that it can withstand a delay (McGregor, Friedman, Reilly, & Newman, 2002). Fast mapping, QUIL and slow mapping can potentially form an overlapping continuum. During initial fast mapping, a rough hypothesis of the novel word-referent association is established. With QUIL, there is the potential to add to this with at least a partial mapping of the word's meaning, and during gradual and prolonged slow mapping repeated encounters work to strengthen the representation. With each subsequent encounter there is the potential for additional information about the referent to be encoded so that the novel word steadily increases in familiarity. Degree of semantic knowledge at any point in this process makes words more or less vulnerable to retrieval failure and is predictive of naming accuracy (McGregor, et al., 2002).

Within this gradual word acquisition context, Camilleri and Law (2007) developed a DA of single word receptive vocabulary and investigated whether it could provide quantitative measures of lexical ability which could not be predicted by static scores alone. The study embedded a graduated prompting approach within a test-teach-retest paradigm and

was administered to 3 ½ to 4 year old children referred to speech-language therapy due to concerns regarding oral language development. Following pre-test administration of the selected static measure, the British Picture Vocabulary Scale (BPVS) (Dunn, Dunn, Whetton, & Burley, 1997), the group was divided into normal range and low scoring. The DA included six vocabulary items; three nouns and three verbs from the BPVS that the children had been unable to identify correctly and was therefore individualised in this respect for each child. Three increasingly assistive levels of mediation, based on the child's ability to use relevance, discrepancy and mutual exclusivity, that is, assume that a known word would not have an overlapping reference (Markman & Wachtel, 1988), were provided in the teaching phase. Presenting each unknown item alongside two known distractors, and cueing as needed around how to make use of this surrounding context created opportunities for independent, implicit and explicit identification. Within the final static post-test, the three targeted words for each grammatical category were presented simultaneously and the child again asked to identify each. Again post-test scores combined with quantitative modifiability ratings provided the most accurate classification of low language learning ability. The authors also highlighted the paradox of requiring adequate comprehension of language in order to develop the metacognitive skills required for completion of a task within a mediation process. This is additionally complicated when the domain targeted is itself language. The child's current language level must be sufficient to cope with the language demands of the task itself. Thus, receptive language levels may be a distinguishing factor between children able to benefit from a MLE, and those who cannot. This reinforces the value first, of DAs which evaluate receptive abilities per se and second, of ensuring that any assistance provided is also pitched at an appropriate linguistic level. Clinically, this is also highly relevant as, as noted previously, it is typically the 2-year old LTs with combined receptive and expressive difficulties who have the most persistent problems.

This same, primarily context mediated approach to the assessment of receptive vocabulary was subsequently re-employed by Hasson, Camilleri, Jones, Smith and Dodd (2013) as a sub-test within their broader Dynamic Assessment of Pre-schoolers Proficiency in Learning English (DAPPLE), still under development. On this occasion, the task was administered with 3 to 5 year old bilingual children. A caseload group was composed of children who had been referred to speech-language therapy (SLT) and a control group of age matched peers. Results indicated a significant difference in amount of mediational assistance required between groups on the receptive task. Two measures were also taken of expressive vocabulary. The first doubled as a teaching task with unspecified support given until correct recall was demonstrated, and the second was simply a naming task administered after the first measure. Expressive differences emerged between the two groups only on the second measure. This suggested that the control group had benefited from the extra exposures provided in the first measure in facilitating later independent recall; however, this was still insufficient to effect change in the caseload group. Number of exposures therefore successfully differentiated the two groups both receptively and expressively but at different points in test administration. As with previous studies cited, the authors concluded that the DA results were most useful when considered in conjunction with static findings.

Continuing with the focus on comprehension, a more naturalistic approach combining the principles of DA and QUIL was explored by Camilleri and Botting (2013). Participants were aged 3 ½ to 4 years of age and again, had been referred to SLT. Their Dynamic Assessment of Word Learning (DAWL) was constructed with view to reducing any aforementioned metacognitive mediation component, and potential bias against those children already demonstrating receptive language difficulties. In order to explore more ecologically valid aspects of word learning, unknown words were introduced within conversational streams of speech which allowed for use of syntactic bootstrapping. A series

of semi-scripted interactions more closely approximating everyday word learning were each directed around a composite picture, within which the unknown word was represented at least twice. Receptively, the embedded prompt hierarchy began with an open elicitation question. This probed for a production which was used to assume understanding. This was followed by a three part description designed to facilitate use of the linguistic and pictorial context in word identification. Last, they received a further description which focused on highlighting semantic features. For children successful at the previous levels, this third level of prompting was used to assess generalisation, that is, the child's ability to identify another occurrence of the word. Increasingly salient joint referencing behaviours constituted the fourth and fifth cues. Again, the overall dynamic phase could be evaluated quantitatively in terms of amount of prompting needed. Participants were tracked for six months and strong correlations were found between the weighted DAWL measures and medium term progress across that time. This strengthens the case for a measure of receptive vocabulary in predicting both receptive and expressive gains.

Kapantzoglou, Restrepo and Thompson (2012) explored the impact of two further methodological adaptations. Firstly, they attempted to address the potential confounding variable of using real words which may still create a level of bias because of variation in amount of previous exposure. The study was completed in Spanish and targeted novel word learning skills in 4 and 5 year old bilingual children. Words were taught using a structured, scripted play activity, and both retention and recall were tested after 9, 18 and 27 exposures. Cue type and increments were selected to quantify rate of learning and to guard against too much support over-riding diagnostically useful word learning differences. Receptively, classification of disorder versus difference was most accurate after 9 exposures. In addition, both typically developing and language impaired children had difficulty recalling the novel words, even after the maximum 27 presentations, suggesting that the expressive naming task

was too difficult. As, however, decreasing this difficulty may compromise the sensitivity of the receptive word retention aspect, the study concluded that these two areas may need to be assessed separately in order for both to be sufficiently sensitive.

Building then on the production perspective, the efficacy of combining DA and QUIL procedures on a recall task was explored by Burton and Watkins (2007). A task was devised to teach four nonsense nouns corresponding to four novel items in a picture book; two referencing whole objects and two, parts of objects. Participants were African-American kindergarten students aged between 5 and 6 ½ years and described as being from high and low risk backgrounds. Participants received a total of eight exposures to each target word and one opportunity for use within a sentence cloze task. Recall was facilitated through the provision of four cues. These were again presented within a graduated, least to most assistive prompt hierarchy and consisted of: elicitation question, semantic cue, phonemic cue, and indirect model. The semantic cue involved linking the novel word back to the familiar context of the book by finding a picture and reiterating its role in the plot. Results indicated that the semantic cue did not provide enough assistance to access partial representation; however, the phonemic cue did seem to effectively tap into fragile representations. The indirect model provided enough assistance for all children to achieve correct productions suggesting that it was not sensitive enough to distinguish between children with fragile representations and those who had not mapped at all. The varying emotional value of target words in terms of story content was also proposed to have contributed to which words were easier to map. Although there was no significant difference between groups on the DA task, the overall low performance of both may confound this finding. The Peabody Picture Vocabulary Test (PPVT-III) (Dunn & Dunn, 1997) was again, employed as the static measure and did, as anticipated, differentiate between groups highlighting its bias against the children in the high risk group. Again, task difficulty and subsequent floor affects with

production of novel words suggest that more clinically useful outcomes might be achieved by the addition of a receptive measure. This might sit alongside the production measure providing more usable data for children unable to engage at this level as well as a more comprehensive profile of the child's overall linguistic potential.

Further, whilst a common theme throughout these word learning studies is the employment of some form of graduated prompting, it is notable that the retention and recall probes selected vary markedly. This is unsurprising given the multiple linguistic, cognitive and socio-pragmatic variables which are known to mediate word learning (McGregor, et al., 2002). It does, however, create a challenge in terms of creating a linear cue hierarchy with somewhat equal increments between levels of assistance or selecting the right prompts to reveal optimal learning potential.

Factors Influencing Two Year Old Word Learning

A number of factors which influence word learning and therefore have the potential to be of clinical use have been introduced. There are also others particularly pertinent to the 2-year old population being considered which warrant further exploration. Relative to other language domains, vocabulary acquisition is linked to prior experience leading to test scores which can reflect life experience more than language ability (Lidz, 2010). This presents a diagnostic challenge with static assessment; however, is advantageous in terms of vocabulary being a domain likely to respond well to manipulations of the learning environment and therefore produce observable modifiability within DA.

First to elaborate on the role of repetition, children with language impairment learn new open class words but more slowly and where difficulties are compensated for by repeated presentations. As touched upon in the previous section, input frequency is one variable which has been reliably used to differentiate between typically developing and

language impaired children. For example, Rice, Oetting, Marquis, Bode and Pae (1994) demonstrated that 3 exposures was sufficient to discriminate between the two populations at age 5 years; however, 10 exposures over-rode any minimal input constraint and was insufficiently challenging for revealing processing limitations. As cited earlier, for Kapantzoglou, et al. (2012), 9 exposures was sufficiently sensitive with the same age group.

With respect to establishing similar quantitative data and a baseline number of exposures for 2 year old children, a single exposure has been shown to be sufficient for retention of one novel word across a 10 minute period (Spiegel & Halberda, 2011) and a maximum of four novel words can be retained over a 5 minute delay with six exposures but only where the repetition is supported by ostensive labelling (Horst & Samuelson, 2008). This influence of gesture is discussed further below. Of note, a minimal delay of 5 minutes between teach and test phases is required for retention to be dependent on some level of representation in long term memory rather than repetition of a just prior selection. In considering the impact of multiple novel name-object mappings, six novel words were introduced in Spiegel and Halberda's (2011) task and eight in Horst and Samuelson's (2008). Processing demands clearly increased with higher number of words to be mapped.

By 2 years of age children appreciate both categorical/taxonomic and thematic relations among objects and use these features to form connections within the semantic component. Taxonomic coordinates tend to be visually similar because they share common functions and/or physical features and if given the choice, young children will interpret a novel noun using this type of superordinate relation. For example, when shown an object which had been assigned a novel label and asked to find another of that item, 25 to 36 month old children demonstrated a preference for taxonomically related over thematically related responses. That is, deer was more likely to be selected as another example of the new nonsense name for dog than bone. This effect was not evident where the object was not

labelled and children simply asked to find another *one* suggesting that it was the introduction of the novel label which served to draw the child's attention to the superordinate relation (Waxman & Kosowski, 1990).

The acquisition process for verbs, already identified as an area of particular difficulty for language-impaired children, may also differ from other word classes. Based on a significant decline in verbs as compared with nouns at retention testing, Rice et al. (1994) hypothesised that techniques to teach verbs may be more effective if they incorporated information about tense marking alongside repetition. Sensitivity to inflectional morphology can provide a kind of bootstrapping, supporting with determination of word class and thus the most likely referent. Children as young as 24 months old can attend to and contrast differing phonotactic structures (Jolly & Plunkett, 2008), whilst the diversity of syntactic frames in which a verb is heard is a positive predictor of subsequent flexible child use in children aged 18 to 29 months. Each frame provides an additional layer of semantic information contributing to the child's overall depth of word knowledge (Naigles & Hoff-Ginsberg, 1998). Within the same study, these multiple, more naturalistic models appeared to over-ride any potential benefit of inputting verbs in the more perceptually salient but syntactically sparse utterance final position.

Gesture and word acquisition are also highly interconnected. Word learning for children aged between 28 and 31 months can be significantly enhanced by the provision of any deictic gesture beginning with gaze alone (Booth, McGregor, & Rohlfing, 2008). Incremental increases in facilitative effect can be achieved with the addition of point, touch and manipulation respectively however with the greatest leap between gaze alone and gaze combined with pointing. Physical interaction with the object serves not only to more effectively draw the child's attention to it but also to enhance their understanding of the speaker's socio-pragmatic intentions. Capone and McGregor (2005) explored the effect of

iconic gesture with children aged 27 to 30 months. Comparing novel words taught with 9 verbal exposures alone to those taught in association with a gesture capturing either shape or function, a further hierarchy of facilitative effect emerged. The provision of any gesture was preferential to no gesture; however, fast mapping was particularly enhanced by those delineating the object's shape.

Finally, sub lexical factors such as high phonotactic probability, that is, the frequency with which phonemes and phoneme sequences occur can also be used to promote word-learning. Two year olds learned novel words composed of familiar sound sequences more rapidly than those composed of rare sound sequences (Storkel, 2001, 2003). Clinically, this is also reflected in established evidence based intervention programs such as Target Word (Earle & Lowry, 2011), which deliberately selects therapeutic targets based on the child's familiarity with the constituent phonemes.

Summary

In conclusion, there are a number of interacting variables which might be manipulated to support the word learning task. The potential for a DA of single word learning which utilises our knowledge of these alongside known predictors of language change to somewhat quantify rate of learning and inform subsequent practice is clear. Two year old LTs would also appear to be an ideal group with which to be implementing such an approach given the current uncertainty surrounding developmental pathways across the pre-school years. Reliable determination of prognosis has considerable implications in terms of service allocation. Finite SLT resources should be concentrated to where they are most needed and parents of children whose difficulties are likely to resolve should not be encouraged to worry unnecessarily. Conversely, given the risk of long term difficulties, whether subclinical or in the language impairment range, the limited vocabulary of a 2 year old LT should not be trivialised and

valuable intervention time wasted. The need for at least some degree of longitudinal follow-up is also emphasised throughout the literature. DA procedures are traditionally associated with potential for immediate change, whilst it can take many months for a LT's more long term language status to be revealed.

An exploratory dynamic assessment for single noun and verb word learning; established criterion in diagnosing language impairment has therefore been devised. Similar to the Kapantzoglou et al. (2012) study, the task takes the form of a structured, scripted play activity and controls for number of exposure. Factors known to support word learning are also embedded and the effects of a range of linguistic prompts are investigated in terms of their role in facilitating retention and recall.

Research questions:

1. Can typically developing 2 year olds be differentiated from late talkers on a dynamic assessment of single word learning across both spontaneous and assisted responses?

   Hypothesis 1: Typically developing 2 year olds will be clearly differentiated from late talkers across both spontaneous and assisted responses.
2. Are dynamic assessment scores for single word learning associated with change in language status across a medium term tracking period of 8 to 9 months?

   Hypothesis 2: PLS change scores from time 1 to time 3 will be positively associated with both time 1 and time 2 dynamic assessment scores.
3. Does the associative value of the dynamic assessment task increase over a 3 month period?

   Hypothesis 3: There will be a stronger relationship between PLS change scores and time 2 dynamic assessment scores when compared with PLS change scores and time 1 dynamic assessment scores.

Chapter 2: Methodology

This study was approved by the University of Canterbury Human Ethics Committee on 14 September 2012 (approval number: HEC 2012/119) (Appendix G) and informed consent was obtained from one parent of each child who participated.

Participants

Participants were recruited using a combination of convenience and snowball sampling. All of the typically developing children (TD) and five of the late talking (LT) children were recruited in response to a request for children at any language level which had been placed on a parent run, Plunket in Neighbourhood (PIN) group, Facebook site. This site was accessed by families in five suburbs in the Greater Wellington area. Plunket is a free support service for the development, health and wellbeing of children under 5 years of age in New Zealand and offers parenting information and support as well as developmental assessments (Plunket, n.d.). One LT responded to the program information sheet which had been distributed around day cares in two further suburbs, and six LTs were recruited when made aware of the study following broader developmental paediatric screening. Three were similarly made aware of the study when they attended early for a standard 2 ½ year old Plunket check. Both paediatric and Plunket contacts were initiated by parents due to specific concerns regarding language development. The remaining LTs were chain referrals arising from contact with the other study participants. Two TD children were also recruited to participate in preliminary trials approximately one month prior to the start of the study proper, and a further three were involved in developing the word learning task approximately five months prior.

Forty children, aged 24 to 29 months participated within the study. Participants were divided into two groups: typically developing (TD) and late talking (LT). Language status

and group allocation were determined on the basis of the MacArthur-Bates Communication Development Inventory: Words and Sentences (Fenson et al., 2007), which had been adapted to New Zealand English (Reese & Read, 2000). Late talking was defined as an expressive vocabulary score at or below the 10th percentile on the CDI vocabulary and typically developing as an expressive vocabulary score above the 10th percentile, relative to normative data collected in the USA (Fenson, et al., 2007). The Expressive Communication (EC) and Auditory Comprehension (AC) components of the PreSchool Language Scale, Fourth Edition (PLS) (Zimmerman, Steiner, & Pond, 2002) were also administered to children in each group in order to obtain a more comprehensive linguistic profile of each child and to track the language abilities of children in the late talking group.

Participant information is detailed in table 1 below.

Table 1. Time 1 Participant information for TD and LT Groups

	TD	**LT**		
	M (SD)	**M (SD)**	**t (df)**	**p**
Age	25.85 (1.84)	26 (1.86)	.26 (38)	.80
CDI Vocabulary	376.00 (161.73)	61.90 (58.64)	8.17 (23.91)***	<.001
PLS AC std	108.00 (14.75)	77.55 (9.12)	7.85 (31.68)***	<.001
PLS EC std	112.70 (16.14)	77.85 (6.80)	8.90 (25.54)***	<.001
PLS Total std	111.75 (15.66)	75.50 (7.74)	9.28 (27.77)***	<.001
PLS AC Raw	33.45 (5.16)	23.75 (2.86)	7.36 (29.70)***	<.001
PLS EC Raw	36.80 (4.87)	24.60 (2.58)	9.89 (28.90)***	<.001

Note. TD = typically developing; LT = late talking; CDI Vocabulary = CDI = MacArthur-Bates Communicative Inventory; PLS = Preschool Language Scale, Fourth Edition; AC = Auditory Comprehension; EC = Expressive Communication; std = standardised

***p<.001

A series of independent t-tests was computed to confirm that groups were matched for age, and that they were significantly differentiated on all standardised measures. There was a strong gender bias for boys within the LT group (*M* age = 25.85 months) which was

comprised of 16 boys and four girls, as compared with 12 girls and eight boys in the TD group (*M* age = 26 months). One LT withdrew from the study following phase 1 reducing this group to 19. Children within the LT group were described as having an isolated language delay of unknown aetiology. Based on parent report, there was no history of hearing impairment, prematurity or bilingualism.

Of note, three LTs scored higher on the Expressive Communication (EC) component of this measure than their CDI score might have suggested. These children, one aged 26 months and two aged 29 months obtained EC standard scores above – 1 SD, therefore within the normal range; a potential reflection of the differing skill sets targeted by the two assessments and the broader age aggregates of the PLS. Acknowledging that there is notable variation within the literature as to how late talking is defined, and that there are valid arguments for use of either assessment, a final decision was made to continue with experimental group allocation based primarily on CDI vocabulary score. This was based on three factors: greater sensitivity to age, frequency of use in the literature, and that a vocabulary measure seemed a more logical fit for a DA of single word vocabulary learning. An exception was made for one 24 month old participant who scored below – 1 SD on the EC component of the PLS however was at the 15th percentile on the CDI. This child was allocated to the LT group on the basis of his PLS EC score. All LT participants therefore scored at or below the 10th percentile on the CDI vocabulary, or at or below the 15th percentile but with a PLS EC score below – 1 SD.

Parents of children in the LT group were made aware of their child's language level and associated risk factors; however, direct input was not provided. Parents who requested further input were supported in making self-referrals to the Ministry of Education Early Intervention team and given the contact details of private speech-language therapists (SLTs) in the area as listed on the New Zealand Speech Therapy Association (NZSTA) website. No

children had received speech-language therapy prior to being recruited to the study. Two of the children proceeded to access private therapy and both received five sessions delivered at fortnightly intervals toward the end of their individual tracking period between times 2 and 3. Only one child accessed government funded therapy. This was for a 6 month period and delivered indirectly via a support worker at day care.

Measures

CDI: Checklists were completed by parents when the principal researcher attended for the time 1 assessment battery. All inventories were scored by the principal researcher.

PLS: The Auditory Comprehension and Expressive Communication strands were administered according to the published procedures. The participants' responses were recorded and scored by the principal researcher. Again, all children were monitored throughout for signs of fatigue and need for breaks.

Procedures

The study was conducted in three phases:

- Time 1: Administration of the CDI, PLS and the DA task with all participants
- Time 2: 3 months after Time 1: the DA task was re-administered to the late talking group only.
- Time 3: 8 to 9 months after Time 1: the PLS was re-administered to the late talking group only.

The Dynamic Assessment Task

The DA task was a scripted, semi-structured play activity delivered in four sections and devised to provide a total of 10 exposures to each novel word. Retention and recall testing

were embedded within the task and took place after each section. The task was introduced and administered according to the script outlined in Appendix C. Part I of the script provided 3 exposures to each novel noun and was followed by the first set of retention and recall testing. Probes used for testing are outlined in Appendix D and included the fourth exposure in the context of an indirect model. This is discussed further below.

Part I of the script was then repeated and extended with part II for each familiar and novel noun in the same sequence. All exposures for each word in each section were provided before progressing to the next word. The 10 exposures were therefore distributed in discrete sets of 3, 1 and 6. Parts I and II of the script also provided 1 and 3 opportunities respectively for verbal rehearsal of both novel and familiar items. Familiar words were included in the script to support with establishing an understanding of task demands and to reduce potential frustration (Kapantzoglou, et al., 2012). Previous studies have indicated that children are more likely to try and name target objects when they can successfully name other objects in the study (Gray, 2005). Administration of parts I and II combined was followed by the second set of retention and recall testing. A short break was then provided (if necessary) before administration of the verb condition. This followed the same format as described for the nouns. Part I of the script was administered for each verb followed by the third set of retention and recall testing. Parts I was then repeated and extended with part II and then the final set of retention and recall testing.

Novel objects throughout were identified through ostensive labelling and associated with both a primary (repeated x2 total) and a secondary (repeated x1 total) function and corresponding iconic gesture(s). Actions were associated with the object that they were acted on and again, an iconic gesture.

In the event of a child generalising an already known word to one of the novel items, some pragmatic direction was provided in order to assist the child with replacing their word with the new term, for example, "yes, it is a kind of horse but today we are giving it a special name. It's a *zutter*". If required, this response substituted a scripted exposure. It was not used as an additional exposure. This approach was based on Clark and Grossman's (1998) findings that linguistic directions referencing inclusion relations were most effective in facilitating acceptance of more than one term for the same referent in 2 to 2 ½ year old children.

The independent variables examined were number of exposures (two levels: 3 and 10) and prompt provision. The contrast between 3 and 10 exposures was selected on the basis of having reliably differentiated between typically developing and language impaired children on previous word learning tasks (Kapantzoglou, et al., 2012; Rice, et al., 1994). Prompt provision was further separated out into retention probing (three levels: elicitation question, gesture cue and semantic cue) and recall probing (three levels: elicitation question, phonemic cue and indirect model). Retention probes from least to most assistive were elicitation question, gesture cue plus elicitation question and function cue plus elicitation question. This hierarchy was selected as both gesture and verbal function cueing are known to contribute to the slow mapping process; however, initial provision of the gesture cue also helps the child to start encoding within initial fast mapping (Booth, et al., 2008). Gray (2005) also noted that some 4 to 6 year old children repeated a semantic cue as they searched successfully for a referent, further strengthening the case for exploring inclusion of the function cueing with a younger 2 year old group. Recall probes from least to most assistive were elicitation question, phonemic cue plus elicitation question and indirect model plus elicitation question. This differing approach to recall was based on Gray's (2005) conclusion that different cues may aid different aspects of word learning. Phonemic cueing, whilst essentially redundant within a comprehension probe, is more effective than semantic cueing in facilitating both whole and

partial word retrieval with 4 to 6 year old children (Burton & Watkins, 2007; Gray, 2005). Within Burton and Watkins' study, indirect modelling also provided by far the highest level of assistance as well as having been effectively used as a recall prompt with children as young as 30 to 36 months of age (Bain & Olswang, 1995). A semantic cue was not given for the verb retention condition as the object was being acted on in the photograph used for testing, and it would not be possible to separate out children for whom this cue was beneficial from those who simply scanned the photos and pointed on the basis of the familiar noun. A 1 point option was therefore not available for the verb retention condition.

The dependent variables were retention and recall modifiability scores (three levels: following 3 exposures, following 10 exposures, and following 3 and 10 exposures combined). Separate recall and retention scores were also computed for nouns and verbs by separating out the 3 and 10 exposures combined scores. A total retention or recall score was therefore equivalent to either, the 3 and 10 exposures combined score, or to the noun score combined with the verb score. Unfortunately, in hindsight, the noun and verb conditions were not counterbalanced and the noun condition was always administered first.

Initial administration of the DA occurred on the same day as the CDI and PLS. At time 1, the PLS was administered prior to the DA task. This also served to ensure that the child was familiar with the concept of pointing to pictures. Breaks were provided as required; however, parents were asked to first help their child refocus before taking a break. Complete administration of the DA task took approximately 40 to 45 minutes. Task structure resulted in approximately 5 minute intervals between focused exposure to the novel words and testing, a similar time lag to that investigated by Horst and Samuelson (2008). This was not, however, strictly controlled and testing simply followed directly on from completion of the play component of the task. There was no change to administration of the DA task at time 2.

Thirty-five participants were seen at each phase within the familiar environs of their own home. Three of the late talking, and two of the typically developing children's families elected instead to use a clinic room at a local Child Development Service. At least one parent remained present throughout for all assessments.

Materials

Props for the DA task included two soft toys; a goat and a giraffe who acted as the protagonists within the play. The familiar nouns were a soft ball, a toy apple with a removable slice attached by Velcro and a set of four pretend flowers which could be removed from their 'flowerbed' base to facilitate a picking action. A wooden knife was also presented (but not named) alongside the apple to facilitate a cutting action. The novel nouns, *tib*, *miggle* and *zutter* were represented by a toy whistle which made a noise when blown, a wooden melon segment with removable fabric peel and a ride on playground toy. Objects for the familiar verbs pushing, jumping and sleeping were a small two wheeled trolley, an equestrian style jump and a felt (picnic) blanket. The novel verbs *heeping*, *wisping* and *nooding* were associated with 4 small wooden building blocks, a toy slide and a toy puppy, small enough to be held in the paws of the goat. All fitted with the context of a fun day out at the park; an event likely to be highly familiar to the two year old sample. As discussed in the following sections, there were 12 laminated photographs and a simple one page scoring form used for testing.

Scoring

A 0 to 3 ordinal scale was used to score both retention and recall. Correct responses to the elicitation questions alone were deemed spontaneous unassisted responses and allocated 3 points. Gesture and phonemic cueing corresponded with 2 points and semantic cueing and indirect modelling with 1 point. 0 scores were given where the child did not respond or

continued to respond incorrectly following the maximum amount of assistance. Each word was scored after 3 exposures and again after 10 exposures resulting in a potential word score of 12 and task score of 72. Examples of completed scoring forms are displayed in Appendix F. This type of hierarchical scoring system is effectively employed throughout the DA word learning literature (Burton & Watkins, 2007; Patterson, Rodríguez, & Dale, 2013) with the only major adaptation being a weighted version (Bain & Olswang, 1995; Camilleri & Botting, 2013). Weighted scores also gave credit for the more supportive cue types, assuming, for example, that a child who was successful at the elicitation question level would easily be able to respond to both a phonemic cue and an indirect model. With regard to the current study, this approach was felt to be undesirable for three reasons; 1) no studies could be found which compared the relative efficacy of these cue types with a 2 year old population therefore whilst there is some evidence for the selected hierarchy with a 3 to 5 year old population, it cannot reliably be said that a 2 year old who is able to demonstrate word retention in response to a gesture cue would also be able to make use of an associated verbal function (semantic) cue; 2) it cannot be assumed that ability to engage in an assistive interaction across multiple turns, a more social correlate of word learning can be inferred from ability to spontaneously map a word to a photo referent, a more straightforward linguistic correlate; and, 3) with many LT children in particular demonstrating floor effects and struggling to utilise any cue type at time 1, creating what would ultimately be large increments between participant scores was felt unnecessary.

Dynamic Assessment Administration

Retention testing for each section was administered before recall testing. For the retention testing components, participants were asked to select the correct referent from photo representations of the objects (Appendix E). Preliminary trials indicated that the 2 year old participants demonstrated a high level of interest in the task; however, demonstrated

difficulty in transitioning from play to testing when the same toys were used throughout. Use of photographs therefore reduced the tendency for any child to simply retrieve their favourite toy from the selection. Further, whilst the photographs depicted the exact toys used, testing with a new representation of each referent extended the children further and could be regarded as a more robust test (Horst, Parsons, & Bryan, 2011). An array of four laminated photographs presented in a four by four quadrant was used to minimize chance responses. Each photograph measured 15.92 x 11.94cm. Distractors were the other two novel words within the same word class, and the real word most semantically and/or temporally associated with the novel word within the play. For example, distractor pictures for the novel noun, *tib* were *miggle*, *zutter* and *ball*. All previous images were removed, the new familiar distractor item substituted, and a new random arrangement presented for each test item. The child's attention was first focused using statements to encourage looking at the photographs and systematic scanning of each before answering. Adapted from the work of Alony and Kozulin (2007) these included use of the child's name, iconic hand signs to sit, look and listen and the verbal statements, "now, let's see if we can remember those new words. Look at all the pictures, listen carefully and then choose just one". Once optimal attention had been established, the child was asked to point to the target item, for example, "Ok, now show me the *tib*… where is it?" If the child was not forthcoming with a clear response, then the examiner would prompt with a guiding statement such as "can you put your hand on it?" This specific direction was provided, as for many of the LT participants in particular, pointing to pictures had been identified within the PLS as an emerging skill. The name of the item was not repeated. Whilst this presented a challenge in terms of temporal distance from the target word growing as level of adult assistance increased, it was felt more important not to unintentionally confound findings by increasing number of exposures further in an uncontrolled for way. If the child did not respond or responded incorrectly, then the examiner

proceeded with provision of the gesture cue. Each subsequent probe was repeated a maximum of two times. This is in line with repetition protocols on the PLS (Zimmerman, et al., 2002). In order to maintain the child's continuing engagement in the task, generic praise such as "great pointing" or "I like the way you chose a picture" was given for any purposeful response. A child was considered to have responded randomly if they identified more than one item or if it was clear that they had not studied any of the pictures before responding.

For recall testing, credit was given to responses most closely aligning with the child's phonological profile as determined by parent report and analysis of speech samples obtained during administration of the PLS. For example, a typically developing child known to be using the phoneme /g/ competently at word level was not credited for use of "middle" (for "miggle") however a late talker, who had not yet acquired this phoneme, was credited if use of /d/ reflected a typical response within this context. During recall testing, if there was any doubt as to whether a recall attempt was the child's optimum production, the next level of prompting was provided. If the production remained the same, then the child received the higher score for the lower level of assistance. If the production was modified so that it more closely approximated the target, then the lower score was given with the rationale that the first production could be attributed to only partial recall of the word rather than to any phonological simplification. For a number of children within the TD group in particular, the phonemic prompt appeared to activate recall of a real word which differed from the target by just one phoneme, for example, "noosing" rather than *nooding* making this differential additionally important. Any uninterpretable utterances which could not be transcribed by either the examiner or the second scorer or those which did not incorporate any consonant sounds were scored as a no response. The phonology of productions obtained after 3 exposures was also compared with those obtained after 10 exposures to ensure consistency of scoring within any given child's profile. If after 3 exposures the child was able to correctly

recall the word either spontaneously or following the phonemic prompt, the indirect model was provided in validation, for example, “yes, *tibs* are for blowing”. This ensured that all children received the same 10 adult delivered exposures prior to the next set of testing. Although the novel word was also heard x1 in retention testing, it was only at this cue level that it was explicitly associated with the referent. Responses throughout were recorded and transcribed as appropriate by the primary researcher using the task scoring form. A number of task administrations were video recorded for later inter-rater agreement checking.

Inter-rater Agreement

Two complete video recorded samples were rescored by the principal researcher with a second scorer present. The second scorer was a qualified and practicing speech-language therapist to whom the task administration and scoring system had been explained. This step was designed as a training exercise for the second scorer. A further 4 complete video recorded administrations of the DA task were then selected randomly and watched independently by the second scorer. Fidelity of administration was checked by the second scorer based on adherence to the script and repetition limits in testing. The task was delivered with 92% accuracy. Point-by-point agreement was calculated separately for total retention and total recall and was 100% and 90% respectively for this 10% sample.

Chapter 3: Results

Question 1: Group Differentiation on DA task (Assisted)

A series of independent t-tests was used to compare performance across groups on dynamic assessment (DA) task score at time 1 (see table 2). Between group differences were significant for all variables. Results therefore support the hypothesis that 2 year old children who are typically developing (TD) can be differentiated from children who are late talkers (LT), both on standardised assessment and on a DA of single word learning.

Table 2. TD and LT Time 1 Comparisons on all DA task variables

	TD	LT			
	M (SD)	**M (SD)**	**t(df)**	**P**	**D**
DA Total	24.35 (10.76)	7.80 (6.54)	7.86 (31.68)***	<.001	1.86
Total Retention	17.95 (6.92)	7.50 (6.53)	4.91 (37.87)***	<.001	1.60
Retention x3	9.35 (3.63)	3.40 (3.63)	5.18 (38.00)***	<.001	1.64
Retention x10	8.60 (4.22)	4.10 (4.49)	3.26 (37.85)**	.002	1.03
Noun Retention	11.40 (4.04)	4.40 (3.66)	5.74 (37.63)***	<.001	1.82
Verb Retention	6.55 (4.02)	3.00 (3.43)	2.93 (36.56)**	.006	.93
Total Recall	6.40 (6.39)	.30 (.57)	4.25 (19.30)***	<.001	1.34
Recall x3	2.50 (3.07)	.25 (.55)	3.23 (20.22)**	.004	1.02
Recall x10	3.90 (4.12)	.05 (.22)	4.18 (19.11)**	.001	1.32
Noun Recall	2.65 (3.66)	.25 (.55)	2.90 (19.86)**	.009	.92
Verb Recall	3.75 (3.34)	.05 (.22)	4.95 (19.17)***	<.001	1.56

Note. DA = dynamic assessment; TD = typically developing; LT = late talking; x3 = following 3 exposures; x10 = following 10 exposures

**p<.01, ***p<.001

Group Differentiation on DA task (Unassisted)

It was also proposed that TD and LT children might differ on number of spontaneous unassisted responses. As each LT and TD received an equal number of opportunities for spontaneous retention and for spontaneous recall, results were again compared using an

independent samples t-test. Results confirmed a significant difference between groups for spontaneous retention and for spontaneous recall (see table 3).

Table 3. TD and LT Time 1 Spontaneous Retention and Spontaneous Recall Comparisons

	TD	**LT**			
	M (SD)	**M (SD)**	**t(df)**	**P**	**D**
Retention	4.45 (2.24)	1.50 (1.76)	4.64 (36.04)***	<.001	1.46
Recall	.65 (1.14)	.00 (.00)	2.56 (19.00)***	<.001	.81

Note. TD = typically developing; LT = late talking

**p<.01, ***p<.001

Question 2: DA scores and Association with PLS Change Scores

The relationship between the DA scores at times 1 and 2 and change in PLS scores over the 8 to 9 month tracking period of the study was explored through computing a series of bivariate correlations. Change was defined as a change in PLS raw score rather than a change in standardised score. This was due to sensitivity to age and the variation in baseline PLS scores at time 1.

Bivariate correlations were computed for both auditory comprehension (AC) and expressive communication (EC) change scores and each DA variable at times 1 and 2. At time 1, no significant correlations were found between either, AC or EC change scores and any DA task variables.

At time 2, significant correlations were found between AC change scores and five of the 11 DA task variables: retention following 10 exposures, noun retention, DA total, total retention and total recall. Significant correlations were also found between EC change scores and three of the DA task variables: noun retention, retention following 10 exposures and total recall. Two variables (retention following 10 exposures and total recall) were strongly

correlated with both AC and EC change scores. Retention following 10 exposures was also the DA task variable most closely associated with AC change ($r = .64$, $p = .003$, $n = 19$). Noun recall was the DA task variable most closely associated with EC change ($r = .61$, $p = .006$, $n = 19$). Correlation values are presented in table 4 below.

Table 4. Correlations between Time 2 DA variables and PLS Time 1 to Time 3 change scores

	Total Retention	**Retention x10**	**Noun Retention**	**Total Recall**	**Recall x10**	**Noun Recall**	**AC Change**	**EC Change**
DA Total	.97**	.94**	.87**	.63**	.67**	.39	.55*	.45
Total Retention		.94**	.92**	.44	.55*	.21	.55*	.36
Retention x 10			.85**	.50*	.64**	.34	.64**	.49*
Noun Retention				.28	.53*	.06	.57*	.27
Total Recall					.76**	.82**	.46*	.48*
Recall x10					.	.55*	.27	.45
Noun Recall							.36	.61**
AC Change								.67**

Note. AC = auditory comprehension; EC = expressive communication; x10 = following 10 exposures

*p<.05, **p<.01

Question 3: Increase in Associative Value of the DA task

As noted in the results reported in the previous section, the DA task for the LT group was strongly correlated with change in language performance at the end of the study; however, no correlation existed between DA scores at time 1 and language change. This suggests that the time 2 DA scores were a more useful measure of language change. Additional analyses were carried out to explore this increase in associative value of the DA task.

The first comparison that was conducted was an analysis of the difference between DA scores at time 1 and time 2 for the LT group. Differences in performance of the LT group on the DA task between time 1 and time 2 was computed using a paired t- test. Results indicated a significant difference in performance across 10 of the 11 DA variables. There was no significant difference between times 1 and 2 on recall following 3 exposures; however, this should be interpreted in the context of the extremely small mean scores, .26 and 1.11 out of 24. Table 5 shows the differences between time 1 and time 2

Table 5. Time 1 versus Time 2 Performance on the DA Task for the LT Group

	Time 1	**Time 2**			
	M (SD)	**M (SD)**	**t(df)**	**P**	**D**
DA Total	8.00 (6.66)	16.32 (10.12)	3.58 (18)**	.002	.97
Total Retention	7.68 (6.65)	14.21 (8.73)	3.34 (18)**	.004	.84
Retention x3	3.47 (3.72)	6.95 (4.82)	2.96 (18)**	.008	.81
Retention x10	4.21 (4.60)	7.26 (4.45)	2.53 (18)*	.021	.67
Noun Retention	4.42 (3.76)	8.05 (4.59)	3.59 (18)**	.002	.87
Verb Retention	3.16 (3.45)	6.16 (4.85)	2.63 (18)*	.017	.71
Total Recall	.32 (.58)	2.11 (2.56)	2.96 (18)**	.008	.96
Recall x3	.26 (.56)	1.11 (1.73)	2.04 (18)	.057	.66
Recall x10	.05 (.23)	1.00 (1.49)	2.67 (18)*	.016	.89
Noun Recall	.26 (.56)	1.32 (2.03)	2.34 (18)*	.031	.71
Verb Recall	.05 (.23)	.79 (1.48)	2.11 (18)*	.049	.70

Note. DA = dynamic assessment; x3 = following 3 exposures; x 10 = following 10 exposures

*p<.05, **p<.01

Bivariate correlations were also computed between each DA variable at time 1 and the same variable at time 2. A small but significant correlation was found for noun retention only ($r = .46$, $p = .05$, $n = 19$). Thus, whilst results indicate significant change, there was no relationship between the performance of the LT children on the DA at time 1, and their performance at time 2.

The small but significant difference between time 1 and time 2 DA scores combined with no significant relationship between the time 1 DA scores and time 1 to 3 PLS change scores would suggest that the task at time 1 was simply pitched at a too high a level for most LTs to fully utilise the assistance provided, and in turn, to produce meaningful associative data.

In the second analysis, a comparison was made between time 2 DA scores for the LT group and the time 1 DA scores for the TD group (which were collected at the beginning of the study). Given that there was a relationship between DA Time 2 and language outcome, it was anticipated that children who show higher DA scores (i.e. are likely to be "language different" rather than disordered) should, in fact, look similar to TD children by time 2. In order to investigate this, time 2 DA total z-scores were created for each of the children in the LT group using the mean and standard deviation of the TD group's time 1 DA scores. Nine of the 11 children in the LT group, whose PLS total scores placed them above – 1 SD, obtained z-scores in the - .78 to .8 range (see table 6). That is, those LT children were within one standard deviation of the TD group mean. These children in the LT group had word learning abilities, as measured by the DA that more closely approximated their TD peers leading to catch up gains sufficient for movement into the normal range over the following five to six months.

Table 6. Late Talker DA z-scores at time 2 depicted alongside their PLS standardised total scores at time 3

Child	Time 2 DA z-score	Time 3 PLS total std score
1	-2.17	75
2	-0.13	66
3	1.80	79
4	-1.52	64
5	0.22*	95
6	-1.05	83
7	-0.31*	105
8	-1.80*	90
9	-1.52	67
10	-1.61*	95
11	-0.59*	85
12	0.80*	109
13	-0.03*	86
14	-1.71	64
15	-1.15	75
16	0.71*	100
17	-0.03*	112
18	0.71*	93
19	-0.78*	109

Note. DA = dynamic assessment; std = standardised; *PLS Total standardised score at, or above – 1 SD

Following the findings that some children in the LT group were now more similar to the TD peers, the LT group was then sub-divided further into two new groups: resolved Late Talkers, known as Late Bloomers (LB); and, those that continued to present as language delayed (LD).

A further series of independent t-tests was therefore conducted between these newly formed LB and LD groups. Children scoring – 1 SD or above on their PLS total standardised score were allocated to the LB condition and all others to the LD condition. Three children who scored above – 1 SD on the PLS at time 1 (in spite of lower CDI vocabulary scores) were excluded as this investigation was around exploring movement into the normal range and including children already at this point biased the data. This resulted in 8 participants within each group. The difference between the two newly formed groups (LB) and (LD)

groups were compared across total DA task score at time 2. Also, given its status as the variable achieving the strongest correlation with any PLS change score at time 2, retention x10 exposures, was the DA variable that was chosen for comparison.

Table 7 – LB and LD between group comparisons on Time 2 DA Total and Retention following 10 exposures with groups based on PLS standardised total scores

	LB	**LT**			
	M (SD)	**M (SD)**	**t(df)**	**P**	**D**
PLS Total std T3	94.13 (7.79)	71.63 (7.33)	5.95 (13.95)***	<.001	2.97
DA Total T2	21.63 (11.05)	9.50 (6.66)	2.66 (11.49)*	.021	1.33
Retention x10 T2	9.50 (4.11)	4.13 (3.76)	2.73 (13.89)*	.016	1.36

Note. TD = typically developing; LT = late talking; PLS Total std = standardised total language score on the Preschool Language Scale, Fourth Edition; x10 = following 10 exposures; T2 = Time 2; T3 = Time 3

**p<.01, ***p<.001

Between group differences were significant for all variables. The LB and LD children were significantly differentiated on PLS standardised total scores at time 3 and, on DA total and Retention following 10 exposures at time 2.

These same analyses were conducted for LD and LB groups for the PLS EC score as well. EC was chosen as groups were originally determined based on an expressive measure. Again, three children were excluded from the analysis due to EC scores already above – 1 SD at Time 1. Two of these children had also been excluded from the previous analysis. Group sizes were 9 for LB and 7 for LD. Two children were re-categorised from LD to LB, one from LB to LD and one from LB to being excluded. The previously excluded child was placed in the LB group.

Table 8 – LB and LD between group comparisons on Time 2 DA Total and Retention following 10 exposures with groups based on PLS standardised EC scores

	LB	**LT**			
	M (SD)	**M (SD)**	**t(df)**	**P**	**D**
PLS EC T3	96.00 (7.94)	72.86 (5.58)	6.84(13.91)***	<.001	3.37
DA Total T2	17.33 (10.76)	11.71 (8.69)	1.16(13.96)	.267	.57
Retention x10 T2	7.67 (3.97)	4.86 (4.30)	1.34(12.48)	.204	.68

Note. TD = typically developing; LT = late talking; PLS EC std = standardised expressive communication score on the Preschool Language Scale, Fourth Edition; x10 = following 10 exposures; T2 = Time 2; T3 = Time 3

**p<.01, ***p<.001

Results again indicated a significant difference on PLS standardised EC score between groups. In contrast to the PLS total score, however, DA Total and Retention x10 Exposure scores were not significant.

Chapter 4: Discussion

Previous literature has shown dynamic assessment (DA) to be a useful tool in clinical decision-making. In this study DA was used to examine the word learning abilities of 2 year old typically developing children (TD) and 2 year old children who were Late Talkers (LT). Further, this study examined the relationship between DA and language change over time. Finally, changes in DA scores at two time points were used to examine when word learning ability might start to approximate that of the TD group.

Children who were TD and LT were administered a four part DA task as well as a standardised language measure at the beginning of the study. The LT children then completed the DA task 3 months later and a standardised language measure 8 or 9 months later at the end of the study. Based on previous literature, three hypotheses were formed: 1) Children in the LT group would differ significantly from the TD group on DA tasks; 2) As DA is meant to reflect "modifiability" to change, DA scores for children in the LT group (at both times 1 and 2) would be correlated with language change (as measured by PLS raw scores); and, 3) DA score be more indicative of language difference versus language disorder in children in the LT group when the children can more effectively complete the DA tasks.

Findings supported the first hypothesis. The DA of single word learning successfully differentiated the TD 2 year olds from the LT 2 year olds across both spontaneous and assisted responses. There was only partial support for the second hypothesis. Contrary to expectation, change in language performance between the start and end of the study was not associated with the DA scores at time 1; however, language change was positively associated on a number of DA variables at time 2. Findings therefore also supported the third hypothesis. Children in the LT group were better able to perform the DA tasks at time 2 with fewer floor affects. That meant that at time 2, the children in the LT group who could demonstrate better learning on the DA task, showed more change in language outcome;

improving more than the children in the LT group who scored poorly on the DA scores at time 2. In fact, the children with higher DA scores at time 2 had scores that more closely resembled their TD peers. On the other hand, the children who continued to perform poorly on the DA tasks at time 2 continued to perform poorly. This has implications for the clinical usefulness of the DA task; and, in particular, reflects the importance of the "level" at which DA tasks are set. The findings relative to each hypothesis are discussed in more detail below.

Dynamic Assessment in Differentiating TD and LT Two year olds

The DA task used in this study successfully differentiated children in the TD group from children in the LT group. This is consistent with a number of other DA studies (Hasson, et al., 2013; Law & Camilleri, 2007; Patterson, Rodriguez, et al., 2013) which concluded that number of cues needed to learn discriminated typically developing and language impaired children for vocabulary.

Closer examination of the results also indicated some subtle differences in the ways the two groups were discriminated. Two variables in particular were identified as being relevant to performance on the DA tasks: 1) "retention versus recall" probes and 2) "number of exposures" to the words. With regard to retention versus recall, in general children in both groups performed better on retention. This is an issue for DA task design because although it has been shown to be beneficial to include both retention and recall measures, when retention tasks are at the appropriate level for measuring "modifiability", then the related recall tasks are often too difficult to get observable change, even with prompts. In this study, the children in the LT group were barely able to perform on the recall component and therefore its usefulness for measuring potential for change was limited. This is, in part, reflected in the differing effect sizes between the retention and the recall tasks; and, in part, in the lack of correlation between DA at time 1 and language change for the LT group.

This dichotomy between retention and recall in task design is also noted in studies by Burton and Watkins (2007), Kapantzoglou, et al. (2012) and Camilleri and Botting (2013). A task which is sensitive to receptive differences between groups can result in floor effects for production. Conversely, making the task easier, to facilitate production, can result in a ceiling effect for retention. Ultimately, differing minimum thresholds at which retention and recall can be demonstrated in both LT and TD children create a challenge in terms of constructing a task which can test at an optimal level for both.

Number of exposures was another task variable that was identified as important, both in the literature (Kapantzoglou, et al., 2012) and, in the current study. In this study, children in the LT and TD groups were clearly differentiated at both the 3 and 10 exposure levels. For retention however, the effect size was smaller at the 10 exposure level and for recall, it was the opposite; effect size increased. For the LT group, increasing exposure number potentially worked to reduce the gap between them and their TD peers; however, the TD children were better placed to make use of increasing exposure in assisting recall. The LT group's extremely low recall scores, at both levels, provided limited opportunity to observe "modifiability". Overall, however, these were very small effect size differences and between group differences remained significant at both levels. Thus, whilst Rice, et al. (1994) demonstrated that increasing number of exposures from 3 to 9 could over-ride receptive word learning difficulties in 5 year old SLI children, this was not the case for the 2 year old LTs profiled. The most likely explanation for this was their younger age and, the previously discussed, competing interaction between number of exposures and number of words targeted.

Looking in more detail at the retention following 10 exposures variable within the LT group, it might be assumed that the 10 exposures strengthened or created representations which had been only partially mapped or not mapped at all at the 3 exposure level. Scanning

the data however, this was not necessarily the case. At both times 1 and 2, children in the LT group successfully demonstrated spontaneous retention of target words after 3 exposures. Many of these were then 'lost' at the 10 exposure level or retention facilitated only through prompting. For some, this response pattern might be attributed to attention difficulties, with the prompt serving to refocus as much as to tap into any fragile spoken word-referent mappings (Alony & Kozulin, 2007); however, for others, these losses were counterbalanced by the retention of new words suggesting that even with 10 exposures, the competing stimuli of 6 novel words in working memory resulted in the loss of those where there had not been adequate opportunity for consolidation.

Dynamic Assessment and Language Change

No significant correlations were found between any of the time 1 DA variables and PLS standardised or change scores at time 3. This would indicate that contrary to the initial study hypothesis, no relationship existed between the DA at time 1 and change in PLS score over an 8 to 9 month period.

Results overall suggest that the task, at time 1, was simply too difficult for the LT group. Two factors would support this. Firstly, the LT children performed significantly better at time 2, than at time 1. Secondly, there was only a very weak relationship, based on a single variable (noun recall), between the LT time 1 and time 2 DA scores, dispelling the idea of related improvements. Together, these factors would contradict the hypothesis that a weak relationship did exist between times 1 and 3; however, not statistically evident within the current small sample.

Of much greater interest are the 7 moderately significant correlations found between the LT time 2 DA task scores and time 3 PLS change scores. These are further explored in terms of clinical value and implications for task design. The strongest correlation identified

was between AC change and retention following 10 exposures (r = .64, p = .003, n = 19). This can be perceived as a positive given that the task, whilst incorporating both retention and recall measures, was particularly interested in capturing retention ability, the more reliable predictor of later expressive language performance, when considering potential for change between 2 and 3 years of age (Thal, Tobias, & Morrison, 1991). Unsurprisingly, the floor effects for verbs resulted in a clear noun advantage both expressively and receptively. Noun retention was positively correlated with AC change scores (r =.57, p = .012, n = 19) and noun recall, with EC change scores (r = .61, p = .006, n = 19). This is interesting in light of the noun-category bias (Waxman & Kosowski, 1990) discussed earlier and the strategies that young children might be able to draw upon in processing novel nouns but not items from other word classes. Whilst significant correlations were also obtained for total retention and total recall, it is suggested that, within the above context, these were achieved on the strength of the noun scores.

DA total scores at time 2 also differentiated children who continued to present as language delayed and those who resolved into Late Bloomers (LB). Further, the LB group's DA total scores at time 2 more closely approximated those of the TD children. This is consistent with the findings of Camilleri and Law (2007), who noted that some 'delayed' children could achieve a similar range of scores on DA to TD children, when static scores were significantly lower.

Given the marked contrast between DA results at times 1 and 2, a series of correlations was also computed to determine any relationship between age, as an independent variable, and PLS (raw) actual or change scores. Results indicated that age at time 1 was significantly correlated with AC change score (r = .554, p = .014, n = 20) but not with AC actual score at either time 1 or time 3. Thus, there is some evidence to suggest that children at the upper end of the 24 to 29 month old age bracket are more likely to make greater gains in

AC over an 8 to 9 month period simply as a function of 'being older' regardless of raw baseline. No significant correlations were obtained between age and EC actual or change scores at either times 1 or 3. With regard to the DA, age was also significantly correlated with a number of time 1 recall components. These were total recall, recall following 3 exposures and noun recall. Given that none of these time 1 recall components were significantly correlated with either AC or EC change or PLS total standardised score at time 3, there is some tentative support for the hypothesis that age combined with language level, rather than language level alone was a factor at time 1 in determining which children could engage with the task and build capacity at a production level and which could not. This may be attributable to maturation in other developmental domains which might previously have been acting as a barrier to task participation.

Limitations of the Current Study

Building on comments made in the previous section, many of the LT group opted out of the task within the verb component. Further, teaching and testing of verbs was not counterbalanced with nouns therefore it is not possible to determine whether poorer performance for verbs was attributable to task difficulty, fatigue, a true noun advantage or attentional difficulties. In the DAWL (Camilleri & Botting, 2013), children received a proportional single word learning score based on some language delayed participants being asked to complete less test items. Given, however, the co-occurrence of attention and language difficulties (Dale, Price, Bishop, & Plomin, 2003), this approach may result in an over-estimate of real world potential for change. Conversely, for some children teaching and testing for nouns and then for verbs, may have provided an unintended verb advantage. The total DA task score reflected overall level of success in completing the task and not any improvement *within* it over trials, a profile highlighted by Patterson, Rodriguez and Dale (2013) as demonstrating a higher level of modifiability than a child who consistently required

the same level of support. A child requiring less prompting on items toward the end of the task, demonstrates greater generalisation or transfer and has ultimately learned to learn. Initial administration of the noun component might mean that more modifiable children had greater awareness of task structure and type of cue provided in the testing phase. As a result, they were likely to be more attuned to this same information, and its prospective value in assisting retention and recall in the subsequent teach phase. This was particularly evident within the recall phase for the TD group. 10 of the 20 TD participants obtained a higher score for verb recall than for noun recall resulting in a higher overall group mean for the verbs (3.75 as compared with 2.65). This visible within task modifiability for recall, as compared with retention, may be attributable to a greater level of difficulty and therefore scope for making gains when provided with the right kind of assistance. Importantly, the TD group were at a developmental stage in their word-learning where they both needed and, were able to make use of this support, explaining why the same pattern was not evident within the LT group. Any future task would therefore need to look at counterbalancing items so that the relative effects of cue type could be more clearly compared across word class.

Also related to task structure, whilst there were 3 distractor items for each child, the task did not fully address the types of errors made and the influence of the competitor items. Whilst many children simply did not respond, or incorrectly selected another novel item, others seemed to have a real word competitor bias, consistently selecting the most semantically or temporally associated real word item. For example, upon hearing the novel word *miggle*, the child might have responded by identifying the apple image. Thus, whilst familiar words were included to encourage task participation and a feeling of success they may for some children have served instead to inhibit the formation of new word-referent associations. Children who formed partial semantic word representations as opposed to forming no representations at all may have been particularly vulnerable in this area as their

representation was sufficient to guide a related response but missing differential levels of detail. Although an alternative scenario is suggested by Schafer (2005), that presenting an object from the same category may facilitate word learning by enabling comparison processes and use of shared commonalities in encoding a richer reference, it is unlikely that the novel and familiar items with the exception of apple and *miggle* fulfilled this criteria. The relative emotional value of each word, novel and familiar was also not factored in. For example, the musical instrument *tib* appeared to be particularly appealing for the majority of children and therefore in addition to its' prime position in order of presentation, may have contributed to the high success rate for this item. Relative to all novel items, *tib* obtained the highest composite word score both for TDs and for LTs at both times 1 and 2. Interestingly, the other 2 nouns, *miggle* and *zutter* obtained almost identical distribution patterns (table 9). A similar bias toward the most 'fun' non word was also highlighted by Burton and Watkins (2007).

Table 9 – Mean novel noun scores for the TD group at time 1 and the LT group at Times 1 and 2

	TD Time 1	**LT Time 1**	**LT Time 2**
	M (SD)	**M (SD)**	**M (SD)**
Tib	5.90 (2.61)	2.15 (1.81)	3.58 (2.24)
Miggle	3.95 (2.26)	1.40 (1.70)	2.68 (2.14)
Zutter	4.20 (2.69)	1.20 (1.74)	3.11 (2.40)

Note: TD Time 1 = Typically Developing group at Time 1; LT Time 1 = Late Talking Group at Time 1; LT Time 2 = Late Talking Group at Time 2

Prompt provision and the proposed hierarchy also posed its own limitations. With regard to retention, it is not clear that the function cue did indeed provide a higher level of assistance than a gesture cue or whether it was the combined linguistic and gesture support that proved effective. Successful utilisation of the function cue assumed either a level of prior linguistic knowledge or that the child had also retained a new verb and its' link with the referent within the teaching component of the task. Given that each gesture directly

corresponded with an associated action rather than aspects of the novel items physical appearance, it is also entirely possible, that for some children, it was this which had been mapped to the referent and not the novel noun. Children were also more likely to have had some level of previous exposure to these everyday actions giving them an advantage in terms of additional exposures serving to strengthen an existing representation. This is important because it would allow children to make correct responses to the function cue without having mapped the novel item.

From a recall perspective, considering first the phonemic cue in isolation, results suggest that whilst very much an emergent skill, older LTs aged 27 to 32 months are able to make use of this support in assisting recall but not younger LTs, aged 24 to 29 months. Inclusion of an isolated phonemic prompt as a first level of support within a DA of single word learning with 2 year old LTs may therefore add little in terms of opportunity to observe modifiability. An alternative scenario however is that, similar to the Burton and Watkins' (2007) study, it was only effective in assessing partial as opposed to complete mapping of words, a sensitivity distinction not addressed by the current study. Greater levels of modifiability may therefore be observed by crediting partial recall. Reliable testing at this level with 2 year-old LTs is however likely to create new challenges in terms of differentiating partial production from a complete production acted on by age typical phonological simplification processes.

Whilst to some extent, it might appear logical that provision of the complete word in the indirect model condition would provide greater support than just the first sound, this advantage may be at least partially offset by having to extract the target from a continuous stream of speech. Hearing the phonemic cue immediately after the indirect model, either as a standalone cue or, within the context of a combined indirect model plus phonemic cue, may have increased the value of this form of assistance. Further, with the older children in Burton

and Watkins' (2007) study, the indirect model provided so much support that all participants were able to use it to achieve a correct production and it was not sensitive enough to reveal any subtle difference in level of mapping. The comparatively low levels of success within the current study suggest that the same is not true for a 2-year old sample and that there is scope for an even higher level of assistance which may reveal more about potential performance.

Although not explored statistically, it was interesting to note that even with a potential total DA score of 72 points, the TD group, time 1 DA total scores were predominantly clustered within the 14 to 35 point range. In spite of many possible response patterns, there appeared to be a task ceiling. Whilst it is stressed that this is an observation rather than a result of statistical significance, it does suggest that a much larger TD sample may have the potential to provide normative data around when change, similar to that observed in typical word learning, can be expected; a level beyond saying that *some* degree of change is imminent.

Finally, this was a small exploratory study. In addition to a small number of participants exhibiting varied profiles, the study was limited to tracking the LT participants for an 8 to 9 month period, insufficient to reliably determine eventual language status.

Implications for the Use of Dynamic Assessment with Two Year Old Late Talkers

The aim was to create a word learning task, sufficiently challenging for prompts to be required but not so difficult that a 2 year LT was unable to perform even with support. Given the response patterns observed, findings support the construct that a DA of single word learning does have potential as a valuable supplement to static assessment but in line with the exploratory nature of the study, a number of modifications are indicated.

Firstly, it is important to consider whether the task was successful in revealing optimal performance. Ultimately, the answer is mixed. In its current format, the task was simply too difficult for LTs with a mean age of 25.85 months however 3 months later, change scores were associated with progress on the PLS, an established standardised assessment. It must however be acknowledged that all correlations were in the moderate range and, increasing the strength of these initial word learning study findings are essential, if the task is to have clinical credibility.

Secondly, it is important to consider whether the selected prompts were as helpful as anticipated. Again, the answer is mixed. Overall, prompts improved performance; however, there were still a large number of 'no responses' within the LT sample at time 2. Even only very occasional semantic errors in spontaneous naming were evident within the TD group, for example, "cake" for *miggle* and "chair" for *zutter*. Neither error type was evident within the LT group. It was therefore possible that these TD children had slow mapped a fragile semantic representation capturing some of the novel object's features but that the LT children were tested much too early in their slow mapping for this to be the case. Further, the TD children were able to access a more robust taxonomical coordinate but the LTs did not have this pool of previously stored words to draw upon.

In particular, a DA targeting retention *and* recall of new words may add more to the assessment picture if it was able to capture and score for partial use of target items across greater range of recall probes.

Whilst there is still much to be said for separating out retention and recall, particularly with view to identifying receptive and expressive discrepancies and devising an individualised intervention plan, within this study there also appeared to be considerable value in the combined DA total score. Given the well documented heterogeneity of the 2-year

old LT population it is possible that this was simply a reflection of how receptive and expressive abilities work together to determine overall language competence.

Future Directions

The study has raised numerous questions that could form the foundation for future research including: reducing task complexity to obtain greater sensitivity to the underlying word learning processes of 24 month old late talking children, exploring the effect of prompt provision, obtaining normative data from a more robust TD sample, investigating smaller age brackets and longitudinal follow-up.

With the TD group demonstrating success rates in the 4.17% to 52.78% range, combined with the previously discussed theory that the task is simply outside the zone of proximal development for the LT group, it would appear sensible as a first step to look at options for simplifying the task and reducing the considerable floor effects observed. This could be done in a number of ways. Firstly, as increasing number of exposures and/or prompts provided would increase the length of the task, counterproductive both for a screening assessment and when considering the attentional capacities of a 2 year old, it would be useful to look at the impact of a reduced number of novel words. Comparing findings to previous studies, it was interesting to note that the TD spontaneous retention mean of 4.45 somewhat approximated Horst and Samuelson's (2008) findings that 24 month old TD children typically retained the first 4 name-object mappings they were exposed to following a similar 5 minute temporal lag between the teaching and testing phases. Unfortunately, more precise data was not provided in relation to this study; however, it would appear that 2 year olds have an upper retention threshold around the 4 word mark. The inclusion of familiar words may also be less helpful than first anticipated. Secondly, two correlations (retention following 10 exposures and total recall), were identified as having greatest associative value

with change scores both expressively and receptively at time 2. There is a role for exploring how these components of the DA task might be adapted to elicit meaningful, predictive data with LT children as young as 24 months.

Further investigation regarding the relative efficacy of each cue type with a 2 year old population is warranted. Cues were not counterbalanced across participants therefore it is not clear whether the proposed hierarchy was accurate. From a recall perspective, no LTs were able to make use of the phonemic prompt at time 1 however 8 found at time 2. In contrast, 5 children demonstrated an ability to utilise the indirect model at time 1and 7 at time 2. Even with the limitations imposed by the study design, this provides tentative support for the hypothesis that in line with previous findings, an indirect model provides a higher level of assistance than a phonemic cue (Burton & Watkins, 2007). Further, a phonemic cue may not provide a sufficiently high level of assistance to warrant inclusion in a DA administered with language delayed children as young as 2 years of age. There is also scope for providing even higher levels of assistance such as reducing the presence of direct competitors and/or other novel items in testing and identifying cues which might have greater facilitative effect. Within the recall condition, many of the TD group responded to the indirect model by repeating the last word heard rather than extracting the target. For example, on hearing ‘tibs’ are for blowing, they responded “blow”. Two higher levels of recall support are documented within the literature. These include a direct model and a direct model plus elicitation statement (Bain & Olswang, 1995), which explicitly prompted the child to produce the novel word. These prompts are also the basis of evidenced based, focused language stimulation (Girolametto, Pearce, & Weitzman, 1996) and milieu teaching interventions (Kouri, 2005), which would make their inclusion additionally advantageous in providing direct carryover to intervention planning. Determining any interaction effect for recall probes and novel word

variables including length, phonological composition and class would also have much clinical relevance in terms of maximising success.

It would also be beneficial to explore 24 to 26 month old LTs and 27 to 29 month old LTs as at least two distinct clinical groups on the DA, that is, develop smaller 3 month age aggregates. It was evident when allocating children to groups that a challenging mismatch was occurring between the CDI vocabulary and the PLS score on group allocation for the older age group in particular and, that in the area of word-learning, a 6 month age aggregate is simply be too broad. This is consistent with previous findings which have indicated that vocabulary acquisition is a domain more suited to 1 month norms (Desmarais, et al., 2010).

Importantly, given the instability of developmental trajectories in the language domain between 2 and 4 years of age (Dale, et al., 2003), it is important to acknowledge that this study was medium term in duration and that there is much need for further longitudinal follow-up. Some LTs had inadequate time to demonstrate catch up gains and for others who made significant gains, it was not clear how rate of progress would stabilise.

Conclusions

In conclusion, whilst there is much scope for task improvement, this early exploratory study would suggest that further development is warranted. A short, easily replicable screening assessment which incorporates many of the features of interventions currently advocated for late talkers would provide highly transferable clinical information around how the child may respond to a similar kind of input within a therapeutic context. Further, by obtaining sufficient normative data, the task has the potential to provide a quantitative score which could reveal if and when word learning abilities had started to approximate those of typically developing peers. The DA would also reveal this information before it became visible within surface language behaviours and therefore assessable via traditional static

measures. Linked to the ‘watch and see’ approach, results from the dynamic assessment have the potential to combine with static parental measures such as the CDI: Words and Sentences (Fenson, 2007) to give increased standardisation and functional value to the watch/monitoring component, all advantageous given the current pressures on speech-language services. Given the heterogeneity of the two year old late talker population as it stands, there is also the potential for transfer to the assessment of language delay in the context of known aetiologies such as prematurity and ASD where there is still much within diagnoses variability and uncertainty around spoken language outcome (Reilly, et al., 2009; Stone & Yoder, 2001)

References

Alony, S., & Kozulin, A. (2007). Dynamic assessment of receptive language in children with Down syndrome. *Advances in Speech Language Pathology, 9*, 323-331. doi: 10.1080/14417040701291415

Bain, B. A., & Olswang, L. B. (1995). Examining readiness for learning two-word utterances by children with specific expressive language impairment: Dynamic assessment validation. *American Journal of Speech-Language Pathology, 4*, 81-91.

Booth, A., McGregor, K. K., & Rohlfing, K. J. (2008). Socio-pragmatics and attention: Contributions to gesturally guided word learning in toddlers. *Language Learning and Development, 4*, 179-202.

Bricker, D., & Squires, J. (with assistance from Mounts, L., Potter, L., Nickel, R., Twombly, E., & Farrell, J.) (1999). *Ages and Stages Questionnaires (ASQ): A parent completed, child monitoring system* (2nd ed.). Baltimore, MD: Paul H. Brookes Publishing Co.

Burton, V. J., & Watkins, R. V. (2007). Measuring word learning: Dynamic versus static assessment of kindergarten vocabulary. *Journal of Communication Disorders, 40*, 335-356. doi: 10.1016/j.jcomdis.2006.06.015

Camilleri, B. (2005). Dynamic assessment and intervention: Improving children's narrative abilities. *International Journal of Language & Communication Disorders, 40*, 240-242.

Camilleri, B., & Botting, N. (2013). Beyond static assessment of children's receptive vocabulary: The dynamic assessment of word learning (DAWL). *International Journal of Language & Communication Disorders, 48*(5), 565-581. doi: 10.1111/1460-6984.12033

Camilleri, B., & Law, J. (2007). Assessing children referred to speech and language therapy: Static and dynamic assessment of receptive vocabulary. *Advances in Speech Language Pathology, 9*, 312-322. doi: 10.1080/14417040701624474

Campione, J., & Brown, A. (1987). Linking dynamic assessment with school achievement C. S. Lidz (Ed.) *Dynamic assessment: An interactional approach to evaluating learning potential*

Capone, N. C., & McGregor, K. K. (2005). The effect of semantic representation on toddlers' word retrieval. *Journal of Speech, Language, & Hearing Research, 48*, 1468-1480.

Carey, S., & Bartlett, E. (1978). Acquiring a single new word. *Papers and Reports in Child Language Development, 15*, 17-29.

Clark, E. V., & Grossman, J. B. (1998). Pragmatic directions and children's word learning. *Journal of Child Language, 25*, 1-18.

Dale, P. S., Price, T. S., Bishop, D. V. M., & Plomin, R. (2003). Outcomes of early language delay: Predicting persistent and transient language difficulties at 3 and 4 years. *Journal of Speech, Language, and Hearing Research, 46*, 544-560.

Desmarais, C., Sylvestre, A., Meyer, F., Bairati, I., & Rouleau, N. (2010). Three profiles of language abilities in toddlers with an expressive vocabulary delay: Variations on a theme. *Journal of Speech, Language & Hearing Research, 53*, 699-709. doi: 10.1044/1092-4388(2009/07-0245)

Donaldson, A. L., & Olswang, L. B. (2007). Investigating requests for information in children with autism spectrum disorders: Static versus dynamic assessment. *Advances in Speech Language Pathology, 9*, 297-311. doi: 10.1080/14417040701413720

Dunn, L., & Dunn, L. (1981). *Peabody Picture Vocabulary Test-Revised*. Circle Pines, MN: American Guidance Service.

Dunn, L., & Dunn, L. (1997). *Peabody Picture Vocabulary Test-Third Edition*. Circle Pines. MN: American Guidance Service.

Dunn, L., Dunn, L., Whetton, C., & Burley, J. (1997). *The British Picture Vocabulary Scale*. Windsor: NFER-Nelson.

Earle, C., & Lowry, L. (2011). *Target Word making Hanen happen leaders guide* (3rd ed.). Toronto, ON, Canada: The Hanen Centre.

Elleman, A. M., Compton, D. L., Fuchs, D., Fuchs, L. S., & Bouton, B. (2011). Exploring dynamic assessment as a means of identifying children at risk of developing comprehension difficulties. *Journal of Learning Disabilities, 44*, 348-357. doi: 10.1177/0022219411407865

Fenson, L. (2007). *MacArthur-Bates Communicative Development Inventories: User's guide and technical manual*. Baltimore, Md: Paul H. Brookes Publishing Company.

Fenson, L., Marchman, V. A., Thal, D. J., Dale, P. S., Reznick, J. S., & E., B. (2007). *MacArthur-Bates Communicative Development Inventories (*CDIs) (2nd ed.): Baltimore, MD: Paul H. Brookes Publishing Co.

Fuchs, D., Compton, D. L., Fuchs, L. S., Bouton, B., & Caffrey, E. (2011). The construct and predictive validity of a dynamic assessment of young children learning to read: Implications for RTI frameworks. *Journal of Learning Disabilities, 44*, 339-347. doi: 10.1177/0022219411407864

Girolametto, L., Pearce, P. S., & Weitzman, E. (1996). Interactive focused stimulation for toddlers with expressive vocabulary delays. *Journal of Speech, Language, and Hearing Research, 39*, 1274-1283. doi: 10.1044/jshr.3906.1274

Glaspey, A., & Stoel-Gammon, C. (2007). A dynamic approach to phonological assessment. *Advances in Speech Language Pathology, 9*, 286-296. doi: 10.1080/14417040701435418

Gray, S. (2005). Word learning by preschoolers with specific language impairment: Effect of phonological or semantic cues. [Article]. *Journal of Speech Language and Hearing Research, 48*, 1452-1467. doi: 10.1044/1092-4388(2005/101)

Grigorenko, E. L., & Sternberg, R. J. (1998). Dynamic testing. *Psychological Bulletin, 124*, 75-111. doi: 10.1037//0033-2909.124.1.75

Gutierrez-Clellen, V. F., & Pena, E. (2001). Dynamic assessment of diverse children: A tutorial. *Language Speech and Hearing Services in Schools, 32*, 212-224. doi: 10.1044/0161-1461(2001/019)

Hasson, N., & Botting, N. (2010). Dynamic assessment of children with language impairments: A pilot study. *Child Language Teaching and Therapy, 26*, 249-272. doi: 10.1177/0265659009349982

Hasson, N., Camilleri, B., Jones, C., Smith, J., & Dodd, B. (2013). Discriminating disorder from difference using dynamic assessment with bilingual children. *Child Language Teaching & Therapy, 29*, 57-75. doi: 10.1177/0265659012459526

Hasson, N., & Joffe, V. (2007). The case for dynamic assessment in speech and language therapy. *Child Language Teaching and Therapy, 23*, 9-25. doi: 10.1177/0265659007072142

Horst, J. S., Parsons, K. L., & Bryan, N. M. (2011). Get the Story Straight: Contextual Repetition Promotes Word Learning from Storybooks. *Frontiers in Psychology 2*, 1-11.

Horst, J. S., & Samuelson, L. K. (2008). Fast mapping but poor retention by 24-month-old infants. *Infancy, 13*, 128-157. doi: 10.1080/15250000701795598

Jitendra, A. K., & Kameenui, E. J. (1993). Dynamic assessment as a compensatory assessment approach - a description and analysis. *Remedial and Special Education, 14*, 6-18.

Jolly, H. R., & Plunkett, K. (2008). Inflectional Bootstrapping in 2-year-olds. *Language & Speech, 51*, 45-59.

Kapantzoglou, M., Restrepo, M. A., & Thompson, M. S. (2012). Dynamic assessment of word learning skills: Identifying language impairment in bilingual children. *Language, Speech, and Hearing Services in Schools, 43*, 81-96. doi: 10.1044/0161-1461(2011/10-0095)

Klee, T., Carson, D. K., Gavin, W. J., Hall, L., Kent, A., & Reece, S. (1998). Concurrent and predictive validity of an early language screening program. *Journal of Speech Language and Hearing Research, 41*(3), 627-641.

Kouri, T. A. (2005). Lexical training through modeling and elicitation procedures with late talkers who have specific language impairment and developmental delays. *Journal of Speech, Language & Hearing Research, 48*, 157-171.

Laing, S. P., & Kamhi, A. (2003). Alternative assessment of language and literacy in culturally and linguistically diverse populations. *Language Speech and Hearing Services in Schools, 34*, 44-55. doi: 10.1044/0161-1461(2003/005)

Larsen, J. A., & Nippold, M. A. (2007). Morphological analysis in school-age children: Dynamic assessment of a word learning strategy. *Language, Speech, and Hearing Services in Schools, 38*, 201-212. doi: 10.1044/0161-1461(2007/021)

Law, J., & Camilleri, B. (2007). Dynamic Assessment and its Application to Children with Speech and Language Learning Difficulties *Advances in Speech Language Pathology, 9*, 271-272.

Lidz, C. S. (1991). *Practitioner's guide to dynamic assessment*. New York, NY US: Guilford Press.

Lidz, C. S. (2002). Mediated Learning Experience (MLE) as a basis for an alternative approach to assessment. *School Psychology International, 23*, 68-84. doi: 10.1177/0143034302023001731

Lidz, C. S. (2010). Review of 'Dynamic assessment: A Vygotskian approach to understanding and promoting L2 development'. *Journal of Cognitive Education and Psychology, 9*, 87-88. doi: 10.1891/1945-8959.9.1.87

Lin, Z. (2010). Interactive dynamic assessment with children learning EFL in kindergarten. *Early Childhood Education Journal, 37*, 279-287. doi: 10.1007/s10643-009-0356-6

Markman, E. M., & Wachtel, G. F. (1988). Childrens use of mutual exclusivity to constrain the meanings of words. *Cognitive Psychology, 20*, 121-157. doi: 10.1016/0010-0285(88)90017-5

McGregor, K. K., Friedman, R. M., Reilly, R. M., & Newman, R. M. (2002). Semantic representation and naming in young children. *Journal of Speech, Language, and Hearing Research, 45*, 332-346.

Moore-Brown, B., Huerta, M., Uranga-Hernandez, Y., & Pena, E. D. (2006). Using dynamic assessment to evaluate children with suspected learning disabilities. *Intervention in School and Clinic, 41*, 209-217. doi: 10.1177/10534512060410040301

Naigles, L. R., & Hoff-Ginsberg, E. (1998). Why are some verbs learned before other verbs? Effects of input frequency and structure on children's early verb use. *Journal of Child Language, 25*, 95-120.

Olswang, L. B., & Bain, B. A. (1991). Clinical forum: Treatment efficacy, when to recommend intervention. *Language, Speech and Hearing Services in Schools 22*, 255-263.

Olswang, L. B., Feuerstein, J. L., Pinder, G. L., & Dowden, P. (2013). Validating dynamic assessment of triadic gaze for young children with severe disabilities. *American*

Journal of Speech-Language Pathology, 22, 449-462. doi: 10.1044/1058-0360(2012/12-0013)

Olswang, L. B., Long, S. H., & Fletcher, P. (1997). Verbs in the emergence of word combinations in young children with specific expressive language impairment. *European Journal of Disorders of Communication, 32*, 15-33.

Olswang, L. B., Rodriguez, B., & Timler, G. (1998). Recommending intervention for toddlers with specific language learning difficulties: We may not have all the answers, but we know a lot. *American Journal of Speech-Language Pathology, 7*, 23-32.

Patterson, J. L., Rodriguez, B., & Dale, P. S. (2013). Response to dynamic language tasks among typically developing latino preschool children with bilingual experience. *American Journal of Speech-Language Pathology, 22*, 103-112. doi: 10.1044/1058-0360(2012/11-0129)

Patterson, J. L., Rodríguez, B. L., & Dale, P. S. (2013). Response to dynamic language tasks among typically developing latino preschool children with bilingual experience. *American Journal Of Speech-Language Pathology, 22*, 103-112.

Paul, R. (1996). Clinical implications of the natural history of slow expressive language development. *American Journal of Speech-Language Pathology, 5*, 5-21.

Paul, R., & Roth, F. P. (2011). Characterizing and predicting outcomes of communication delays in infants and toddlers: Implications for clinical practice. *Language Speech and Hearing Services in Schools, 42*, 331-340. doi: 10.1044/0161-1461(2010/09-0067)

Pena, E., Gillam, R. B., Malek, M., Ruiz-Felter, R., Resendiz, M., Fiestas, C., & Sabel, T. (2006). Dynamic assessment of school-age children's narrative ability: An experimental investigation of classification accuracy. *Journal of Speech Language and Hearing Research, 49*, 1037-1057. doi: 10.1044/1092-4388(2006/074)

Pena, E., Iglesias, A., & Lidz, C. S. (2001). Reducing test bias through dynamic assessment of children's word learning ability. *American Journal of Speech-Language Pathology, 10*, 138-154. doi: 10.1044/1058-0360(2001/014)

Pena, E., Quinn, R., & Iglesias, A. (1992). The application of dynamic methods to language assessment: A nonbiased procedure. *Journal of Special Education, 26*, 269-280.

Royal New Zealand Plunket Society. (n.d.). Who we are. *plunket.org.nz*. Retrieved March 12, 2014 from http://www.plunket.org.nz/what-we-do/who-we-are/

Ram, G., Marinellie, S. A., Benigno, J., & McCarthy, J. (2013). Morphological analysis in context versus isolation: Use of a dynamic assessment task with school-age children. *Language, Speech & Hearing Services in Schools, 44*, 32-47. doi: 10.1044/0161-1461(2012/11-0023)

Reese, E., & Read, S. (2000). Predictive validity of the New Zealand MacArthur Communicative Development Inventory: Words and Sentences. *Journal of Child Language, 27*, 255-266.

Reilly, S., Bavin, E. L., Bretherton, L., Conway, L., Eadie, P., Cini, E., . . . Wake, M. (2009). The Early Language in Victoria Study (ELVS): a prospective, longitudinal study of communication skills and expressive vocabulary development at 8, 12 and 24 months. *International Journal of Speech-Language Pathology, 11*, 344-357.

Rescorla, L. (1989). The Language-Development Survey: A screening tool for delayed language in toddlers. *Journal of Speech and Hearing Disorders, 54*, 587-599.

Rescorla, L. (2002). Language and reading outcomes to age 9 in late-talking toddlers. *Journal of Speech, Language, and Hearing Research, 45*, 360-371. doi: 10.1044/1092-4388(2002/028)

Rescorla, L. (2009). Age 17 language and reading outcomes in late-talking toddlers: Support for a dimensional perspective on language delay. *Journal of Speech, Language, and Hearing Research, 52*, 16-30.

Rescorla, L., & Achenbach, T. M. (2002). Use of the Language Development Survey (LDS) in a national probability sample of children 18 to 35 months old. *Journal of Speech Language and Hearing Research, 45*. doi: 10.1044/1092-4388(2002/059)

Rescorla, L., Roberts, J., & Dahlsgaard, K. (1997). Late talkers at 2: Outcome at age 3. *Journal of Speech & Hearing Research, 40*, 556-566.

Rice, M. L., Oetting, J. B., Marquis, J., Bode, J., & Pae, S. Y. (1994). Frequency of input effects on word comprehension of children with specific language impairment. *Journal of Speech and Hearing Research, 37*, 106-122.

Rice, M. L., Taylor, C. L., & Zubrick, S. R. (2008). Language outcomes of 7-year-old children with or without a history of late language emergence at 24 months. *Journal of Speech, Language, and Hearing Research, 51*, 394-407. doi: 10.1044/1092-4388(2008/029)

Schafer, G. (2005). Infants can learn decontextualized words before their first birthday. *Child Development, 76*, 87-96. doi: 10.2307/3696710

Semel, E., Wiig, E., H. & Secord, W. A. (2000). *Clinical Evaluation of Language Fundamentals,Third Edition*, London: Psychological Corporation

Spiegel, C., & Halberda, J. (2011). Rapid fast-mapping abilities in 2-year-olds. *Journal of Experimental Child Psychology, 109*, 132-140. doi: 10.1016/j.jecp.2010.10.013

Stone, W. L., & Yoder, P. J. (2001). Predicting spoken language level in children with Autism Spectrum Disorders. *Autism, 5*, 341-361. doi: 10.1177/1362361301005004002

Storkel, H. L. (2001). Learning New Words: Phonotactic Probability in Language Development *Journal of Speech, Language & Hearing Research, 44*, 1321-1337.

Storkel, H. L. (2003). Learning new words II: phonotactic probability in verb learning. *Journal of Speech, Language & Hearing Research, 46*, 1312-1323.

Swanson, H. L. (2011). Dynamic testing, working memory and reading comprehension growth in children with reading disabilities. *Journal of Learning Disabilities, 44*, 358-371. doi: 10.1177/0022219411407866

Thal, D. J., Tobias, S., & Morrison, S. (1991). Language and gesture in LT: A one-year follow-up. *Journal of Speech and Hearing Research, 34*, 604-612.

Ukrainetz, T. A., Harpell, S., Walsh, C., & Coyle, C. (2000). A preliminary investigation of dynamic assessment with native American kindergartners. *Language Speech and Hearing Services in Schools, 31*, 142-154.

Vygotsky, L. S. (1978). 'Mind in society' V. J.-S. M. Cole, S. & a. E. S. Scribner, (Eds.), *The development of higher psychological*

processes.

Watt, N., Wetherby, A., & Shumway, S. (2006). Prelinguistic predictors of language outcome at 3 years of age. *Journal of Speech Language and Hearing Research, 49*, 1224-1237. doi: 10.1044/1092-4388(2006/088)

Waxman, S. R., & Kosowski, T. D. (1990). Nouns mark category relations: Toddlers' and preschoolers' word-learning biases. *Child Development, 61*, 1461-1473.

Zimmerman, I., Steiner, V., & Pond, R. (2002). *Preschool Language Scale, Fourth Edition (PLS-4)* San Antonio, TX: The Psychological Corporation.

Zubrick, S. R., Taylor, C. L., Rice, M. L., & Slegers, D. W. (2007). Late language emergence at 24 months: An epidemiological study of prevalence, predictors, and covariates.

Journal of Speech, Language, and Hearing Research, 50, 1562-1592. doi: 10.1044/1092-4388(2007/106)

Appendix A: Information Sheet

A Dynamic Assessment of Single Word Learning in Two Year Old Late Talkers

Information for parents/whānau

An invitation to participate in a research project on children's early language development

The researcher: Victoria Singer
Phone: 021 127 5716
Email: Victoria.singer@gmail.com
Project Supervisor: Professor Thomas Klee
E-mail: Thomas.klee@canterbury.ac.nz
Phone: 03 364 2987 #8501
University Website: www.cmds.canterbury.ac.nz

Kia ora! Hello!

I would like to invite you to take part in a research project about children's early language development. Before deciding if you'd like to participate, please read through this leaflet. If you have any questions about the project, please feel free to get in touch with me by phone or email. My contact details are above.

This research project is for my MSc thesis in Speech and Language Sciences. I am a qualified speech-language therapist having obtained a BSc in Speech and Language Sciences in 2001 and worked in both health and education settings in New Zealand for the past nine years. My work will be supervised by Professor Thomas Klee and Dr Catherine Moran from the Department of Communication Disorders at the University of Canterbury. Professor Klee's contact details and the university web-site are listed above.

The project has been reviewed and approved by the University of Canterbury Human Ethics Committee (Email human-ethics@canterbury.ac.nz or post to Human Ethics Committee, University of Canterbury, Private Bag 4800, Christchurch).

How can I volunteer to take part?

If you would like to take part in this project, please just contact me using the details above. If you have any concerns about your child's speech and language development, you can talk about them with me. If your child is a late talker, that is, they have less than 50 spoken words and/or no word combinations I would like to track their progress for nine months. If you would like to access additional SLT support for your child within this time frame I can help you with organising this.

There are no risks to you or your child as a result of participating in this study.

What is the purpose of the project?

Some children learn to talk quickly and some take their time. Many two-year-olds are joining words together in sentences while others are saying only a few or no words. Some children who start off slowly catch up over time, while others have on-going problems with language. The purpose of this study is to explore whether a play based assessment task can be used to

differentiate between two-year old late talkers who are language impaired and therefore going to benefit most from early intervention and those who are late bloomers who will catch up in their own time without this support.

What is involved?
The project will take place in three parts. In the first part, I am searching for children between 24 and 30 months of age. If your child will be in this age range between now and March 2013, you can participate! You will need to fill in one form about your child's use of words and sentences (this is called the MacArthur-Bates Communication Development Inventory) and I will administer both a standardized language assessment (this is called the Pre-school Language Scales) and a new play based language assessment which I have devised. In the play based assessment, I want to see how quickly your child learns new words and what kind of support is helpful in terms or retaining and later recalling these new words. With your consent, I will video record the play based assessment. Consent to participating within the project does not oblige you to consent to having your child video-recorded. If your child has 'typically developing' language skills then this is all you need to do. Completing this part of the project will take approximately two hours. I will need to see you and your child in a quiet space with as few distractions as possible. This can be in your own home or I can arrange an alternative space at your local community or health centre. You will need to be present throughout both assessments.

If your child is a 'late talker', I will ask you to participate in both part 2 and part 3 of the project. In part 2, I will repeat the play based assessment. This will be approximately three months after your child first completed it. In part 3, I will complete the standardized assessment. This will be approximately nine months after your child first completed it. Each of the follow-up sessions will take approximately one hour. Participating in the first part of the project does not oblige you to participate in follow-ups and there will be no penalty for withdrawing at any time.
I need at least 40 parents in the Wellington region to participate. Whether your child is talking a lot or has not yet begun to talk, I would like to hear from you.

What will happen to the information that I and my child provide?
None of the personal information you provide us will be made public. Your responses on the questionnaires will be kept confidential. They will be combined with those of other participants and summarised. Your questionnaire, child's assessment scores and any video recorded information will be kept in a secure locked cabinet. Only I and my supervisors at the University of Canterbury will have access to them. At the end of the project, all this information will be kept for five years and then destroyed.
I will write up the results of the research as my Master of Science thesis in Speech and Language Sciences. If you would like to receive a brief summary of my findings, please tick your preferred form of contact box on the consent form and I will send you one when the study is completed. Findings may also be published in speech-language therapy related scientific and/or professional journals or presented at conferences. This is so I can share my findings with other speech-language therapists. Please tell me if you would like to know more or have any concerns about this aspect of the study.

Thank you and look forward to hearing from you, **Victoria Singer**

Appendix B: Informed Consent Form

Consent Form

A Dynamic Assessment and Word Learning in Two Year Old Late Talkers

I have read and understood the information that was given to me about the research project named above. I have had a chance to ask questions and have had them answered. I understand that my participation in this project is voluntary and that I can withdraw without penalty at any time. I understand that the information you collect from me will remain confidential and will be securely stored. I understand that any presentations or publications resulting from this project will not refer to me or anyone in my family by name. I understand that this project has been reviewed ***and approved*** by the University of Canterbury Human Ethics Committee. On this basis, I agree to participate in this research project.

MY CHILD'S NAME (please print): .....
MY NAME (please print): .....
My Signature: .....
Date: .....
Address: .....
.....

❑ I would like you to send me a brief summary of your findings when the study is complete.
My preferred form of contact is:

❑ Email _____
❑ Cell phone _____
❑ Landline _____
❑ Post _____
❑ Other (please state) _____

Main Researcher:
Victoria Singer
Email: victoria.singer@gmail.com
Phone: 021 127 5716

Project Supervisor:
Professor Thomas Klee
Email: Thomas.klee@canterbury.ac.nz
Phone: 03 364 2987 #8501

Appendix C: DA Task Script:

Nouns:

Part I

We are going to play and learn some new words. Let's try and remember the new words and the things that they go with. Remembering the names of the new things that we see is very important. We're going to play and I'm going to help you.

It's a lovely sunny day. Giraffe and Goat are at the park.

It's time for a game (toy category).

Look, Giraffe has a **ball**. Fun! What is it? (Imitation Opportunity) Yes, **ball**. Giraffe is kicking the **ball** to you, *child's name*. (Examiner manipulates objects)

Part II

Look, he wants you to have a turn with the **ball**. What is it? (Opportunity provided for child led object manipulation and third imitation) Yes, **ball**. **Balls** are for kicking and rolling. (Examiner supports verbal with iconic gesture)

Part I

It's time for a snack (food category).

Look, Giraffe has an **apple**. Yum! What is it? (Imitation opportunity) Yes, **apple**. Giraffe is cutting the **apple** *child's name.* (Examiner manipulates objects)

Part II

Look, she's giving you some **apple**. What is it? (Opportunity provided for child led object manipulation and third imitation) Yes, **apple**. **Apples** are for eating. (Examiner supports verbal with iconic gesture)

Part I

It's time for another game (toy category).

Look, Goat has a **tib**. Fun! What is it? (Imitation Opportunity) Yes, **tib**. Goat is blowing the **tib** *child's name*. (Examiner manipulates objects)

Part II

Look, she wants you to have a turn with the **tib**. What is it? (Opportunity provided for child led object manipulation and third imitation) Yes, **tib**. **Tibs** are for blowing and making music. (Examiner supports verbal with iconic gesture)

Part I

It's time for another **snack** (food category).

Look, Goat has a **miggle**. Yum! What is it? (Imitation opportunity) Yes, **miggle**. Goat is peeling the **miggle** *child's name.* (Examiner manipulates objects)

Part II

Look, she's giving you some **miggle**. What is it? (Opportunity provided for child led object manipulation and third imitation) Yes, **miggle**. **Miggles** are for peeling and licking. (Examiner supports verbal with iconic gesture)

Part I

I wonder what other fun things are in the park (everyday objects).

Look, Giraffe sees a **flower**. Pretty! What is it? (Imitation opportunity) Yes, a **flower**. Giraffe is picking the **flower** *child's name*. (Examiner manipulates object)

Part II

Look, she's giving you a **flower**. What is it? (Opportunity provided for child led object manipulation and third imitation) Yes, **flower**. **Flowers** are for picking and smelling. (Examiner supports verbal with iconic gesture)

Part I

What else is in the park?

Look, Goat sees a **zutter**. Yeah! What is it? (Imitation opportunity) Yes, a **zutter**. Goat is sitting on the **zutter** *child's name*. (Examiner manipulates object)

Part II

Look, she wants you to have a turn on the **zutter**. What is it? (Opportunity provided for child led object manipulation and third imitation) Yes, **zutter**. **Zutters** are for sitting on and riding. (Examiner supports verbal with iconic gesture)

Verbs:

Part I

It's time for another game with Goat and Giraffe! This time let's see if we can remember all the things that they like to do!

Giraffe decides to do some **pushing**. What's he doing? (Imitation opportunity) Yes, **pushing**. Giraffe is **pushing** the trolley. (Examiner manipulates objects)

Part II

Look, Giraffe needs some help with **pushing**. What's he doing? (Opportunity provided for child led object manipulation and third imitation) Yes, **pushing**. He's **pushing** very fast! (Examiner supports verbal with iconic gesture)

Part I

Giraffe decides to do some **jumping**. What's he doing? (Imitation opportunity) Yes, **jumping**. Giraffe is **jumping** over the stripy pole. (Examiner manipulates objects)

Part II

Look, Giraffe loves **jumping**. What's he doing? (Opportunity provided for child led object manipulation and third imitation) Yes, **jumping**. He's **jumping** high, high, high! (Examiner supports verbal with iconic gesture)

Part I

Goat decides to do some **heeping**. What's he doing? (Imitation opportunity) Yes, **heeping**. Goat is **heeping** with blocks. (Examiner manipulates objects)

Part II

Look, Goat needs some help with **heeping**. What's he doing? (Opportunity provided for child led object manipulation and third imitation) Yes, **heeping**. He's **heeping** up, up, up! (Examiner supports verbal with iconic gesture)

Part I

Goat decides to do some **wisping**. What's he doing? (Imitation opportunity) Yes, **wisping.** Goat is **wisping** on the slide. (Examiner manipulates objects)

Part II

Look, Goat loves **wisping**. What's he doing? (Opportunity provided for child led object manipulation and third imitation) Yes, **wisping**. He's **wisping** down, down, down! (Examiner supports verbal with iconic gesture)

Part I

After all that play, it's time for a rest.

Giraffe is **sleeping**. What's he doing? (Imitation opportunity) Yes, **sleeping**. Giraffe is **sleeping** on the picnic blanket. (Examiner manipulates objects)

Part II

Look, Giraffe is still **sleeping**. What's he doing? (Opportunity for child led object manipulation and third imitation) Yes, **sleeping**. He's **sleeping**. Zzzzz. (Examiner supports verbal with iconic gesture)

Part I

Goat is **nooding**. What's he doing? (Imitation opportunity) Yes, **nooding**. Goat is **nooding** the puppy. (Examiner manipulates objects)

Part II

Look, Goat is still **nooding**. What's he doing? (Opportunity for child led object manipulation and third imitation) Yes, **nooding**. He's **nooding** with his goat hands. (Examiner supports verbal with iconic gesture)

Appendix D: Retention and Recall Probes

Retention prompts for nouns demonstrated with tib:

1. (Elicitation question) *Where's the tib? Can you find it?*
2. (Gesture cue) *Remember it was the one for (blow gesture)? Can you find it?*

   Examiner provides gesture cue(s) *without* labelling object or action.
3. (Semantic cue)*Which one was for blowing (and making music)? Can you find it?*

   Examiner provides primary function in first retention trial and both functions in second retention trial.

Recall prompts for nouns demonstrated with tib:

1. (Elicitation question) *What is it?*
2. (Phonemic Cue) *It's a /t/…?*
3. (Indirect Model) *Tibs are for blowing… can you remember what it is?*

Retention prompts for verbs demonstrated with nooding:

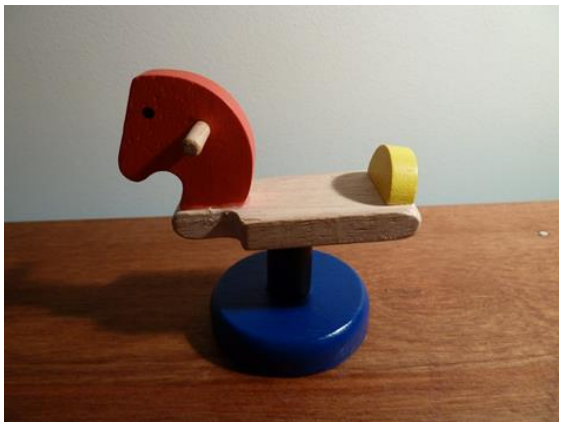
1. (Elicitation question)*Which one is nooding? Can you find it?*
2. (Gesture cue) *Remember it was the one (nood gesture)? Can you find it?* Examiner provides gesture cue(s) *without* labelling object or action.

Recall prompts for verbs demonstrated with nooding:

1. (Elicitation question) *What's goat doing?*
2. (Phonemic Cue) *Goat is /n/…?*
3. (Indirect Model) *Goat is nooding his puppy… can you remember what he's doing?*

Appendix E: Photographs of the DA Task Materials used in Testing

Nouns:



Verbs:

Appendix F: Scoring Form Completed Example

			Retention				***Recall***			
Child's Name and Date	*Number of Exposures*	*Word*	*Spontaneous*	*Gesture Cue*	*Function Cue (nouns only)*	*Exposure Score*	*Spontaneous*	*Phonemic Cue*	*Indirect Model*	*Exposure Score*
	3	Tib	Zutter	✓		4	-	-	-	3
		Miggle	Zutter	Apple	-		-	/m/	✓	
		Zutter	Miggle	✓	-		-	✓ /z/ - /d/		
	10	Tib	Zutter	✓		8	-	/t/	✓ /t/ - /d/	4
		Miggle	✓				-	/m/	✓	
		Zutter	✓				-	✓ /z/ - /d/		
	Combined		6	6	0	12	0	4	3	7
	3	Heeping	Wisping	✓		7	Building	/h/	Block /bl/ - /l/, /k/ - /t/	0
		Wisping	Nooding	✓			-	/w/	Slide /sl/ - /l/	
		Nooding	✓				-	/n/	Puppy /p/ - /b/	
	10	Heeping	Pushing	✓		7	Building	/h/	Block	0
		Wisping	Jumping	✓			Go	/w/	-	
		Nooding	✓				-	/n/	Puppy	
	Combined		6	8	-	14	0	0	0	0
	Total					***26***				***7***

UNIVERSITY OF CANTERBURY

HUMAN ETHICS COMMITTEE

Secretary, Lynda Griffioen
Email: human-ethics@canterbury.ac.nz

Ref: HEC 2012/119

14 September 2012

Victoria Singer
Department of Communication Disorders
UNIVERSITY OF CANTERBURY

Dear Victoria

The Human Ethics Committee advises that your research proposal "A dynamic assessment of single work learning in late talking two year olds" has been considered and approved.

Please note that this approval is subject to the incorporation of the amendments you have provided in your email of 10 September 2012.

Best wishes for your project.

Yours sincerely

pp LGriffioen

Lindsey MacDonald
Chair
University of Canterbury Human Ethics Committee

University of Canterbury Private Bag 4800, Christchurch 8140, New Zealand. www.canterbury.ac.nz

E 8