ASSESSING PRODUCTIVITY AND STIMULABILITY OF TENSE AND AGREEMENT MORPHEMES IN TYPICALLY DEVELOPING CHILDREN AND AAC SPEAKERS

by

Thomas Kovacs

B.A., University of Iowa, 2004

M.A., University of Colorado, 2006

Submitted to the Graduate Faculty of

The School of Health and Rehabilitation Sciences in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2017

UNIVERSITY OF PITTSBURGH

School of Health and Rehabilitation Sciences

This dissertation was presented

by

Thomas Kovacs

It was defended on

November 1, 2017

and approved by

Michael Walsh Dickey, Ph.D., Associate Professor, Department of Communication Science and Disorders, University of Pittsburgh

J. Scott Yaruss, Ph.D., CCC-SLP, Professor, Department of Communicative Sciences and Disorders, Michigan State University

Eric Nyberg, Ph.D., Professor, Language Technologies Institute, Carnegie Mellon University

Dissertation Advisor: Katya Hill, Ph.D., CCC-SLP, Associate Professor, Department of Communication Science and Disorders, University of Pittsburgh

ASSESSING PRODUCTIVITY AND STIMULABILITY OF TENSE AND AGREEMENT MORPHEMES IN TYPICALLY DEVELOPING CHILDREN AND AAC SPEAKERS

Thomas Kovacs, Ph.D.

University of Pittsburgh, 2017

Expressive morphology skills are rarely assessed in children who use augmentative and alternative communication (AAC). Pediatric AAC speakers may therefore have undocumented difficulty acquiring morphemes in the English tense and agreement system. Optional infinitive (OI) theory and gradual morphosyntactic learning (GML) theory predict different patterns of development within the tense and agreement system and different patterns of cross-morpheme generalization. This dissertation tested these competing hypotheses and used theoretically driven tasks to assess the tense and agreement systems of typically developing children and pediatric AAC speakers with cerebral palsy.

Experiment 1: Play-based language samples were collected from typically developing 30-54 month olds. Tense and agreement morpheme use and productivity were measured in samples of 150 multi-morpheme utterances. Tense marker totals and productivity scores increased with age. Morpheme category productivity grew at different rates in different categories, which was most consistent with predictions of GML theory. COPULA BE productivity was at ceiling. *-3s*, *-ed*, and AUXILIARY DO productivity were all significantly higher than AUXILIARY BE productivity. An unexpected divergence between *-3s* and *-ed* productivity was not consistent with either theory.

Experiment 2: Tense and agreement morpheme stimulability was assessed in the children from Experiment 1. Communication modality (spoken or graphic symbol) was randomly

assigned. No significant age-by-communication modality interactions were found. When significant main effects of age were found, stimulability increased with age. When significant main effects of communication modality were found, stimulability was higher in the spoken modality. Morpheme category stimulability grew at similar rates across categories, which was consistent with predictions of OI theory.

Experiment 3: Pretest tense marker use, productivity, and stimulability were assessed in pediatric AAC speakers. A target tense marker that was stimulable and not used in the pretest assessment was identified for each participant. A short course of intervention focused on production of the target tense marker and copula *is*. Both participants used target tense markers and copula *is* more productively in a posttest assessment. Possible cross-morpheme generalization was found for tense markers that shared grammatical features with the target or copula *is*. This was most consistent with predictions of GML theory.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

PREFACE

DEDICATION

To the dozens of families who contributed to this project:

Thank you all so very much for welcoming me into your homes to share your lives for a few brief hours. Observing such a diverse group of excellent parents in action was a truly unique and blessed experience. The lessons and joys of this experience extend far beyond the text of any research paper.

To my own family:

Thank you Mom, Dad, Alan, Cris, and Aia. You are blessings from God and my greatest inspiration. You have helped push me through the seemingly infinite challenges of a PhD program and helped me reach farther than I thought I could in totally new directions. Thank yinz from the bottom of my heart.

ACKNOWLEDGEMENTS

Thanks to the numerous people at all stages who helped make this dissertation a reality. Thank you to my advisor, Dr. Katya Hill and dissertation committee, Dr. Michael Dickey, Dr. J. Scott Yaruss, and Dr. Eric Nyberg. This dissertation was supported by the AY 2017-2017 Provost's Development Fund at the University of Pittsburgh and the School of Health and Rehabilitation Science Development Fund. This project was included in the Pitt+Me® Registry, and institutional research participant registry supported by the NIH through Grant Number UL1TR001857. Thank you to Mark Zucco for creating original graphic art, photography, and video to this project and Semantic Compaction Systems for sponsoring time for Mark's efforts. Additional picture stimuli were provided by Laureate Learning Systems. Software was donated by Saltillo. Thank you to Catharine Burns for assistance with reliability testing. Additional thanks to Bruce Baker, Bob Conti, Kay Chen, Sheila Pratt, Nicole Pero, Suzanne Feinstein, and Hannah Famili.

LIST OF ABBREVIATIONS

***-3s*:** regular third person singular, used in reference to both the *-3s* tense marker and the THIRD PERSON SINGULAR morpheme category

AAC: augmentative and alternative communication

AGR: agreement (as a functional category)

AUXILIARY: when written in all caps, AUXILIARY refers to a morpheme category containing multiple tense markers: AUXILIARY DO

auxiliary: when written in lower case, auxiliary refers to a specific tense marker: auxiliary *do*

COPULA: when written in all caps, COPULA refers to the COPULA BE morpheme category, containing multiple tense markers

copula: when written in lower case, copula refers to a specific tense marker, such as copula *is*

C-unit: clausal unit

***-ed*:** regular past tense, used in reference to both the *–ed* tense marker and the REGULAR PAST TENSE -ED morpheme category

GML: gradual morphological learning

I: inflection (as a constituent/functional category)

MLU/MLUm: mean length of utterance/mean length of utterance in morphemes

MSL: mean syntactic length

OI: optional infinitive

Prog: progressive aspect

TNS: tense (as a functional category)

V: verb (as a constituent)

GLOSSARY

Accent 1000: a high-tech SGD manufactured by The Prentke Romich Company, which supports several versions of the Unity™ Minspeak™ Application Program.

access method: Strategies and interface technologies that an AAC speaker uses to select language content on an AAC system. Examples include selecting symbols using a finger on a touch screen or scanning through an array of symbols using switches.

aided language stimulation: providing an AAC speaker with language input and using the AAC system to explicitly model production of target language structures in the AAC modality by selecting symbols (Goosens', 1989). Can be supplemented with verbal speech or speech output from the AAC system. Also referred to as aided language modeling (K. D. R. Drager et al., 2006) and System for Augmenting Language (Romski & Sevcik, 1996; Wilkinson, Romski, & Sevcik, 1994).

communication modality: the modality that a person uses for expressive communication or production of linguistic output, such as the spoken, written, or graphic symbol modality

control interface: an umbrella term for technology features that a human user uses to control a piece of technology, such as features used for making selections on an AAC system.

C-unit: an utterance segmented from a language sample following syntactic rules, structurally defined as "an independent clause and its modifiers" (Loban, 1976, p. 9)

language representation methods: methods of representing language content on AAC systems (Hill, 2010). Across devices and manufacturers, there are only three known language representation methods:

alphabet based methods: methods of representing language content using orthography, including spelling, word prediction, and orthographic word selection (selection of whole words represented with text).

single meaning pictures: graphic symbols that have a 1:1 mapping to a referent. The referent may be a pre-stored vocabulary word or a pre-programmed utterance.

semantic compaction: a closed set of sequenced multi-meaning icons, each of which is associated with a variety of different referents. An AAC speaker may produce a target pre-stored vocabulary word by selecting a corresponding sequence of two or more multi-meaning icons in a specific order.

message window: an area on the display of as SGD that displays messages generated by an AAC speaker in real-time

Minspeak™: A commercial name for AAC systems using the semantic compaction language representation method.

obligatory context: a token context with linguistic and situational constraints that would oblige a fluent adult speaker to use a given morpheme (Brown, 1973). Obligatory contexts are specified by broader conversational contexts. Obligatory contexts are counted in token-based measures of morpheme accuracy.

pre-stored vocabulary: words and morphemes that are stored in the available language content of an AAC system as separate units so that they can be produced using single meaning pictures, orthographic word selection, or sequenced, multi-meaning icons. The pre-stored vocabulary can include multiple inflected forms of one lemma, such as different forms of the same verb.

pre-programmed utterance: whole utterances that are stored in the available language content of an AAC system as a single unit so that they can be produced using single meaning pictures, orthographic word selection, or sequenced, multi-meaning icons. . For example, "My name is Tom Kovacs." Pre-programmed utterances are excluded from most language sample analyses because they are not considered spontaneous.

speak display: a control function directing an SGD to speak the text displayed in a message window

speech output: speech produced by an AAC system using synthesized or digitized speech

spelling mode: a mode used to generate text (or speech) on an AAC system using spelling. Many SGDs allow AAC speakers to freely toggle back and forth between a mode for producing pre-stored vocabulary/pre-programmed utterances and a spelling mode.

spontaneous novel utterance/spontaneous utterance: an utterance that is generated as it is produced, rather than a non-spontaneous utterance that is retrieved from memory.

spontaneous novel utterance generation: The process of combining linguistic units to generate a spontaneous novel utterance. AAC speakers generate spontaneous novel utterances by combining spelled and/or pre-stored vocabulary words to form larger utterances. By definition, spontaneous novel utterances are generated using multiple language events (multiple spelled characters and/or pre-stored vocabulary words).

sufficiently different contexts: use of a tense marker in different semantic contexts, such as production of different regular past tense verbs or production of COPULA BE forms in different subject/tense marker combinations. Sufficiently different contexts of morpheme use are counted in type-based measures of morpheme productivity.

stimulability: observable evidence of learning potential, indicated by the ability to complete a task following a model or cue. For example, ability to use a tense marker in the graphic symbol modality following aided language stimulation.

tokens: instances of an action or a behavior, such as the number of times a tense marker is used

types: different examples of an action or a behavior, such as the number of different ways a tense marker is used (i.e. number of different past tense verbs)

Unity™ Language Application Program: A series of commercially available language application programs for SGDs that provide access to all language representation methods.

user interface: an umbrella term for technology features that provide a human user with sensory feedback about the current status of the technology, such as display features that provide information about words that can be produced with an AAC system.

zone of proximal development: in Vygotsky's (1934/2012) domain-general learning theory, tasks that a child can currently complete with guidance but cannot complete independently

1.0 STATEMENT OF PURPOSE

Children who use augmentative and alternative communication (AAC) systems for expressive language production (pediatric AAC speakers) are at risk for developing expressive morphology impairments. These children may present with relative weaknesses in expressive morphology skills that are similar to weaknesses observed in their verbally speaking peers. No *a priori* reasons exist for assuming that the heterogeneous population of pediatric AAC speakers develops a fundamentally different morphosyntactic system and different patterns of strengths and weaknesses than their verbally speaking peers. Morphemes that verbally speaking children with developmental language disorders are known to have difficulty acquiring should be accounted for any time the expressive morphology skills of pediatric AAC speakers are assessed. For example, there is robust evidence showing that children with specific language impairment have difficulty using morphemes in the tense and agreement system (Bedore & Leonard, 1998; Gladfelter & Leonard, 2013; Hadley & Rice, 1996; Hadley & Short, 2005; Rice, Wexler, & Cleave, 1995; Rice, Wexler, & Hershberger, 1998).

The tense and agreement system should be assessed in pediatric AAC speakers, who may experience similar, but undiagnosed difficulties. Pediatric AAC speakers present with a wide range of etiologies that are associated with different language profiles in verbally speaking children. For example, many pediatric AAC speakers present with severe motor speech disorders secondary to primary diagnoses of Down syndrome or cerebral palsy. Verbally speaking children with Down syndrome are known to have relative weaknesses in morphosyntax, with selective weaknesses in expressive morphosyntax relative to both lexical skills and receptive morphosyntax (Chapman, 2006; Chapman & Hesketh, 2000; Chapman, Schwartz, & Kay-Raining Bird, 1991; Chapman, Seung, Schwartz, & Kay-Raining Bird, 1998). In contrast, verbally speaking children with cerebral palsy often have relatively intact language systems. No studies have investigated whether or not the morphosyntactic skills of pediatric AAC speakers are comparable to verbally speaking peers with similar diagnoses. Assessment of expressive morphology skills is necessary for detecting and treating morphological disorders in pediatric AAC speakers. The lack of assessment in this area corresponds to a growing body of intervention studies for pediatric AAC speakers that primarily focuses on building lexical skills and teaching early-developing multi-word constructions that do not require the use of grammatical morphemes. This critical assessment and intervention gap has direct health and educational consequences for children with undetected and untreated delays in morphological development.

The only study to date that has assessed the expressive morphology skills of pediatric AAC speakers focused on establishing baseline levels of performance using specific morphemes in isolation (Blockberger & Johnston, 2003). No studies have used any other form of expressive morphology assessment to inform the selection of intervention goals for pediatric AAC speakers or characterize the way that pediatric AAC speakers use a morphosyntactic system observed in verbally speaking children. This lack of assessment is striking, given the critical role of expressive morphology assessment in detecting specific language impairment and other developmental language disorders. Without proper assessment, these disorders may go unidentified in pediatric AAC speakers with concomitant motor speech impairments. No studies have compared language skills across communication modalities to test the null hypothesis that expressive language skills develop in similar ways across modalities. If this null hypothesis is true, then the robust external evidence on spoken language development can be used to guide intervention programs for AAC speakers. For example, known strategies for assessing morphosyntactic systems in verbally-speaking children could potentially be adapted for assessing morphosyntactic systems in pediatric AAC speakers and yield crucial assessment data to guide decisions about intervention targeting developmentally appropriate skills for individual children.

Researchers studying language development in verbally speaking children have held that groups of functionally related morphemes that share an underlying grammatical system will emerge and develop together as a morphosyntactic system. The earliest formal description of morphosyntactic systems may be found in Radford's (1990) formative study on the development of functional categories. If functionally related morphosyntactic systems emerge and develop together, then clinicians and researchers assessing the expressive morphology skills of any child should seek to measure the child's expressive use of unified morphosyntactic systems instead of focusing exclusively on the expressive use of isolated morphemes.

Research on the expressive morphology of verbally speaking children with developmental language impairments suggests that children can have selective difficulties acquiring specific morphosyntactic systems rather than global difficulties acquiring all grammatical morphemes in their language (Bedore & Leonard, 1998; Kamhi, 2014). For example, Bedore and Leonard (1998) used discriminant function analysis to determine whether measures of grammatical morpheme use could accurately differentiate between verbally speaking preschoolers with specific language impairment and an age-matched group of children with typical language development. A composite measure of noun morphology measuring children's overall accuracy using possessive *'s*, plural *-s*, and articles *the, a*, and *an* had low sensitivity for classifying children with specific language impairment. A composite measure of finite verb morphology measuring children's overall accuracy using morphemes marked for tense and agreement (past tense *-ed,* third person singular *-3s*, COPULA BE, and AUXILIARY BE) had higher sensitivity and specificity, but still misclassified some children with specific language impairment as typically developing.

In a follow-up study replicating the analyses of Bedore and Leonard (1998) in age-matched groups of school-age children with specific language impairment and typical language development, both composite measures had lower sensitivity (Moyle, Karasinski, Ellis Weismer, & Gorman, 2011). The finite verb morphology composite may have had lower sensitivity in samples of older children because children with specific language impairment eventually acquire expressive use of these morphemes, even if development is delayed (Moyle et al., 2011). The combined results from Bedore and Leonard (1998) and Moyle et al. (2011) suggest that regular past tense (*-ed*), regular third person singular (*-3s*), COPULA BE, and AUXILIARY BE are part of an integrated tense and agreement morpheme system that is particularly difficult for some children to acquire on a typical developmental timeline. No prior studies have assessed development of the unified tense and agreement system in pediatric AAC speakers. The tense and agreement systems of pediatric AAC speakers should also be assessed to identify AAC speakers who present with similar patterns of selective weaknesses. Some pediatric AAC speakers may have undiagnosed difficulties acquiring a unified tense and agreement system, which may persist beyond the typical developmental timeline.

Several theories have been proposed to explain how several different functionally related morphemes grow together as a unified tense and agreement system. The optional infinitive (OI) theory originally posited by Wexler (1994) predicts that several morphemes in a unified tense and agreement system will grow with highly similar developmental trajectories across a genetically determined maturational timeline. In the OI theory, systematic developmental changes happen in tandem across individual surface forms (Rice et al., 1998). In contrast, the gradual morphosyntactic learning (GML) theory predicts that certain categories of morphemes in a larger tense and agreement system will grow together because they have similar positional processing contexts, while morphemes with different positional processing contexts will grow at different rates (Rispoli & Hadley, 2011; Rispoli, Hadley, & Holt, 2012). The OI theory is a null hypothesis that morphemes in a unified system will mature simultaneously, while the GML theory is an alternative hypothesis that morphemes in the same unified system will grow at different rates through processes of maturation and learning.

Expressive language assessment provides direct evidence about how children express themselves using language (Paul, 2001). Grammatical morphology skills should be accounted for any time a child's expressive language skills are assessed, regardless of the modality that the child uses for expressive language production. A clinician cannot adequately characterize or efficiently treat an expressive morphology impairment without obtaining direct evidence of a child's expressive morphology skills. This direct evidence is necessary for four basic purposes: screening for potential disorders, establishing baseline levels of expressive morphology skills, selecting specific goals for intervention targeting expressive morphology skills, and measuring ongoing change during intervention (Paul, 2001; Westby, Stevens Dominguez, & Oetter, 1996).

Many clinicians recommend conducting an assessment that allows one to identify skills within a given child's zone of proximal development (Vygotsky, 1934/2012) when trying to select specific goals for intervention targeting expressive morphology. Stimulability testing, which is primarily used in the area of phonology, can be used to identify morphemes that a child can produce with adult guidance. The process of stimulability testing can provide overt evidence of learning potential for individual morphemes, indicating that a morpheme is an appropriate target for therapy in a child's zone of proximal development.

The OI theory predicts that all tense markers in the system grow and develop together (Rice et al., 1998). This suggests that there is one zone of proximal development across tense markers. In this case, tense markers should have similar levels of stimulability across morpheme categories. OI theory describes a maturational system. In such a system, intervention focused on increasing production of one tense-marker should not generalize to other tense markers because generalized learning does not contribute to maturational growth.

In contrast, the GML theory predicts that tense markers in different categories will grow at different rates (Rispoli & Hadley, 2011; Rispoli et al., 2012). This suggests that there are multiple zones of proximal development in the tense and agreement system. In this case, tense markers in different morpheme categories should have different levels of stimulability. In addition, intervention targeting increased production of one tense marker should generalize to increased production of other tense markers with similar features in positional processing.

The purposes of this dissertation project were to test competing hypotheses of the OI and GML theories, and to use theoretically driven tasks to assess the expressive use of the tense and agreement system in typically developing children and pediatric AAC speakers acquiring English as a first language. This project included three experiments, each of which tested against predictions of the OI theory as a null hypothesis. The purposes and research questions of each experiment are summarized in Table 1.1.

The first two experiments of this dissertation provide critical evidence on growth and assessment of tense and agreement morphemes in typically developing children. Experiment 1 tested the competing predictions of OI and GML theory in spontaneous language samples from a cross-sectional sample of typically developing children. Experiment 2 used a structured stimulability task to test the predictions of these same hypotheses in the same cross-sectional sample. Experiment 2 compared growth of tense and agreement morpheme stimulability across the spoken and graphic symbol modalities. This was the first experiment to test the broad null hypothesis that language skills develop in similar ways across communication modalities. These experiments lead to a new strategy for systematically assessing the development of the tense and agreement system in pediatric AAC speakers and selecting developmentally appropriate intervention goals targeting expressive morphology skills in pediatric AAC speakers. Experiment 3 demonstrated this process in two pediatric AAC speakers with cerebral palsy and tested for cross-morpheme generalization of treatment effects. Experiment 3 demonstrated development and implementation of customized intervention programs based on assessment results.

Table 1.1. Purpose and research questions for Experiments 1-3

Purpose	Research Questions
The purpose of Experiment 1 is to investigate tense and agreement morpheme emergence and productivity in spontaneous language samples from typically developing 30-54 month olds learning English as a first language. This was the first study using one sampling method across age groups this full age range.	A. Does the number of tense markers children use in language samples with a fixed number of multi-morpheme utterances increase with age between 30 and 54 months? B. Does productivity score increase with age between 30 and 54 months in language samples with a fixed number of multi-morpheme utterances? C. Does morpheme category productivity increase at the same rate across morpheme categories?
The purpose of Experiment 2 was to characterize the effects of age and communication modality on tense marker stimulability and tense and agreement morpheme category stimulability in typically developing children acquiring English as a first language.	A. Do the odds of a child being stimulable for each tense marker increase at the same rate across communication modalities? B. For each morpheme category, does morpheme category stimulability increase at the same rate across communication modalities? C. Does morpheme category stimulability increase at the same rate across morpheme categories?
To use direct evidence of tense and agreement morpheme use and stimulability in pediatric AAC speakers with cerebral palsy to establish baseline levels of use and stimulability, ensure access to an SGD with potential to support a fully productive system, select specific goals for intervention, and measure ongoing change during intervention, and to compare patterns of cross-morpheme generalization to predictions made by OI theory, GML theory, and the hypothesis that tense marker stimulability is a prognostic indicator of generalization.	A. Does intervention focused on increased production of a target tense marker in pediatric AAC speakers generalize across tense markers? If so, does stimulability testing predict patterns of cross-morpheme generalization?

2.0 REVIEW OF THE LITERATURE

This chapter will present background information on the optional infinitive (OI) and gradual morphological learning (GML) theories, stimulability testing to assess learning potential, and prior research on the developmental morphology skills of pediatric AAC speakers. First, shared commonalities assumed by the OI and GML theories will be defined. Then, the central framework of each theory will be presented. Each framework will be presented in the context of developmental data from verbally speaking children with typical language development and specific language impairment. Next, a short section will compare the token and type-based measures used for testing the predictions of the OI and GML theories, respectively and present a rationale for using type-based measures in a project comparing these competing theories. This will be followed by a discussion of the role of developmental readiness in each framework. Finally, a summary of prior research on the morphology of AAC speakers will be presented.

2.1 COMMON ASSUMPTIONS

This section summarizes shared commonalities assumed by the OI and GML theories. Although there are critical differences between theories, there are several common assumptions that are central to both accounts of morpheme growth. First, both theories assume that children are equipped with an innate language acquisition device, and that the process of language acquisition is guided by a universal grammar (e.g., Chomsky, 1995). Specifically, both theories assume that structure dependence is innate, or that human languages depend on a hierarchical structure rather than a linear word order. Movement operations in a language with structure dependence, such as V → I movement of the copula in GML theory (Rispoli et al., 2012), subject raising in OI theory (Wexler, 1998), and subject-auxiliary inversion in both theories rely on this hierarchical structure. In addition, both theories hold that certain central components of the tense and agreement system are part of the innate universal grammar including grammatical features for tense and agreement.

Next, the OI and GML theories agree that there is a period of growth between the moment of emergence, where a morpheme appears for the first time in a child's expressive language output and the moment that a child achieves an adult level of mastery using that morpheme. Both theories predict that there will be gradual and continuous growth during this period. The child is not expected to acquire morphemes rapidly or in discrete stages. This period of growth is well-established, and was central to Brown's (1973) classic study on morphological development. Brown (1973) used a longitudinal corpus of naturalistic language samples to track the development of 14 grammatical morphemes in the spontaneous language of three typically developing children acquiring English. Brown (1973) tracked acquisition of each morpheme using a token-based measure of accuracy, which was operationally defined as the percentage of correct use in obligatory contexts as shown in (1).

1. Accuracy = (Correct Use Tokens / Obligatory Context Tokens) × 100%

Brown (1973) defined obligatory contexts as contexts with linguistic and situational constraints that would oblige a fluent adult speaker to use a given morpheme. These constraints were defined broadly in relation to conversational contexts, including utterances produced by a child’s communication partners. Brown’s (1973) accuracy measure requires accounts for both tokens of correct morpheme use in the child’s utterances and obligatory contexts broadly specified by conversational context. Brown (1973) used a criterion of mastery operationally defined as 90% correct use in obligatory contexts as an indication that any given morpheme was acquired. Brown (1973) found that all grammatical morphemes he studied developed gradually, rather than suddenly appearing at a high level of accuracy. Each morpheme initially emerged in a few contexts and then became more accurate over an extended growth period. Some morphemes did not meet Brown’s criterion of mastery for over a year after the point of initial emergence. However, the order in which these morphemes were “acquired” or met Brown’s criterion of mastery was highly consistent across children. These findings from Brown’s (1973) longitudinal data were replicated in a cross-sectional sample of typically developing children (de Villiers & de Villiers, 1973).

Finally, the OI and GML theories both hold that the English tense and agreement morphemes emerge simultaneously in a unified morphosyntactic system. An initial description of the English tense and agreement morphemes as a unified system was presented in Radford’s (1990) landmark text on the emergence of functional categories. Radford (1990) predicted that a group of functionally related morphemes which shared an underlying grammatical system would emerge and develop together as a morphosyntactic system. Radford (1990) described several morphosyntactic systems that exist in adult grammars but are not included in children’s earliest grammars, including a determiner system, an inflectional system, and a case system. Radford (1990) argued that children would have no compelling reason to use a morpheme in one of these systems until the system matures (Borer & Wexler, 1987) and the underlying grammatical features become available for use in the child’s grammar. At this time, all of the morphemes in a newly matured morphosyntactic system could potentially emerge. Once this happens, Radford (1990) expected to find direct evidence of this growing grammatical knowledge in the child's linguistic output. The direct evidence Radford (1990) described consisted of productive morpheme use across a diverse range of semantic contexts (e.g., using the same inflectional form to conjugate different verbs).

The inflectional system described by Radford (1990) includes a head-functional inflection constituent (I), which projects into the child's syntax. In this system, I carries the features [αTNS, αAGR], where [α] is a variable across features that must be either positive in both occurrences or negative in both occurrences. Inflectional morphemes located at I include infinitival *to*, which is not marked for tense or agreement [-TNS, -AGR] and several finite forms marked for both tense and agreement [+TNS, +AGR]; modals, finite verb inflections (*-3s*, *-ed*), AUXILIARY DO, COPULA BE, AUXILIARY BE, and AUXILIARY HAVE. Radford (1990) found no evidence of productive use for any of these morphemes in young children acquiring English until several morphemes emerged simultaneously at 24 months (± 4 months). This finding suggests that several diverse morphemes in a unified morphosyntactic system emerge simultaneously, but says nothing about the development of this system after the initial point of emergence (Rispoli & Hadley, 2011).

The OI and GML theories hypothesize that there is a functionally related tense and agreement morpheme system consisting of finite morphemes marked for tense and agreement. The morphosyntactic systems postulated by the OI and GML theories both consist of the same set of 15 tense marked forms (referred to throughout this document as tense markers). These tense markers include: past tense *–ed*, third person singular *-3s*, copula and auxiliary *am*, *is*, *are*, *was*, *were*, and auxiliary *do*, *does*, *did*.

Some researchers investigating the OI theory have argued that infinitival *to* (Norris, 2004; Wexler, 2011) and AUXILIARY HAVE (Harris & Wexler, 1996; Wexler, 2011) are part of this same morphosyntactic system and subject to the same processes of developmental growth. However, these tense markers have not been accounted for in studies investigating optional infinitives in children with developmental language disorders (e.g., Rice et al., 1995; Rice et al., 1998) and are not included in standardized assessments of expressive morphology based on the OI theory (Rice & Wexler, 2001). This paper acknowledges the possibility that infinitival *to* and AUXILIARY HAVE are part of a larger tense and agreement system, but focuses on the 15 tense markers that have been accounted for by both theoretical frameworks.

2.2 THE OPTIONAL INFINITIVE THEORY

The OI theory assumes a maturational model of grammatical development with an innate genetic program for a universal grammar guiding grammatical growth, as initially proposed by Borer and Wexler (1987, 1992). In maturational systems, some aspects of universal grammar grow over time, with a growth pattern guided by an underlying genetic program (Borer & Wexler, 1987). OI theory specifies that the aspects of universal grammar guiding growth of the tense and agreement system are guided by a genetic program and a genetically determined timeline (Rice et al., 1998; Wexler, 1994, 1998, 2011).

The central hypothesis of the OI theory is that children acquiring language experience an optional infinitive stage of development, in which they use finite and root infinitive verb forms interchangeably in contexts that would be obligatory contexts for finite verb forms marked for tense and agreement in adult grammars (Wexler, 1994, 1998, 2011). Tense marking is obligatory in root declarative clauses for adult speakers of all languages that mark tense on the verb, including English (Wexler, 2011). Children acquiring these languages frequently omit tense from verbs in root clauses and produce infinitives instead of tense marked verbs. These sentences are common in child language and ungrammatical in adult language. As a child grows through the optional infinitive stage, the proportion of correctly marked finite verb forms he produces in obligatory contexts gradually and progressively grows. Wexler's (2011, p. 57) revised characteristics of the optional infinitive stage are shown in (2):

2. The properties of the OI stage are the following:
    a. Root infinitives (non-finite verbs) are possible grammatical sentences for children in this stage
    b. These infinitives co-exist with finite forms
    c. The children nevertheless know the relevant grammatical principles and have set the relevant parameters early.

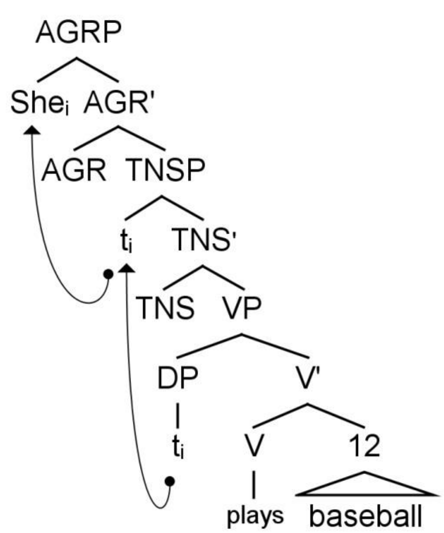
According to Wexler (1998), properties of many of the important inflectional elements are present at the earliest observable stages, when children are in the early two-word stage of development. These properties include basic grammatical parameters and grammatical and phonological properties of inflectional elements that must be learned from language input using perceptual learning (Wexler, 1998). If grammatical learning occurs before children produce linguistic output, then children do not rely on negative evidence or corrective feedback for grammatical learning. The child may continue to learn lexical items throughout his lifetime, but much grammatical learning occurs very early in development. This very early learning does not explain developmental changes in children's grammars over time, well beyond the early two-word stage. Wexler's solution follows a maturational model (Borer & Wexler, 1987, 1992) and claims that “a certain amount of inflectional development unfolds over time according to a genetic blueprint” (Wexler, 1998, p. 85).

Wexler (1998) proposed a unique checking constraint, which exists in young children’s grammars but not adult grammars and argued that this developmental constraint can explain universal grammar-guided growth of finiteness in the OI stage with a single biological mechanism. The unique checking constraint is a developmental constraint on the computational system of language that fades over time, and has been compared to parallel biological constraints that do not allow young children to walk (Wexler, 2003). The unique checking constraint restricts production of many verb inflections across languages that cannot be inserted into a node unless both [+ TNS] and [+ AGR] are present (Wexler, 2003).

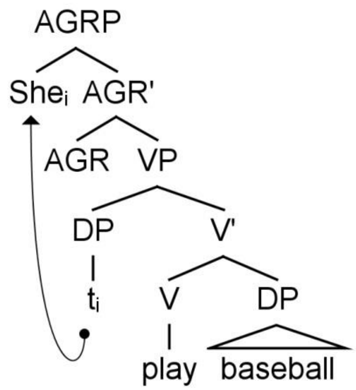
The unique checking constraint (Wexler, 1998) builds on assumptions of Chomsky’s (1995) Extended Projection Principle (Wexler, 2011). First, all verb arguments are generated within a verb phrase. The subject determiner phrase is an argument generated in the specifier of the verb phrase. Second, Wexler (1998) assumes that tense (TNS) and agreement (AGR) are distinct abstract functional categories generated outside of the verb phrase, rather than the single functional category (I) proposed by Radford (1990). TNS and AGR each have an uninterpretable determiner feature (D-feature) that must be checked by an interpretable D-feature. If one of these uninterpretable D-features is not checked, then the corresponding functional category (TNS or AGR) is deleted. Third, children adhere to rules of distributed morphology and the elsewhere principle (Halle & Marantz, 1993). Morphemes compete for insertion at nodes. The maximally specified morpheme with no extraneous features that are not present on the node will be inserted. In the event that a TNS or AGR category is deleted, a root infinitive is inserted as the maximally specified morpheme because tense markers include an extraneous feature that is not present on the node. Following the extended projection principle (Chomsky, 1995), the subject determiner phrase is forced to raise out of the verb phrase in order to check the D-features on TNS and AGR (Wexler, 1998).

Wexler (1998, 2011) argues that double-checking is permitted in adult grammars, so the subject determiner phrase checks the D-features on both functional categories (TNS and AGR) as it is raised, allowing adults to correctly produce tense and agreement marking in all cases. An example of a sentence with double-checking is shown in (3a). In contrast, children have a unique checking constraint in their initial grammar that only allows a determiner phrase to check the D-feature of one functional category. Any time the unique checking constraint is applied, a root infinitive will be inserted (Wexler, 1998). Examples (3b) and (3c) show instances where the unique checking constraint is applied and TNS and AGR are omitted, respectively. Evidence of English-speaking children producing optional infinitives in sentences with nominative subjects (3b) and sentences with non-nominative subjects (3c) is interpreted as evidence that TNS and AGR are separate functional categories and that either of these functional categories can be omitted (Schütze & Wexler, 1996).

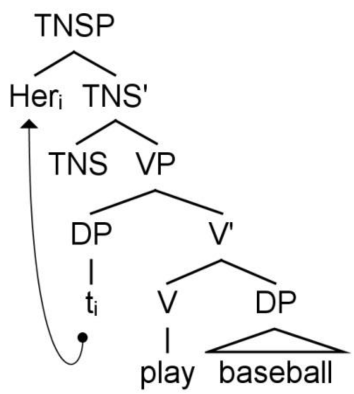
3. a.



b.



c.



The unique checking constraint gradually withers on a genetically defined time-line. Initially, the unique checking constraint is broadly applied and a vast majority of the child's utterances are infinitive. Finiteness rates are expected to tend towards zero in younger populations (Wexler, 1994). At least one team of researchers has used language samples from the single-word stage to argue that children produce 100% infinitive verbs at the very onset of verb production (Blom & Wijnen, 1999). As the unique checking constraint gradually withers, it is applied to a progressively smaller proportion of utterances and the child correctly produces a progressively larger proportion of finite forms until adult levels of accuracy are reached.

Wexler (1994) predicted that the optional infinitive stage in children acquiring English would be evident in a tendency to omit several tense-marked morphemes, including *-3s*, *-ed*, AUXILIARY DO, COPULA BE, and AUXILIARY BE. Overt evidence of optional infinitives has been reported for several languages that form verbs by affixing a specific non-finite inflection onto a verb stem (e.g., Wexler, 1994). Overt evidence of optional infinitives is difficult to observe in English because the surface form of the English infinitive is not differentiated from the bare stem form (Wexler, 1994). For example, an English speaking child in the optional infinitive stage would produce the positive forms (4a-4b), where (4a) is finite and (4b) is infinitive. The same child also would produce the infinitive negative form (4c), which clearly contains no evidence of finite tense marking because medial negation is present and *do*-support is not supplied (Wexler, 1994). The child may use *do*-support in production of finite negatives like (4d), but some children do not produce these sentences until they are older (Wexler, 1994). The child would be expected to set parameters before entering the optional infinitive stage (Wexler, 1998), and would not produce sentences like (4e) that violate known parameters. In contrast, an adult English speaker would only produce the finite forms (5a) and (5d).

4. a. Mary plays baseball
   b. Mary play baseball
   c. Mary not play baseball
   d. ?Mary does not play baseball
   e. *Mary not plays baseball
5. a. Mary plays baseball
   b. *Mary play baseball
   c. *Mary not play baseball
   d. Mary does not play baseball

e. *Mary not plays baseball

Young children's infinitive forms are not strictly limited to forms that are traditionally considered infinitives in adult language (Wexler, 1994, 1998, 2003), such as English forms with infinitival *to*. Norris (2004) and Wexler (2011) argue that infinitival *to* is subject to the unique checking constraint, and would not be inserted into main clauses produced with optional infinitives, as in (6).

6. #Mary to eat candy

Children's grammars may include a variety of infinitive or non-finite forms characterized by an absence of tense marking. In English sentences with progressive aspect, verbs are marked with an *–ing* suffix and finiteness are only marked by a form of AUXILIARY BE (7a-7b). Wexler (1994) suggests that children's early productions of progressive verbs in the absence of AUXILIARY BE (7c) are optional infinitive forms produced when the unique checking constraint is applied and the auxiliary is dropped.

7. a. Mary is going
   b. Mary was going
   c. Mary going

The primary metrics used to measure growth of the tense and agreement system in research investigating optional infinitives are token-based measures similar to Brown's (1973) morpheme-specific measure of accurate use in obligatory contexts defined in (1). These measures are typically reported as composite measures of accuracy or optional infinitive use across all tense markers. Unlike Brown (1973), researchers investigating optional infinitives consider productions of both tense-marked verbs and optional infinitives accurate within the constraints of the child's current grammar. The "error" term is a measure that captures the percentage of utterances containing optional infinitives instead of a finite tense marker. Computational formulas for calculating the percentage of accurate finite clauses and the percentage of optional infinitives are shown in (8a-8b).

8. a. Accuracy = (Finite Clauses / Clauses That Should Be Finite) × 100%

   b. Optional Infinitives = (Infinitives That Should Be Finite / Clauses That Should Be Finite) × 100%

The OI theory predicts that a child's overall accuracy of finite tense and agreement morpheme production will gradually and continuously increase until the child out-grows the optional infinitive stage. This should be evident in a gradual and continuous increase in accurate production of finite clauses over time or a gradual and continuous decrease in percentage of optional infinitives over time.

Evidence of a gradual and continuous increase in accurate production of finite clauses is found in a longitudinal study by Rice et al. (1998). Rice et al. (1998) studied acquisition of the tense and agreement system in 5-year olds with specific language impairment, an age-matched group of 5-year olds with typical language development and an MLU-matched group of 3-year olds with typical language development. All participants were sampled every 6 months for seven intervals. The combined results from the two groups of typically developing children encompass a developmental period spanning from 3 to 8 years. Rice et al's (1998) composite measure of tense marker accuracy included spontaneous and elicited productions of *-3s*, *-ed*, COPULA and AUXILIARY BE and elicited productions of AUXILIARY DO. Between the ages of 3 and 4, the composite measure of tense marker accuracy increased from 50% to 90+%, exceeding Brown's (1973) criterion of mastery.

A steady and continuous reduction in overall percentage of optional infinitives was found in a cross-sectional sample of typically developing children learning Dutch (Wexler, Schaeffer, & Bol, 2004). In these children, percentage of optional infinitives decreased from 83% in 19-24 month olds to only 7% in 37-43 month olds.

Rice et al. (1998, p. 1417) make a strong prediction about the growth of the various morphemes in the tense and agreement system as children progress through the optional infinitive stage: "Growth curves for individual morphemes should be highly similar to each other and to a composite measure." Their longitudinal data showed evidence of growth for accurate use of each finite morpheme in spontaneous or elicited productions in children with typical language development between 3 and 4 years of age. A period of growth is evident for each tense marked measure, but details of growth trajectories are unclear because samples were only collected once every six months. A different growth trajectory was found for a control morpheme (plural -*s*), which was consistently used accurately in over 90% of obligatory contexts by typically developing children.

Delayed development of the tense and agreement system is often considered a clinical marker for children with specific language impairment. Rice and colleagues (Rice, 2003; Rice & Wexler, 1996, 2001; Rice et al., 1995; Rice et al., 1998) proposed a hypothesis that children with specific language impairment experience an extended optional infinitive stage, which some children may not ever fully emerge from. These children may never achieve adult levels of accuracy for finite tense marking. Longitudinal data on English-speaking 5-8 year olds with specific language impairment have lower levels of accuracy than MLU-matched peers for tense marked morphemes and high levels of accuracy for a plural –*s* morpheme that does not require tense marking (Rice et al., 1998). These children still produced some optional infinitives at age

8 (Rice et al., 1998). At age 5, these same children showed high levels of accuracy for production of several unrelated morphemes that did not require tense marking (Rice & Wexler, 1996). High levels of accuracy producing unrelated morphemes also were found in groups of age-matched controls and MLU-matched controls (Rice & Wexler, 1996). Similar evidence of an extended optional infinitive stage has been reported in cross-sectional samples of children learning English (Rice & Wexler, 2001) and Dutch (Wexler et al., 2004). This evidence from children with specific language impairment indicate that children can have selective difficulty acquiring morphemes in a unified tense and agreement system and that these morphemes should be accounted for during assessment.

In the maturational system postulated by OI theory, input-driven learning does not contribute to tense and agreement system growth after parameters are set. If the subsequent development of this system is purely maturational, then broad-targeted intervention programs focused on increasing accurate use of tense marking in general would fail to produce a treatment effect. More specific intervention programs focused on increasing production of specific tense markers would fail to produce a treatment effect for any given tense marker or generalize across tense markers. Increased use of the specific tokens practiced in intervention would be possible if those tokens are stored in memory as lexical items.

In a maturational system, intervention focused on teaching a child to produce tense markers with an AAC system is not expected to further development of the child's tense and agreement system. This intervention program can teach a child to use tense markers that have already developed maturationally in a new communication modality. However, the potential benefits of this intervention at any point in time will be subject to maturational constraints, and some children may use optional infinitives on their AAC systems.

2.3 THE GRADUAL MORPHOSYNTACTIC LEARNING THEORY

The GML theory (Rispoli & Hadley, 2011) holds that the diverse set of structures in the English tense and agreement system is an abstract system unified at the level of grammatical features (Bock & Levelt, 1994), which develops gradually in a relatively uniform partial sequence, and is learned within the constraints of an initial hypothesis space defined by universal grammar (Hadley, Rispoli, Fitzgerald, & Bahnsen, 2011). The GML theory makes an empirical prediction that tense and agreement morphemes will develop in partial sequence with growth rates that vary across morphemes. Differences in growth rates across tense and agreement morphemes arise because different tense and agreement morphemes are processed differently during grammatical encoding (Bock & Levelt, 1994; Rispoli & Hadley, 2011).

In the Bock-Levelt model of language production (Bock & Levelt, 1994), grammatical encoding happens incrementally across two levels of processing. First, grammatical features of an intended message are generated at the level of functional processing through the processes of lexical selection and function assignment. The process of lexical selection involves identifying lexical concepts and lemmas carrying grammatical information associated with individual lexical concepts. The process of function assignment assigns grammatical functions to the selected lemmas. Positional processing specifies a set of procedures necessary for assembling a hierarchically organized message frame with slots for selected message constituents at the level of positional processing. Syntactic and morphological contexts related to the order of constituents in this message frame are generated at the level of positional processing.

The GML theory predicts that "morpheme categories will increase in productivity together, provided that they are similar enough in positional processing" (Rispoli et al., 2012, p. 1009). Learning procedures for assembling a frame at the level of positional processing can contribute to learning other procedures for assembling similar frames. Conversely, morpheme categories should increase in productivity at different rates if they are not similar enough in positional processing because learning requirements for these procedures differ to a greater degree.

The distribution of functional and positional processing features for all 15 tense markers and the stem forms of verbs are shown in Figure 2.1 (adapted from Rispoli et al., 2012). The forms fall in complementary distribution: any given combination of functional and positional processing features will only map to one form. These forms are unified at the level of functional processing by a shared set of tense and agreement features. Productive use of this set of tense markers can emerge simultaneously as procedures for using grammatical features associated with tense and agreement are learned. In the literature, the tense markers are typically grouped into five morpheme categories, as shown in Figure 2.2. The tense marker categories are primarily differentiated at the level of positional processing by syntactic and morphological context: COPULA BE, AUXILIARY BE, AUXILIARY DO, and the THIRD PERSON SINGULAR -3S and PAST TENSE –ED verb affixes. The COPULA and AUXILIARY BE categories each contain five tense markers: *am*, *is*, *are*, *was*, *were*. The AUXILIARY DO category contains three tense markers: *do*, *does*, *did*. The THIRD PERSON SINGULAR -3S and PAST TENSE –ED only contain one tense marker each (*-3s*, *-ed*). For consistency, the terms *-3s* and *–ed* are used in reference to these two categories and the corresponding tense markers throughout this dissertation. In this category structure, there are only two cases where categories are differentiated at the level of functional processing: *-3s* and *-ed* differ in tense, and AUXILIARY BE differs from all other categories because it includes an additional feature.

The COPULA BE category is uniquely specified in positional processing as a frame that lacks a lexical verb [- Verb] (Rispoli et al., 2012) following hypotheses in generative linguistic theory that COPULA BE undergoes V → I movement (Pollock, 1989). Tense markers in the COPULA BE category are treated as a closed class of morphemes marking tense and agreement features that do not behave as lexical verbs. All four remaining categories have a [+ Verb] feature because they demand a lexical verb at the level of positional processing.

The AUXILIARY BE category is demanded by progressive aspect [+ Prog] in functional processing, along with an *–ing* suffix on the lexical verb stem (Rispoli et al., 2012). The addition of a grammatical feature for progressive aspect makes this the most complex morpheme category in functional processing because it is the only morpheme category that requires both an auxiliary form and an affix on the lexical verb (LaPointe & Dell, 1989; Rispoli et al., 2012).

The individual tense markers in the COPULA BE and AUXILIARY BE categories are each uniquely specified by grammatical features at the level of functional processing (Rispoli et al., 2012). The present tense [- Past] forms *am*, *is*, and *are* are specified for first person singular, third person singular, and other contexts of agreement, respectively. The past tense [+ Past] forms *was*, *were* are specified for first and third person singular agreement, and other contexts of agreement, respectively.

The *-ed*, *-3s* and AUXILIARY DO categories all have a nonspecific aspect [- Prog] in functional processing that does not require an additional grammatical feature (LaPointe & Dell, 1989; Rispoli et al., 2012). Lexical verbs with nonspecific aspect are differentiated by tense and agreement features at the level of functional processing in a tripartite paradigm: *-3s*, *-ed*, and the *stem* form (Rispoli et al., 2012). *-3s* and *–ed* are [- Aux] and directly attached to verb stem as a suffix in contexts that do not require a tense marked auxiliary. AUXILIARY DO is a surrogate

<table>
<tr><th>Tense[a]</th><th>Number[a]</th><th>Person[a]</th><th>-V[b]
-Aux[b]
-Prog[a]</th><th>+V
-Aux
-Prog</th><th>+v
+Aux
-Prog</th><th>+V
+Aux
+Prog</th></tr>
<tr><td>-Past</td><td>Singular</td><td>3rd</td><td>cop is</td><td>-3s</td><td>aux does</td><td>aux is</td></tr>
<tr><td></td><td></td><td>1st</td><td>cop am</td><td rowspan="5">Stem</td><td rowspan="5">aux do</td><td>aux am</td></tr>
<tr><td></td><td></td><td>2nd</td><td rowspan="4">cop are</td><td rowspan="4">aux are</td></tr>
<tr><td></td><td>Plural</td><td>3rd</td></tr>
<tr><td></td><td></td><td>1st</td></tr>
<tr><td></td><td></td><td>2nd</td></tr>
<tr><td>+Past</td><td>Singular</td><td>3rd</td><td rowspan="2">cop was</td><td rowspan="6">-ed</td><td rowspan="6">aux did</td><td rowspan="2">aux was</td></tr>
<tr><td></td><td></td><td>1st</td></tr>
<tr><td></td><td></td><td>2nd</td><td rowspan="4">cop were</td><td rowspan="4">aux were</td></tr>
<tr><td></td><td>Plural</td><td>3rd</td></tr>
<tr><td></td><td></td><td>1st</td></tr>
<tr><td></td><td></td><td>2nd</td></tr>
</table>

Figure 2.1. Complementary distribution of English tense markers

Note. [a]Functional processing feature. [b]Positional processing feature.

<table>
<tr><th>Tense[a]</th><th>Number[a]</th><th>Person[a]</th><th>-V[b]
-Aux[b]
-Prog[a]</th><th>+V
-Aux
-Prog</th><th>+v
+Aux
-Prog</th><th>+V
+Aux
+Prog</th></tr>
<tr><td>-Past</td><td>Singular</td><td>3rd</td><td rowspan="12">COPULA
BE</td><td>*-3s*</td><td rowspan="12">AUXILIARY
DO</td><td rowspan="12">AUXILIARY
BE</td></tr>
<tr><td></td><td></td><td>1st</td><td rowspan="5"></td></tr>
<tr><td></td><td></td><td>2nd</td></tr>
<tr><td></td><td>Plural</td><td>3rd</td></tr>
<tr><td></td><td></td><td>1st</td></tr>
<tr><td></td><td></td><td>2nd</td></tr>
<tr><td>+Past</td><td>Singular</td><td>3rd</td><td rowspan="6">*-ed*</td></tr>
<tr><td></td><td></td><td>1st</td></tr>
<tr><td></td><td></td><td>2nd</td></tr>
<tr><td></td><td>Plural</td><td>3rd</td></tr>
<tr><td></td><td></td><td>1st</td></tr>
<tr><td></td><td></td><td>2nd</td></tr>
</table>

Figure 2.2. Complementary distribution of tense and agreement morpheme categories

Note. [a]Functional processing feature. [b]Positional processing feature.

auxiliary category [+ Aux] necessitated at the level of positional processing by syntactic contexts that require *do*-support for tense and agreement marking (Rispoli et al., 2012). The three tense markers in the AUXILIARY DO category (*does*, *did*, *do*) have a tripartite paradigm differentiated at the level of functional processing in the same way that *-3s*, *-ed*, *stem* are differentiated for nonspecific lexical verbs (Rispoli et al., 2012).

The predictions of the GML theory have been studied in a series of longitudinal studies investigating the growth of the tense and agreement system in a cohort of 20 typically developing toddlers acquiring English as a first language (Hadley et al., 2011; Rispoli, Hadley, & Holt, 2009; Rispoli et al., 2012). Play-based language samples were collected from each child during dyadic interaction with a primary caregiver at 21, 24, 27, 30, and 33 months. Each sample was collected in two 1-hour sessions spaced no more than two weeks apart. Growth of the tense and agreement system was characterized for individual children and at the cohort level.

The primary metrics used to measure growth of the tense and agreement system across this set of studies were measures of spontaneous productivity defined by Hadley and Short (2005) and clarified with examples by Rispoli and Hadley (2011). These measures were adapted from Radford's (1990) type-based method of documenting the emergence of morpheme productivity to capture the growth of a productive morphosyntactic system beyond the point of initial emergence. Five spontaneous category productivity scores were found to characterize the productivity of each morpheme category in the children's spontaneous language samples. In each category, productivity was measured by counting the number of sufficiently different uses of corresponding tense markers (see Appendix A for details). A composite productivity score was found by adding the five different category productivity scores. The composite productivity score correlates with a token-based composite measure of accuracy used to track progress towards mastery of these same morphemes (Hadley & Short, 2005).

Type-based measures of productivity are used to characterize tense and agreement morpheme use in a language sample because they reflect distributed use of grammatical encoding to produce morphemes in a diverse range of semantic contexts. The GML theory holds that language can be produced through either direct activation of associative connections between referential and phonetic content (Bock, 1982) or through grammatical encoding processes (following Bock & Levelt, 1994) that require sentence formulation (Rispoli & Hadley, 2011; Rispoli et al., 2009, 2012). Both generative and usage-based theories of language recognize that high-frequency strings of words and morphemes may be stored in the lexicon as whole units and directly activated as if they were a single morpheme (e.g., Bybee, 2006; Pinker, 1999; Pinker & Prince, 1994; Ullman, 2001). As an example, Radford (1990) shows that adults store high-frequency phrases from foreign languages in memory as single words without realizing that the phrase contains several distinct morphemes. Phrases retrieved via direct activation do not require morphological representation. In contrast, language produced using grammatical encoding processes requires incremental planning and use of morphosyntactic representations (Rispoli & Hadley, 2011; Rispoli et al., 2012). Token-based measures of accuracy may be systematically inflated when children repeatedly use direct activation to retrieve multi-morphemic units containing high-frequency tense-marked forms (Rispoli & Hadley, 2011; Rispoli et al., 2009, 2012). Type-based measures of productivity avoid this systematic inflation.

Evidence of gradual growth in productivity is reported in Rispoli et al. (2009) for the main longitudinal cohort of 20 toddlers and in a replication study with a second longitudinal cohort of 37 toddlers (Hadley, Rispoli, Holt, Fitzgerald, & Bahnsen, 2014). Rispoli et al. (2009)

and Hadley et al. (2014) used hierarchical linear modeling to model longitudinal growth of composite productivity scores in the two independent cohorts. No significant differences in growth models were found models between cohorts. Productivity scores were not significantly higher than zero at 21 months, so both growth models assumed a zero-intercept with no tense marker use at 21 months. Average growth in both cohorts was characterized by instantaneous linear growth of less than 1 productive morpheme per month beginning at 21 months, with accelerating quadratic growth. Significant individual variation was found in both cohorts.

Analyses investigating relationships between individual tense markers in different categories suggest unification of tense markers at the level of functional processing. Rispoli et al. (2012) tested correlations between spontaneous productivity of copula *is* and spontaneous productivity of other tense markers sharing the [+ Agr3s, - Past] feature bundle (auxiliary *does*, -*3s*, auxiliary *is*) and a control tense marker with few shared features (-*ed*). Productivity of copula *is* was significantly correlated with productivity of auxiliary *does* and -*3s*, but not -*ed* or auxiliary *is*. Correlations between tense markers with similar features in functional processing did not extend to the auxiliary form with an added [+ Prog] feature.

Rispoli et al. (2012) investigated the hypothesis that morpheme categories with similar positional processing would increase in productivity together by analyzing the spontaneous category productivity scores of all categories at each time point. All categories had essentially zero productivity at 21 months and became more productive over time. Significant contrasts between spontaneous category productivity scores emerged at 27 months and persisted at 30 and 33 months. COPULA BE was significantly more productive than all four [+ Verb] categories. The three [+ Verb] categories with nonprogressive aspect (AUXILIARY DO, -*ed* and -*3s*) grew together in productivity and were significantly more productive than AUXILIARY BE. [+ Verb] categories with similar positional processing grew in productivity together. However, the AUXILIARY BE category with greater complexity at the level of functional processing did not grow in productivity at the same rate as the other [+ Verb] categories.

Gladfelter and Leonard (2013) found different patterns of spontaneous category productivity scores in cross-sectional samples of 4 and 5 year olds with typical language development and specific language impairment. No significant differences were found between age groups. However, significant differences in overall productivity and different patterns of category productivity were found between diagnostic groups. In the children with specific language impairment, productivity of the COPULA BE category was significantly higher than all other categories and productivity of the *-3s* category was significantly higher than the AUXILIARY BE category. This indicates that the children with specific language impairment developed less productivity in categories with [+ Verb] contexts in positional processing than their typically developing peers.

In Gladfelter and Leonard's (2013) typically developing children, category productivity scores were highest for the COPULA BE category, followed by the AUXILIARY DO and *-3s* categories. The category productivity scores of the AUXILIARY DO and *-3s* categories were significantly higher than the category productivity scores of the *–ed* category. The productivity of the *–ed* category was not significantly higher than the productivity of the AUXILIARY BE category. The finding that *–ed* category productivity patterns with AUXILIARY BE category productivity instead of *-3s* category productivity in 4 year olds cannot be explained by Risploi et al.'s (2012) prediction that tense marker categories with similar positional processing features will grow in productivity together.

One possible explanation relates to differences in language sampling protocols across studies. Rispoli et al. (2012) used a fixed time interval to collect conversational samples of parent-child dyads with as little examiner participation as possible. The number of child utterances in Rispoli et al.'s (2012) toddler samples were highly variable, ranging from a minimum of 147 utterances at 21 months to 881 utterances at 27 months. In contrast, Gladfelter and Leonard (2013) collected samples of child-examiner dyads and permitted examiners to occasionally ask specific questions to elicit additional, spontaneous utterances. Gladfelter and Leonard (2013) selected a total of 152 spontaneous utterances (the smallest sample generated) from each participant to control the total number of child utterances in each sample. Sample size was controlled to avoid potential ceiling effects in measures influenced by each instance of a scorable morpheme. Low productivity of *–ed* may be related to Gladfelter and Leonard's (2013) small sample size or examiner-administered elicitation prompts biased against past tense forms. These possibilities cannot be ruled out unless language samples are collected across age groups using one sampling method that allows for direct comparisons.

Finally, the GML theory holds that the tense and agreement system and morphosyntactic growth in general are gradually learned within the constraints of an initial hypothesis space defined by universal grammar (Hadley et al., 2011). Hadley et al. (2011) acknowledge a prominent role of input and statistical learning in morphosyntactic growth that is similar to the variational learning mechanism described by Legate and Yang (2007). The learning mechanism described by Hadley et al. (2011) is innately guided and specific to the language domain, while the variational learning mechanism described by Legate and Yang (2007) is domain general. In this learning mechanism, the parameters of children's grammars are gradually adjusted based on input. The gradual learning process of the GML theory differs from the relatively rapid learning process of the OI theory, which assumes that relevant parameters are set at the earliest observable point in time.

The GML theory assumes that children's initial hypothesis space is constrained by an innate universal grammar, and that children have some innate knowledge for organizing language input (Hadley et al., 2011). Specifically, the GML theory assumes that children are equipped with distinctions between predicates, arguments, and adjuncts (Van Valin, as cited in Hadley et al., 2011), which guides the learning of clausal structure (Hadley et al., 2011). The GML theory assumes that learning of phrasal and clausal structure is guided by the principle of structure dependence (Crain & Nakayama, 1987), including the requirement that "tense must have scope over a constituent" (Hadley et al., 2011, p. 551). In this framework, knowledge of what counts as a clause and knowledge of what does not count as a clause guide children as they learn what scope the tense morpheme has in their language (Hadley et al., 2011). This knowledge of clausal structure and structure dependency allows children to learn from distributed evidence of morphemes that appear in complimentary distributions across diverse syntactic contexts (Hadley et al., 2011).

Hadley et al.'s (2011) innately guided learning mechanism is described as a competition between binary parameter values, such as a competition between a [+ Tense] grammar with obligatory tense marking and a [- Tense] grammar without tense marking. In this system, learning to set a parameter is an iterative process of rewarding and punishing competing grammars as the child processes relevant evidence from the input stream. In the initial hypothesis space, there is an equal probability that the child will select either of the competing grammars to analyze an input sentence. For each input sentence, the learner selects a grammar and analyzes the sentence with the selected grammar. If the analysis is successful, the probability of selecting the same grammar again increases and the probability of selecting the competing grammar decreases. If the analysis is unsuccessful, the probability of selecting the same grammar again decreases and the probability of selecting the competing grammar increases. This iterative process continues until the correct grammar is selected consistently and the competing grammar is driven to extinction.

In this system, the child gradually learns from input to configure the parameters of an adult grammar. In a longitudinal study of vocabulary growth in 42 American children, Hart and Risley (1995) presented compelling evidence that the quantity and quality of day-to-day parental input during the first three years of life were strong predictors of long-term vocabulary growth and language development. Hart and Risley (1995, p. 95) recognized that "not all talk is equally informative" and sought to define and quantify features of parent behaviors that added informative quality to the input stream and encouraged learning. In this process, they considered both net amounts of input and proportional experiences with different types of input relative to the total number of utterances in the input stream.

Legate and Yang (2007) conducted a cross-linguistic analysis to explore the possibility that proportional input informativeness can influence the acquisition of a morphosyntactic system with a [+ Tense] grammar. They argued that input providing unambiguous evidence of tense marking, such as *-3s* would consistently reward a [+ Tense] grammar while input providing ambiguous evidence of tense marking, such as bare stem verbs that could be interpreted as infinitives, would consistently reward whatever grammar was selected for analysis. If this were the case, they predicted that children exposed to substantially more informative (unambiguous) evidence of tense marking than uninformative (ambiguous) evidence of tense marking would outgrow the optional infinitive stage faster than children exposed to relatively similar proportions of unambiguous and ambiguous evidence of tense marking. They found that children learning Spanish were exposed to more unambiguous evidence of tense marking than children learning French or English and outgrew the optional infinitive stage at approximately 24 months. In contrast, children learning English were exposed to almost as much ambiguous evidence of tense marking as unambiguous evidence of tense marking and still continued to produce optional infinitives at 41 months.

Hadley et al. (2011) studied the contribution of input informativeness on the growth of productivity scores in their cohort of children acquiring English. They found that the proportion of unambiguous tense marking in parental input at 21 months was positively related to early linear growth of productivity and predicted productivity scores at 30 months. They also found that the frequency of ambiguous tense marking in parental input that could potentially reward a [- Tense] grammar was related to slower growth in productivity scores over time. A 4-variable model that accounted for both the proportion and frequency of unambiguous and ambiguous tense marking in parental input at 21 months accounted for 78.7% of the total variance in productivity scores at 30 months, indicating that both net and proportional input contributed to learning. This evidence is consistent with the prediction that the [+ Tense] parameter is gradually configured in an iterative manner as the child processes evidence of tense marking in the input stream. This suggests that providing unambiguous evidence of tense marking is useful for teaching a [+ Tense] grammar in an intervention program.

Recent studies have explored potential intervention programs that increase children's environmental exposure to unambiguous evidence of tense marking. Hadley and Walsh (2014) found that parents used significantly more third person subjects, lexical noun phrase subjects, and unambiguous third person present tense singular tense markers when they were trained to have conversations with their children using a “toy talk” strategy during play time. The toy talk strategy consisted of (a) talking about the toys the child is currently playing with by commenting on the states, actions, and properties of the toys and (b) naming the item instead of referring to the item with a pronoun. This strategy increased the frequency and diversity of informative tense markers in children’s language environments, and may be a viable intervention strategy for supporting overall development of the tense and agreement system.

Some preliminary evidence exists to indicate that measures of productivity can differentiate between children with specific language impairment and children with typical language development (Gladfelter & Leonard, 2013; Hadley & Short, 2005). Hadley and Short (2005) found that 2 year olds with language abilities in the low end of the average range had higher productivity scores than 2 year olds at risk for specific language impairment. In addition, they found that low average children had higher tense marker totals than children at risk for specific language impairment, demonstrating initial use of a wider variety of tense markers than the children at risk for specific language impairment. Based on these dual findings, Hadley and Short (2005) recommended using both of these measures to characterize the breadth and depth of tense marker use and identify children with specific language impairment based on a two-dimensional metric.

Gladfelter and Leonard (2013) found that 4 and 5 year old children with specific language impairment had significantly lower productivity than age-matched peers with typical language development. They also found that children with specific language impairment had significantly lower category productivity for the *-3s*, AUXILIARY DO, and AUXILIARY BE categories but not for the COPULA BE and *-ed* categories, suggesting that productivity may be more limited for some tense markers than others. If a pediatric AAC speaker has similar difficulties with tense marking, limitations may vary across tense markers in similar patterns. For example, a pediatric AAC speaker may use COPULA BE productively, while the four lexical verb form categories are still emergent. Comprehensive assessment can identify individual strengths and limitations within the tense and agreement system.

In the system postulated by the GML theory, probabilistic learning informs tense and agreement system growth. Some evidence indicates that children learn functional processing features as they learn individual tense markers. Specifically, early development of one tense marker may facilitate later development of other tense markers with shared functional processing features. Rispoli (2016) used hierarchical multiple regression analysis to explore the relationship between early productivity of copula *is* later productivity of the *-3s* and *–ed* categories. He found that copula *is* productivity at 24 months explained a significant unique portion of variance in *-3s* productivity at 27 months after variance explained by MLU at 24 months was accounted for. In contrast, copula *is* productivity at 24 months did not uniquely explain any of the variance in *–ed* productivity at 27 months after variance explained by MLU at 24 months was accounted for. This indicated that learning the grammatical features of copula *is* facilitated later development of verb inflections with shared functional processing features (*-3s*) but not verb inflections with different functional processing features.

In such a system, treatment targeting tense and agreement system growth could be streamlined by targeting a subset of tense markers and capitalizing on patterns of generalization. Cross-morpheme generalization across tense markers with similar features has been reported in at least two treatment studies (Leonard, Camarata, Brown, & Camarata, 2004; Leonard, Camarata, Pawtowska, Brown, & Camarata, 2006). Leonard et al. (2004) studied cross-morpheme generalization in two groups of children with specific language impairment. One group received intervention targeting *-3s*. The other group received intervention targeting auxiliary *is*, *are* and *was*. In the *-3s* group, treatment effects generalized to improved production of non-target auxiliary forms (*is*, *are*, *was*). In the AUXILIARY group, treatment effects generalized to improved production of *-3s*. Cross-morpheme generalization to *–ed* or a control morpheme (either infinitival *to* or nonthematic *of*) was not found for either group. This same pattern was replicated in a follow-up study that accounted for maturation effects over an extended intervention period (Leonard et al., 2006).

If cross-morpheme generalization patterns with grammatical features, then intervention programs focused on improving tense and agreement morpheme use could potentially enhance tense and agreement system development in predictable ways by building knowledge of grammatical features that are shared across tense markers. In this case, intervention teaching children to produce tense markers on an AAC system could potentially enhance grammatical development and increase knowledge of non-target tense markers with shared grammatical features. This could potentially lead to observable patterns of generalization to non-target tense markers that share grammatical features with target tense markers in pediatric AAC speakers as Leonard et al. (2004; 2006) reported in verbally speaking children with specific language impairment.

2.4 TOKEN AND TYPE-BASED MEASURES

The token-based measures of accuracy and the type-based measures of productivity that researchers investigating the OI theory and researchers investigating GML theory use to measure morpheme growth can produce very different pictures of development in the same language sample. As an example, auxiliary *do* is used a total of 36 times in *Green Eggs and Ham* (Dr. Seuss, 1960). A striking contrast is found when token-based measures of accuracy and type-based measures of spontaneous productivity are used to interpret the use of AUXILIARY DO use in *Green Eggs and Ham*.

If Brown's (1973) token-based measure is used, Dr. Seuss demonstrates full mastery of auxiliary *do* in *Green Eggs and Ham*. AUXILIARY DO is used to mark tense and agreement in clauses containing a lexical verb and no other auxiliary forms. AUXILIARY DO is obligatory in syntactic contexts separating the tense marker from the verb (subject-auxiliary inversion, negation, and grammatical ellipsis) and in pragmatic contexts requiring emphasis. In these obligatory contexts, AUXILIARY DO is used to overtly mark tense and agreement and the lexical verb is produced as a bare stem. *Green Eggs and Ham* contains a total of 36 obligatory contexts requiring the use of auxiliary *do* for *do*-support. Auxiliary *do* is correctly used in all of these contexts. Of these, 34 tokens require *do*-support because the tense marker is separated from the verb by negation (9a-9b), 1 token requires *do*-support because the tense marker is separated from the verb by subject-auxiliary inversion to produce a yes/no question (9c), and 1 token requires *do*-support for emphasis (9d).

9. a. I do not like green eggs and ham
   b. You do not like green eggs and ham
   c. Do you like green eggs and ham
   d. I do so like green eggs and ham

In this analysis, auxiliary *do* is fully mastered and used correctly in 100% of obligatory contexts. *Green Eggs and Ham* contains zero obligatory contexts requiring the use of auxiliary *does* or *did* for *do*-support. Therefore, no interpretation can be made about Dr. Seuss' level of accuracy using auxiliary *does* or *did*. The OI theory would predict that other tense and agreement morphemes have been mastered because they grow with similar developmental trajectories towards a criterion of mastery (Rice et al., 1998).

If Hadley and Short's (2005) type-based measures of productivity are used, Dr. Seuss' use of the AUXILIARY DO category in *Green Eggs and Ham* is less robust than indicated by token-based measures of accuracy. In this analysis, the AUXILIARY DO category contains three different tense markers: auxiliary *do*, *does*, and *did*. Auxiliary *do* is used 36 separate times, but auxiliary *does* and *did* are not used at all. This direct evidence indicates that *do* is at least emergent (used at least once) and no direct evidence to indicate that *does* and *did* have emerged. The productivity measure reports the number of different subject/tense marker combinations using *do*. *Green Eggs and Ham* contains 33 separate tokens of *I do* (10a), which are all counted as one subject-tense marker combination. *Green Eggs and Ham* also contains 2 tokens of *you do* (10b) and 1 token of *do you* (10c). Following (Rispoli & Hadley, 2011) the three tokens of *you do/do you* are counted as one subject-tense marker combination even though the tense marker has been moved via subject-auxiliary inversion in one token (10c).

10. a. I do not like green eggs and ham

b. You do not like green eggs and ham

c. Do you like green eggs and ham

In this analysis, one out of three tense markers in the AUXILIARY DO category is used in two sufficiently different contexts (i.e. subject-tense marker combinations). This is evidence of initial productivity (Hadley & Short, 2005), but not a fully productive morpheme category or a full set of tense markers. The GML theory would predict similar levels of productivity for the *-3s* and *-ed* categories because these three morpheme categories have similar positional processing contexts and grow together at similar rates, but different levels of productivity for the COPULA BE and AUXILIARY BE categories, which have different positional processing contexts and different developmental trajectories (Rispoli & Hadley, 2011; Rispoli et al., 2012).

Token and type-based measures clearly capture fundamentally different elements of any given language sample. Composite measures of productivity and tense marker accuracy are correlated in naturalistic language samples (Hadley & Short, 2005). Both types of measures have been used to describe the growth of the tense and agreement system (e.g., Hadley et al., 2014; Rice et al., 1998; Rispoli et al., 2009, 2012; Wexler et al., 2004) and to differentiate between young children with typical language development and children with specific language impairment (e.g., Gladfelter & Leonard, 2013; Hadley & Short, 2005; Rice et al., 1998; Wexler et al., 2004). These measures are used in research studies testing the predictions of two different theoretical frameworks that attempt to explain the acquisition of the same morphosyntactic system. The OI theory is a null hypothesis predicting that the different tense and agreement morphemes will grow in accuracy together with similar growth rates, while the GML theory is an alternative hypothesis predicting that different morphemes will grow in productivity at different rates. Both measures could feasibly be used in a statistical test against the null hypothesis of the OI theory. However, neither measure addresses theoretical predictions related to both accuracy and productivity. Measures of productivity are preferred for testing a null hypothesis that morphemes grow together because productivity measures minimize the influence of high-frequency forms retrieved through direct activation, a phenomenon widely recognized by both generative and usage-based theories of language (Bybee, 2006; Pinker, 1999; Pinker & Prince, 1994; Ullman, 2001).

2.5 STIMULABILITY TESTING

Two critical purposes of expressive morphology assessment are to establish baseline levels of expressive morphology skills prior to intervention and to select specific goals for intervention targeting expressive morphology skills (Paul, 2001; Westby et al., 1996). Many clinicians recommend conducting an assessment that allows one to identify skills within a given child's zone of proximal development (Vygotsky, 1934/2012) when trying to select specific goals for intervention targeting expressive morphology. For example, Fey, Long, and Finestack (2003) and Long and Olswang (1996) recommend selecting specific goals that the child is developmentally ready to learn.

Vygotsky (1934/2012) proposed a domain-general learning theory claiming that children's level of development in any given area is not manifested as a discrete point along a continuum. At any given time, a developing child will have an established set of mental skills at the low end of this continuum, which can be performed independently without any adult assistance. At the time, there will be a set of more complex skills at the high end of this continuum that the child is not capable of performing, even with adult assistance. These skills, which exceed the child's capabilities cannot be learned yet, even with maximal support from well-designed adult intervention programs. Between the low level skills that the child can perform independently and the high level skills that are beyond the child's capabilities is a zone of proximal development consisting of emerging skills that the child can perform with assistance in optimal circumstances. Vygotsky (1934/2012) argues that children will only master additional skills if instruction targets emerging skills within the zone of proximal development. As this happens, the zone of proximal development advances and new skills begin to emerge. Vygotsky

(1934/2012) believed that this focused learning and instruction can be used as a ‘motor’ to drive development (Kozulin, 2011).

Kozulin (2011) suggests that Vygotsky (1934/2012) hinted at two fundamentally different areas of research in his early writing on the zone of proximal development, and that subsequent researchers have followed both of these research tracks with interchangeable and often conflicting terminology. One area of research has been aimed at developing new methods to evaluate children’s thinking along a continuum and determine the degree to which children’s thinking is modifiable. Research in the area of dynamic assessment has focused on the modifiability of cognitive functions and children’s readiness to transition between cognitive/developmental stages. The second area of research has been aimed at developing methods to assess children’s learning potential and evaluate children’s ability to benefit from models, cues, problem solving with adult guidance, and examples during different learning tasks (Kozulin, 2011). Assessment of learning potential is useful for identifying skills to target as specific goals in intervention programs that the child is currently ready to learn with guidance.

Researchers and clinicians in the field of speech-language pathology have borrowed central concepts from Vygotsky’s (1934/2012) domain-general learning theory to establish procedures for assessing the development of domain-specific speech and language skills at a given point in time. Dynamic assessment and assessment of learning potential are both well-established in the field of speech-language pathology. These fundamentally different types of assessment are typically used to investigate children’s levels of development in different language domains. The central concepts of dynamic assessment and learning potential assessment are useful for determining how to teach a child speech and language skills within a zone of proximal development and what skills to teach at a given point in time. However, these concepts are only useful for informing clinical practice if they are applied appropriately in a domain-specific theory of speech or language acquisition.

Dynamic assessment strategies are frequently used to assess the modifiability of children's early language skills and determine the amount and type of instructional support necessary to help children learn to use developing language skills more efficiently. The primary focus of a dynamic assessment approach is to characterize how a child engages in learning processes while building skills at a given level of development (Lidz & Peña, 1996). Through this dynamic assessment process, clinicians can "determine the type of help that is best provided" for a given child (Lidz & Peña, 1996, p. 369). Dynamic assessment has been used to characterize instructional support necessary to help diverse populations of children learn a wide range of communicative behaviors. Dynamic assessment has been used to identify teaching strategies to help verbally speaking children learn to produce new vocabulary words (Olswang, Bain, Rosendahl, Oblak, & Smith, 1986) and two-word utterances (Long & Olswang, 1996). Dynamic assessment also has been used to characterize the amount of support necessary to teach early semantic relations to pediatric AAC speakers (King, Binger, & Kent-Walsh, 2015) and to identify strategies for teaching triadic eye gaze (Olswang, Feuerstein, Pinder, & Dowden, 2013) and early symbolic communication behaviors (Snell, 2002) to individuals who could potentially benefit from AAC intervention. In most of these studies, one language skill was selected for use in dynamic assessment tasks after other measures were used to establish baseline levels of performance. Although dynamic assessment methods can provide information about modifiability, they provide no information about a child's learning potential.

A different strategy, known as stimulability testing (Millisen, 1954) is well-established as an essential component of assessment for phonological and articulatory disorders (for reviews, see Powell & Miccio, 1996; Rvachew, 2005). Early forms of stimulability testing were implemented long before Vygotsky's writing on the zone of proximal development was popularized in the English-speaking world (Millisen, 1954). Later researchers highlighted stimulability testing as a means of assessing learning potential within the zone of proximal development (Bain, 1994; Powell & Miccio, 1996) and began using stimulability testing to establish specific goals for intervention at the phoneme level (Powell, Elbert, & Dinnsen, 1991; Rvachew & Nowak, 2001). Unlike many dynamic assessment models, the primary focus of stimulability testing is to identify skills to target in intervention, or to determine what to teach at a given point in time.

In phonological stimulability tests, a child is prompted to watch and listen while a clinician models a target phoneme. The child is then prompted to repeat the target phoneme. The phoneme may be presented in isolation, or in a wide range of contexts. In all variants of this task, the child is cued and expected to follow an adult model to complete a task that cannot be completed independently. This process differentiates between developing phonemes that a child can produce with guidance and phonemes that are beyond the child's capabilities at the time of an assessment. Overt evidence of stimulability is considered a demonstration of learning potential and a prognostic indicator for improvements in articulatory accuracy (Rvachew, 2005) that reflects underlying phonological knowledge about a phoneme (Powell & Miccio, 1996). The adult model provided during stimulability testing is a highly invasive maximal prompt. Overt evidence of stimulability indicates that a child can at least produce a target with maximal support. Further intervention may provide evidence that the child requires a lower degree of support. Stimulability tests are frequently implemented in informal assessment tasks, and integrated into standardized protocols for assessing a broad range of phonemes (Goldman & Fristoe, 2015).

Powell et al. (1991) established procedures for selecting specific phonemes to target in intervention based on a given child's baseline phonetic inventory and phoneme-specific measures of stimulability for phonemes that were not part of the phonetic inventory at baseline. First, a baseline phonetic inventory of major English consonants in prevocalic, intervocalic, and postvocalic environments was established by collecting a spontaneous conversational speech sample and eliciting productions of target words representing each phoneme. This baseline phonetic inventory provided an index of phonemes that the child had acquired and could produce independently without adult assistance. Next, a series of phoneme-specific stimulability tasks was administered to assess stimulability of each phoneme that did not appear in at least two contexts in the child's phonetic inventory. This follow-up stimulability task differentiated between stimulable phonemes that could be produced following an adult model and phonemes that were beyond the child's capabilities at the time of the baseline assessment.

In a randomized controlled trial, Rvachew and Nowak (2001) investigated the relative effectiveness of two different target selection strategies on phonological learning over the course of 12 weeks of speech-language therapy for children with moderate-severe delays in phonological development. One group received treatment targeting highly stimulable, early developing phonemes. A second group received treatment targeting late-developing phonemes with low stimulability to test a hypothesis that treatment of later-developing phonemes may result in system-wide generalization (Gierut, Morrisette, Hughes, & Rowland, 1996). Both groups showed progress towards mastery of target phonemes. However, the group receiving treatment targeting highly stimulable targets showed greater progress towards acquisition of target sounds and better post-treatment maintenance of these gains than children in the low stimulability group. No significant differences between groups were found for measures of generalization to non-target phonemes. This randomized controlled trial provides evidence that treatment targeting the targets a child is most stimulable for is likely to result in greater treatment effects than treatment targeting the targets a child is least stimulable for (Rvachew, 2005; Rvachew & Nowak, 2001).

The central concepts of stimulability testing may be extended to assessment of expressive skills in other language domains, including morphosyntax. For instance, there are 15 different tense markers in the English tense and agreement system. Testing the stimulability of each tense marker is an objective means of identifying tense markers with learning potential in the child's zone of proximal development.

Developmental researchers have never assumed that children have a single, unified zone of proximal development spanning tasks and learning domains. Children may develop skills in different tasks at different rates as they build experience, leading to several different zones of proximal development. Skills in one area may contribute to the development of skills in another area with overlapping requirements.

Stimulability of tense markers in the tense and agreement system may develop together as if the tense markers develop within a single zone of proximal development. This would be consistent with null hypothesis of the OI theory predicting that all tense markers in the system grow and develop together (Rice et al., 1998). If this is the case, then tense markers should have similar levels of stimulability across morpheme categories.

In contrast, the alternative hypothesis of the GML theory predicts that tense markers in different categories will grow at different rates (Rispoli & Hadley, 2011; Rispoli et al., 2012).

This suggests that there are multiple zones of proximal development in the tense and agreement system. If this is the case, then tense markers in morpheme categories that differ at the level of positional processing should have different levels of stimulability.

2.6 RESEARCH ON DEVELOPMENTAL MORPHOLOGY SKILLS OF PEDIATRIC AAC SPEAKERS

This section reviews prior research on the developmental morphology skills of pediatric AAC speakers. AAC speakers are people who use communication strategies other than verbal speech for expressive communication because natural speech alone does not adequately meet their expressive communication needs. This dissertation focuses specifically on bridging an assessment gap for children who use assistive technologies such as dedicated SGDs, communication apps on tablet devices, and manual communication boards as their primary means of expressive communication. Therefore, the term "pediatric AAC speakers" refers to children who use assistive technologies as their primary means of expressive communication throughout this document.

Like their verbally speaking peers, pediatric AAC speakers interact with communication partners by generating spontaneous novel utterances. Verbally speaking children can combine a finite set of linguistic elements to generate an infinite set of spontaneous novel utterances using their natural speech (Chomsky, 1965). In ideal circumstances, pediatric AAC speakers would have a similar ability to generate an infinite number of spontaneous novel utterances using finite means. Pediatric AAC also speakers generate spontaneous novel utterances by combining linguistic elements. These linguistic elements consist of orthographic symbols and pre-stored vocabulary words stored as a whole unit in the "dictionary" of the child's AAC system as the child's external lexicon (Klein, 2017). Pre-stored vocabulary words are typically accessed by selecting graphic symbols (single meaning pictures or sequenced multi-meaning icons). Highly fluent pediatric AAC speakers may use a small number of pre-programmed utterances, or complete sentences stored as whole units in the external lexicon in pragmatically appropriate contexts. However, a majority of the utterances produced by pediatric AAC speakers are generated spontaneously.

A review of the external evidence on syntax and grammatical morphology in pediatric AAC speakers was conducted to identify the best evidence on grammatical morphology in pediatric AAC speakers. This review found minimal evidence of expressive morphology assessment. However, clear guidelines for adapting known expressive morphology assessment strategies to accommodate the needs of AAC speakers were identified.

During this review process, structured searches were conducted in three electronic databases to identify peer-reviewed publications on expressive language development in AAC speakers with congenital impairments during Brown's (1973) Stages. The databases searched include Education Resources Information Center, Linguistics and Language Behavior Abstracts, and PsycINFO. Searches were conducted in each database using Boolean phrases for all permutations of ["augmentative communication" "alternative communication"] AND [morphology syntax "language development"] to identify pairs of search terms that occurred anywhere in the full text of a document. In addition, hand searches were conducted in three journals that were known to publish papers on AAC intervention with pediatric AAC speakers: *Augmentative and Alternative Communication*, *Disability and Rehabilitation*, and *Journal of*

Speech, Language, and Hearing Research. Morphology assessments for pediatric AAC speakers identified in the external evidence are summarized in Table 2.1.

Several researchers have assessed the morphology skills of AAC speakers using receptive morphology tasks without measuring productive morpheme use (Berninger & Gans, 1986; Binger, Kent-Walsh, Berens, Del Campo, & Rivera, 2008; Binger & Light, 2007; Binger, Maguire-Marshall, & Kent-Walsh, 2011; Bruno & Trembath, 2006; Cumley & Swanson, 1999; Redmond & Johnston, 2001; Tönsing, Dada, & Alant, 2014; Wilkinson, Romski, & Sevcik, 1994). Some studies have reported results of composite receptive language assessment batteries that include receptive morphology components (Binger et al., 2008; Binger & Light, 2007; Bruno & Trembath, 2006; Cumley & Swanson, 1999; Tönsing et al., 2014; Wilkinson et al., 1994). Other studies have reported measures specific to receptive morphology (Berninger & Gans, 1986; Binger et al., 2011; Blockberger & Johnston, 2003; Redmond & Johnston, 2001).

Receptive morphology assessment has been used to identify targets for an intervention program focused on building the expressive morphology skills of AAC speakers (Binger et al., 2011). Binger et al. (2011) conducted an item analysis of participant performance on the Grammatical Morphemes subtest of the *Test for Auditory Comprehension of Language- Third Edition* (TACL-3; Carrow-Woolfolk, 1999) in order to identify three target morphemes that were within the developmental capabilities of each participant. These three targets were treated in a multiple baseline across-behaviors design for each participant. High within-subject variability and a lack of within-subject replication were found in the first post-treatment maintenance phase. All participants met a criterion level of >80% accuracy in post-treatment maintenance probes for at least one target morpheme. At the same time, all participants had 0% post-treatment

Table 2.1. Morphology assessment for pediatric AAC speakers in the external evidence

Reference	Expressive Morphology Measures	Receptive Morphology Measures	AAC Potential to Support Morpheme Use
Berninger & Gans (1986)	None	Comprehensive report of spoken & written language comprehension	No details related to morphology provided
Binger & Light (2007); Binger et al (2008)	None	Composite receptive language measure (TACL-3)	No details related to morphology provided
Binger et al. (2011)	Performance in treatment targeting expressive morpheme use.	Composite receptive language measure (TACL-3). Grammatical Morphemes subtest scores reported & used to select treatment goals.	Clear report of method of producing target morphemes on participants' SGDs.
Blockberger & Johnston (2003)	Production of suffixes (possessive *-s*, past tense *-ed*, third person singular *-3s*) in structured writing task where first letter of target word is provided, only used to assess literate participants	Grammaticality judgment and grammatical morpheme comprehension task for three suffixes (possessive *-s*, past tense *-ed*, third person singular *-3s*)	Access to spelling described for each participant. No information on device potential to support morpheme use in participants without spelling.
Bruno & Trembath (2006)	Elicited immitation task, some items include grammatical morphemes	Composite receptive language measure (TACL-3)	Photos of manual communication boards and SGDs used by participants show some morphemes
Cumley & Swanson (1999)	None	Composite receptive language measure (TACL-R)	Samples communication boards include prepositions
Redmond & Johnston (2001)	None	Several grammaticality judgment tasks	Method of producing grammatical morphemes defined for each participant, w/examples.

Table 2.1. (Continued)

Reference	Expressive Morphology Measures	Receptive Morphology Measures	AAC Potential to Support Morpheme Use
Tonsing (2014)	None	Composite receptive language measure (CELF-P(UK))	No grammatical morphemes included on system
Wilkinson et al. (1994)	None	Composite receptive language measure (ACLC)	No grammatical morphemes included on system

Note. TACL-3 = *Test for Auditory Comprehension of Language, 3rd Edition* (Carrow-Woolfolk, 1999); TACL-R = *Test for Auditory Comprehension of Language, Revised Edition* (Carrow-Woolfolk, 1985); CELF-P(UK) = *Clinical Evaluation of Language Fundamentals-Preschool UK* (Wiig, Secord, & Semel, 2000); ACLC = *Assessment of Children's Language Comprehension, 1983 Revision* (Foster, Giddan, & Stark, 1983).

maintenance for at least one target morpheme and required a second intervention phase to re-inforce use of that morpheme before any post-treatment maintenance was observed. Using measures of expressive morphology skills to identify targets for an intervention program may increase the likelihood that new skills will be maintained after an initial course of intervention.

Only one study assessing expressive use of grammatical morphemes in pediatric AAC speakers was identified in the external evidence. Blockberger and Johnston (2003) investigated the acquisition of possessive *–s*, *–ed*, and *-3s* in receptive vocabulary-matched groups of pediatric AAC speakers, typically developing children, and verbally speaking children with developmental language delays. Three separate tasks were used to assess receptive comprehension and expressive production of each target morpheme. A grammaticality judgment task and a sentence-picture matching task were used to assess receptive morpheme comprehension. A structured writing task was used to assess expressive production of each target morpheme. In the structured writing task, participants were asked to complete a short story by filling in the ends of incomplete words (the first letter was provided) representing obligatory contexts for target grammatical morphemes. The group of pediatric AAC speakers had significantly lower performance than the matched groups of typically developing children and verbally speaking children with developmental language delays in all three tasks.

Unfortunately, Blockberger and Johnston's (2003) structured writing task can only provide limited information about AAC speakers' expressive use of morphemes. This task was only useful for eliciting productions of suffixes bound to a root word and sufficient for assessing specific morphemes in isolation. Variants of this task can only be used to assess expressive use of five out of 14 grammatical morphemes studied by Brown (1973) and two out of 15 tense markers studied by Hadley and Short (2005). This structured writing task is not sufficient for assessing development of a morphosyntactic system that includes free morphemes.

In addition, Blockberger and Johnston's (2003) structured writing task assumes that all participants have a requisite level of basic literacy skills access to a spelling system. Only 40% of the AAC speakers in their study met inclusion criteria for participation in the structured writing task. This is evidence of a device-dependent contingency: expressive morphology assessment required AAC system features that allowed for selective production of the target morpheme (and selective production of a bare stem that omits the target morpheme) in the assessment task.

This device-dependent contingency is central to some AAC intervention protocols, which are designed around the assumption that children must learn to select preferred graphic symbols from arrays with at least one competing non-preferred graphic symbol in order to establish relationships between symbols and referents (Bondy & Frost, 1994; Frost & Bondy, 2002). The Picture Exchange Communication System (Bondy & Frost, 1994; Frost & Bondy, 2002) is a structured intervention protocol for teaching children to use graphic symbols for basic communicative behaviors, such as requesting. Symbol discrimination is a crucial step in this training protocol used to teach selective use of graphic symbols as new lexical items. During this step, children are required to discriminate between multiple symbols and select their preferred symbol to communicate their desired message. Selective use of a target symbol for expressive production of a specific word or message is often considered evidence of successful lexical learning in the graphic symbol modality.

In a paper reviewing evidence on the morphology and syntax of AAC speakers, Binger and Light (2008) extend this concept of selective graphic symbol use to expressive morphology assessment and intervention. Binger and Light (2008) suggest that many strategies for teaching and assessing expressive use of morphemes can be adapted for AAC speakers. They give several examples of known expressive morphology assessment and intervention strategies, which can be adapted for AAC speakers using a system that allows for selective morpheme production. In all of their assessment examples, an AAC speaker's ability to use *–ed* is measured by eliciting productions of the inflected form of a lexical verb in contrast to the uninflected form of the same lexical verb. In all of their intervention examples, an AAC speaker is taught to selectively produce the inflected form of a lexical verb on a system that allows for production of both inflected and uninflected forms. These strategies may be difficult or impossible to adapt for AAC speakers using systems that do not allow for selective morpheme production.

Some compelling evidence shows that AAC speakers' receptive knowledge of grammatical morphemes is not determined *a priori* by the scope of the pre-stored language content on their AAC systems. Redmond and Johnston (2001) used a series of grammaticality judgment tasks to measure receptive knowledge of affixes in four 11-15 year old children with cerebral palsy or quadriplegia secondary to a brainstem aneurysm who used a range of different AAC systems for communication. These participants consistently recognized that aspect marking is obligatory, auxiliaries must agree in person and number with their subjects, irregular verbs must be marked for tense, and overregularization errors are ungrammatical. These same participants all had difficulty detecting bare stem regular verb errors. The authors argue that morphological weaknesses in AAC speakers may be localized or specific to particular morphemes and grammatical features rather than a generalized limitation in affixation.

The AAC systems used by Redmond and Johnston's (2001) participants were highly idiosyncratic and used diverse methods of affixation. Participant performance was not related to the method of affixation available on the participant’s AAC system, or the complete absence of a method of affixation. Most notably, one 15-year old participant had an AAC system which was limited to a simple voice output communication aid with single pre-programmed utterance and a manual eye-gaze system with single meaning pictures but no method of affixation. No conclusions can be made about whether this adolescent participant’s receptive knowledge of morphology developed in a typical time-line without a method of morpheme production. However, this participant’s robust comprehension of morphological errors on a grammaticality judgment task indicates that receptive knowledge of grammatical morphemes can develop without an accessible method of morpheme production.

In their review on morphology and syntax in AAC speakers, Binger and Light (2008) clearly show that adapting expressive morphology assessment and intervention strategies for AAC speakers requires integration of aided language stimulation (Goosens', 1989) into the assessment/intervention protocol. Aided language stimulation is an intervention technique used to support development of expressive communication skills in pediatric AAC speakers. Aided language stimulation is closely adapted from spoken language intervention strategies in which a clinician shapes conversations to model production of target language structures. In aided language stimulation, the clinician adapts these strategies by selecting symbols on the child’s AAC system to explicitly model production of target words or structures in the graphic symbol modality (Goosens', 1989). When an AAC speaker uses a voice output communication aid or SGD, clinicians often model utterances by selecting symbols and triggering the device to produce speech output in conjunction with their models (Romski & Sevcik, 1996; Wilkinson et al., 1994).

Aided language stimulation provides comprehension experience in the modality that the AAC speaker uses to produce linguistic output, eliminating a frustrating input-output asymmetry for a short period of time. Pediatric AAC speakers experience a distinct asymmetry between the primary channel for language input and their primary channel for language output, which is not found in typical language development (Smith & Grove, 1999, 2003; Sutton, Soto, & Blockberger, 2002). Pediatric AAC speakers have little or no exposure to language models in the modality they use for generating linguistic output because they are immersed in a spoken language environment and receive spoken language input through the auditory channel (Smith & Grove, 2003; Sutton et al., 2002; von Tetzchner & Jensen, 1996). AAC speakers may only receive language input that is compatible with their output modality in structured teaching situations involving aided language stimulation (Sutton et al., 2002; von Tetzchner & Jensen, 1996).

Mirenda (2008) strongly concludes that aided language stimulation is an inherent and requisite component of any AAC intervention using an SGD based on typical language development because clinicians should always expect language comprehension experiences to precede expressive language production. Clinicians who implement aided language stimulation provide pediatric AAC speakers with exposure to language input in their output modality before expecting productive graphic symbol use. Aided language stimulation may be faded over the course of treatment to encourage independent production of targets while other techniques are then implemented.

Aided language stimulation techniques are consistently used in AAC intervention studies targeting use of language structures above the single word level. A total of 8 intervention studies targeting expressive language skills above the single word level in pediatric AAC speakers with congenital disabilities were identified in the external evidence (Binger et al., 2008; Binger & Light, 2007; Binger et al., 2011; Iacono, Mirenda, & Beukelman, 1993; Kent-Walsh, Binger, & Buchanan, 2015; Nigam, Schlosser, & Lloyd, 2006; Tönsing et al., 2014; Wilkinson et al., 1994). Aided language stimulation was an integral component of every intervention strategy reported across all 8 identified studies.

The modeled examples of AAC system use provided in aided language stimulation give pediatric AAC speakers focused examples of language input in the graphic symbol modality and may be used as a maximal prompt in a prompting hierarchy designed to elicit productions of target language structures in the graphic symbol modality. Some treatment studies have specifically used aided language stimulation to provide language input in the graphic symbol modality without utilizing a prompting hierarchy to elicit responses (Binger et al., 2008; Binger & Light, 2007; Binger et al., 2011; Kent-Walsh et al., 2015; Wilkinson et al., 1994). Other treatment studies have used aided language stimulation as a maximal prompt in a least-to-most prompting hierarchy (Iacono et al., 1993; Nigam et al., 2006; Tönsing et al., 2014).

Although the presence of aided language stimulation is 100% consistent across intervention studies, no forms of aided language stimulation were provided in Blockberger and Johnston (2003) expressive morphology assessment task. This is logical for a traditional elicitation task that uses contextual information to prime a child for production of a target morpheme.

Explicit modeling of a target form is an integral component of stimulability tasks. In a morpheme stimulability test, this explicit modeling would include production of a target morpheme in the context of a grammatically correct sentence. If this modeling is provided in the graphic symbol modality, at least the portion of the modeled sentence containing the target morpheme should be provided in the graphic symbol modality as an explicit model. In both spoken and graphic symbol modalities, the models provided in stimulability tests are a maximal prompt serving the same purpose of modeled input using the spoken modality in phonological stimulability tests. Such stimulability tests may be useful for identifying target morphemes that children can produce following a highly invasive maximal prompt. Subsequent intervention programs would focus on helping the child produce the target morphemes with a greater degree of independence. Aided language stimulation may be an integral component of those intervention programs.

An improved method of assessing the expressive morphology skills of pediatric AAC speakers is needed. Given the evidence that verbally speaking children can have particular difficulty developing productive use of the tense and agreement morpheme system, this same system should be assessed in pediatric AAC speakers using measures that capture development. This improved method should meet the criteria in (11). Table 2.2 summarizes how each of the criteria in (11) are addressed in Experiments 1-3 of this dissertation project.

11. Criteria for assessing expressive use of tense and agreement morphemes in pediatric AAC speakers.
    a. This method should use graphic symbol-based tasks that do not require literacy skills so that children with poor literacy skills are not systematically excluded from assessment as they were in Blockberger and Johnston (2003).
    b. This method should use device-agnostic tasks that are not contingent on the use of a specific AAC system, as in Redmond and Johnston (2001).
    c. This method should account for the potential of a child’s AAC system to support productive morpheme use, as in Redmond and Johnston (2001).

d. This method should follow stimulability testing procedures to establish baseline levels of productivity and stimulability (Powell et al., 1991; Rvachew & Nowak, 2001).

e. Measures of productivity should be taken to assess present levels of morpheme use (Hadley & Short, 2005; Rispoli & Hadley, 2011).

f. Measures of stimulability (Millisen, 1954) should be taken to assess readiness to learn additional morphemes in intervention and identify morphemes to target as specific goals for intervention (Fey et al., 2003; Long & Olswang, 1996).

g. Explicit modeling of morphemes in the graphic symbol modality should be used to adapt stimulability tasks for the graphic symbol modality (Binger & Light, 2008).

h. All stimulability tasks should require children to selectively produce target morphemes using an array of graphic symbols that allows for selective production of at least one contrasting morphological form (Binger & Light, 2008; Bondy & Frost, 1994).

A growing body of research studies has investigated the use of graphic symbols for communication in groups of typically developing children and nondisabled adults. Several studies have investigated production of graphic symbol sentences in typically developing children and nondisabled adults using experimental AAC systems with limited pre-stored vocabularies (Nakamura, Newell, Alm, & Waller, 1998; Smith, 1996; Sutton, Gallagher, Morford, & Shahnaz, 2000; Sutton & Morford, 1998; Trudeau, Sutton, Dagenais, de Broeck, & Morford, 2007). Other studies have investigated typically developing children's use of graphic symbols to produce single words on AAC systems with pre-stored vocabulary words arranged in different configurations (Drager et al., 2004; Drager, Light, Speltz, Fallon, & Jeffries, 2003; Light et al., 2004). These studies did not compare use of any specific expressive language skills across spoken and graphic symbol modalities, and only used measures of spoken language skills

Table 2.2. Summary of ways assessment criteria are accounted for in Experiments 1-3

Criterion	Experiment 1	Experiment 2	Experiment 3
11a	NA	Stimulability tasks use graphic symbols with no text gloss	• Spontaneous production tasks use child's own SGD • Stimulability tasks use graphic symbols with no text gloss
11b	NA	Stimulability tasks use custom device-agnostic vocabulary pages	Stimulability tasks use custom device-agnostic vocabulary pages
11c	NA	NA	SGD is reviewed to ensure the device supports a fully productive tense & agreement system.
11d	Language sample task contributes to comprehensive protocol in Experiment 3	Stimulability tasks contribute to comprehensive protocol in Experiment 3	• Measures of use: language sample from Experiment 1 & elicitation probe. • Measures of stimulability: stimulability tasks from Experiment 2
11e	Measures of productivity from language sample	NA	Measures of productivity from language sample & elicitation probe
11f	NA	Stimulability tasks obtain measures of stimulability for all 15 tense markers and all 5 morpheme categories.	Stimulability tasks obtain measures of stimulability for all 15 tense markers and all 5 morpheme categories.
11g	NA	All stimulability tasks incorporate modeling in the graphic symbol modality.	All stimulability tasks incorporate modeling in the graphic symbol modality.
11h	NA	All stimulability tasks require the child to select between multiple competing tense markers presented simultaneously on the SGD vocabulary page.	All stimulability tasks require the child to select between multiple competing tense markers presented simultaneously on the SGD vocabulary page.

to verify that participants met inclusion criteria. These studies do not address hypotheses about differences in language skills across modalities, and provide no reason to reject a null hypothesis that language skills develop in similar ways across communication modalities. Studies investigating naturalistic language samples from pediatric AAC speakers with congenital disabilities have shown that pediatric AAC speakers can develop emergent use of grammatical morphemes (Ortloff, 2010) and syntactic structures involving English auxiliaries (Kovacs & Hill, 2013) in a typical developmental sequence.

The overarching null hypothesis that expressive language skills develop similarly across communication modalities is a crucial but generally unstated hypotheses with critical implications for pediatric AAC assessment and intervention. If the null hypothesis is true, then the robust external evidence on normal and disordered spoken language development can be used to guide assessment and intervention programs for AAC speakers acquiring developmental skills. For example, known strategies for assessing the development of spoken language skills could be adapted into an accessible format and implemented in the graphic symbol modality. If the null hypothesis is not true, then extensive research is needed to characterize developmental milestones in the graphic symbol modality and provide an evidence base for guiding modality-specific assessment and intervention strategies. Experiment 2 of this dissertation project began addressing this null hypothesis by assessing tense marker stimulability in a cross-sectional sample typically developing children in two communication modalities and modeling growth of tense marker stimulability across communication modalities.

2.7 SUMMARY

This chapter outlined two different theories explaining the growth of a unified morphosyntactic system that includes a diverse set of 15 English tense markers. OI theory holds that the tense and agreement system matures as a unique checking constraint withers, allowing children to produce finite tense marked forms in a progressively larger proportion of obligatory contexts (Wexler, 1998, 2003, 2011). OI theory predicts that the tense markers will grow together with highly similar trajectories (Rice et al., 1998). In contrast, GML theory holds that the tense and agreement system is unified at the level of grammatical features, and that the system will grow in partial sequence (Rispoli & Hadley, 2011; Rispoli et al., 2012). Specifically, GML theory predicts that different categories of tense markers will grow together in productivity if they have enough similarity at the level of positional processing (Rispoli & Hadley, 2011; Rispoli et al., 2012).

Evidence from spontaneous language samples of children acquiring English is consistent with the predictions of both of these theories. Rice et al. (1998) found that several morphemes in the tense and agreement system grow together in accuracy in longitudinal data from typically developing children and children with specific language impairment. In contrast, Rispoli et al. (2012) found that these same morphemes grow in productivity at different rates in longitudinal data from typically developing children, and that morpheme categories with similar positional processing contexts grew in productivity together. One cross-sectional study unexpectedly found patterns of productivity across morpheme categories that were not consistent with the predictions of either theory (Gladfelter & Leonard, 2013).

These conflicting findings were obtained from similar language samples using two fundamentally different measures: a token-based measure of accuracy adapted from Brown

(1973) and a type-based measure of productivity adapted from Radford (1990). These two measures can lead to different interpretations about the development of the same morphemes in the same language sample. Measures of tense marker use in spontaneous language samples may not be sufficient for differentiating between theories. Both measures are useful for establishing baseline levels of tense and agreement morpheme use in children who may require intervention improving expressive use of the tense and agreement system. Measures of productivity are preferred because they minimize the influence of high-frequency forms retrieved through direct activation.

These two different theoretical perspectives, which are summarized in Figure 2.3, make different predictions about patterns of cross-morpheme generalization. This may lead clinicians to different intervention strategies for the same child at the same point in time. Cross-morpheme generalization of learned skills cannot occur in the purely maturational system postulated by OI theory. Clinicians working within OI theory may provide treatment targeting tense and agreement system growth by broadly targeting all tense markers to support localized learning across all morphemes in the system. Alternatively, they may target lexical skills or other language skills with the expectation that the tense and agreement system will gradually mature without intervention.

In the system postulated by GML theory, treatment effects are expected to generalize across tense markers with similar grammatical features. In this case, treatment may be streamlined by targeting a subset of tense markers and capitalizing on predicted patterns of cross-morpheme generalization.

Expressive morpheme assessment strategies may be adapted for pediatric AAC speakers by incorporating explicit modeling of target morphemes in the graphic symbol modality into assessment protocols in a template that follows the pattern of phonological stimulability tests. In phoneme stimulability testing, evidence of stimulability is a prognostic indicator of generalization. Researchers have not investigated whether or not the prognostic characteristics of stimulability tests generalize across language domains. If morpheme stimulability is prognostic, then stimulability tests can be used to predict individualized patterns of cross-morpheme generalization and guide individualized treatment targeting improved morpheme production. Research is needed to determine whether or not patterns of morpheme stimulability are consistent with patterns predicted by either the OI or GML theory.

- Acquisition is maturational: innate genetic program for universal grammar guides grammatical growth on a genetically determined timeline.
- Tense & agreement morphemes emerge simultaneously in a unified morphosyntactic system after parameters are set.
- Gradual and continuous period of growth is expected after initial emergence of morphemes
- Child experiences an optional infinitive stage, and uses finite and root infinitive verbs interchangeably in contexts where finiteness is obligatory. A unique checking constraint leads to inconsistent productions of root infinitives. Accuracy increases as the constraint withers developmentally.
- Grammatical learning occurs very early, parameters are set before the two-word stage.
- Unique checking constraint is a singular mechanism guiding growth of all tense and agreement morphemes. Patterns of growth are consistent across individual morphemes in the system and a composite measure.
- One zone of proximal development across all tense & agreement morphemes. Levels of stimulability are similar across morphemes.

- Tense & agreement system is unified at the level of grammatical features, develops gradually in a relatively uniform partial sequence, and is learned from input within an initial hypothesis space defined by universal grammar.
- Tense & agreement morphemes emerge simultaneously in a unified morphosyntactic system as grammatical features at level of functional processing are learned.
- Gradual and continuous period of growth is expected after initial emergence of morphemes
- Children begin with equal probabilities of selecting either a [+ Tense] or a [- Tense] grammar, and may use both. Probabilities of competing grammars are adjusted iteratively as child learns from input until parameter is set.
- Grammatical learning is a gradual, iterative process of statistical learning.
- Differences in growth rates arise because the different tense and agreement morphemes are processed differently in grammatical encoding. Morphemes with similar positional processing grow together.
- Multiple zones of proximal development across tense & agreement morphemes. Morphemes with similar positional processing have similar levels of stimulability.

Figure 2.3. Tense and agreement morpheme growth in an OI child and a GML child

3.0 METHOD

This chapter describes procedures for three experiments testing the competing hypotheses of the OI and GML theories. In each experiment, predictions of OI theory were treated as a null hypothesis that there would be no differences in participant performance across morpheme categories. The predictions of GML theory were treated as an alternative hypothesis that there would be a predictable pattern of differences across morpheme categories consistent with the pattern observed by Rispoli et al. (2012). Experiments 1 and 2 also tested for age-by-morpheme category interactions implied by the different patterns of morpheme category productivity observed in toddlers (Rispoli et al., 2012) and preschoolers (Gladfelter & Leonard, 2013).

The three experiments measure an inter-related set of dependent variables characterizing tense and agreement system development. Operational definitions for the dependent variables in all experiments are provided in Appendix A. Experiment 1 assessed the emergence and growth of tense and agreement productivity in a cross-sectional sample of typically developing children ages 30-54 months. Experiment 2 assessed stimulability of the tense and agreement system in the same children; critically, it did so by testing these children using both spoken- and novel graphic-symbol-modality measures, to establish the utility and reliability of the graphic-symbol-modality measures to assess grammatical morphology production. Experiment 3 assessed emergence, productivity, and stimulability of the tense and agreement system in pediatric AAC speakers with cerebral palsy and used the results of this assessment to select specific tense markers to target in a short course of intensive intervention. All three experiments were approved by the institutional review board at the University of Pittsburgh. A brief summary of all three studies is presented in Table 3.1.

3.1 PARTICIPANT RECRUITMENT & SCREENING (EXPERIMENTS 1 AND 2)

A cross-sectional sample of typically developing, monolingual English speaking children was recruited to participate in Experiments 1 and 2. The sample included 20 children ages 30-42 months and 20 children ages 43-54 months. This cross-sectional sample spanned the full age range represented by two prior studies of morpheme category productivity in young children (Gladfelter & Leonard, 2013; Rispoli et al., 2012) and an intermediate period of development that was not accounted for by these prior studies. One parent for each child also was included as a participant in the language sample task of Experiment 1. Parents were screened during the initial screening interview to verify that they were at least 18 years old and spoke fluent English. A flow-chart summarizing all Experiment 1 and 2 tasks involving participants is shown in Figure 3.1.

Participants were recruiting using a variety of strategies. Recruitment notices were distributed on social media, posted in hard-copy on community bulletin boards, and distributed in person at family-centered community events with permission from event organizers. In addition, the study was listed in the Pitt+Me Research Participant Registry.

Parents who expressed interest in participating with their child were contacted via telephone or email. After informed consent was obtained, an initial screening interview was

Table 3.1. Summary of Experiments 1-3.

	Experiment 1	Experiment 2	Experiment 3
Prticipnts	TD children n = 20 ages 30-42 months n = 20 ages 43-54 months	Same as Experiment 1	Pediatric AAC speakers w/cerebral palsy (n = 2)
Tasks	Language sample	Stimulability tasks, randomly assigned to communication modalities	Experiment 1 & 2 tasks Elicitation probes Review SGD potential to support tense marker use
Measures	Tense marker total Productivity score Category productivity score	Tense marker stimulability Category stimulability score	All Exp. 1 & 2 measures Potential TM total Potential prod. score Potential category prod. Score
H_0 (OI)	No differences between category productivity scores for either age group.	No differences between category stimulability scores for either age group.	No evidence of cross-morpheme generalization
H_1 (GML)	Category productivity varies across categories, as in Rispoli et al. (2012). Category productivity scores may diverge, but general pattern holds.	Category stimulability scores vary across categories, as in Rispoli et al. (2012). Category stimulability scores may diverge, but general pattern holds	Cross-morpheme generalization will pattern with tense markers that share features with the target tense marker.
H_2	Age-by-modality interaction reflecting different patterns of category productivity in prior studies of children in different age groups.	Age-by-modality interaction reflecting different patterns of category productivity in prior studies of children in different age groups.	
H_{Stim}			Cross-morpheme generalization will pattern with stimulable tense markers

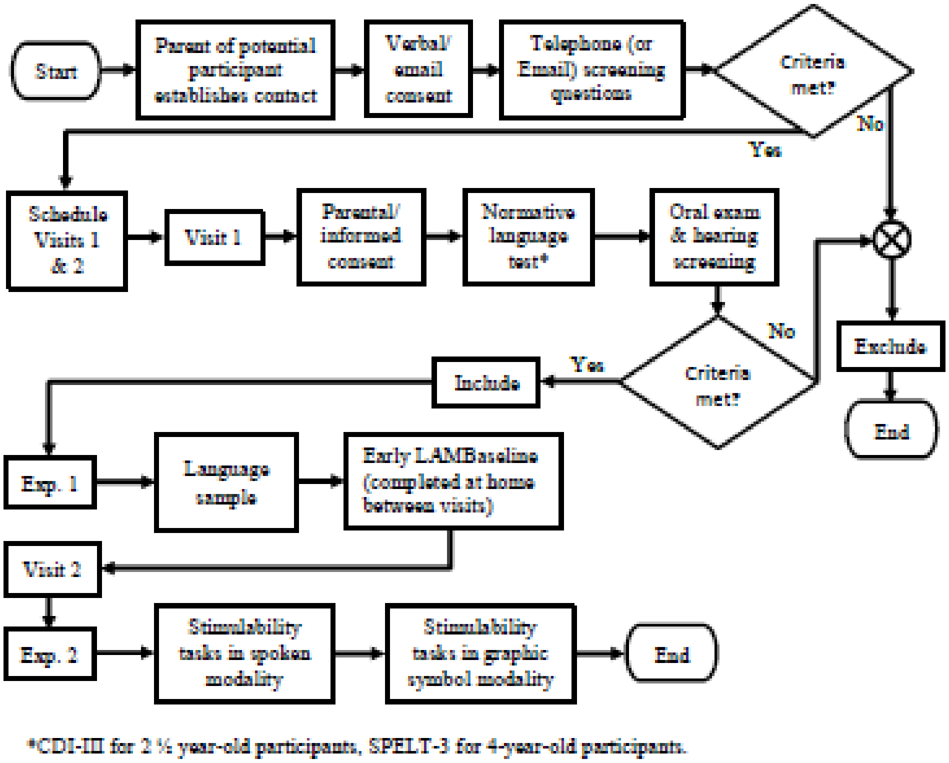


Figure 3.1. Flow chart for Experiments 1 and 2.

conducted via telephone or email. The screening interview verified that the parent-reported inclusion criteria in (12) were met.

12. Inclusion criteria for Experiments 1 and 2 verified by parent report over the telephone and/or email:

    a. The child is in one of two target age ranges (30-42 months, or 43-54 months).

    b. The child is monolingual and acquiring English as a first language.

    c. The child has no history of speech or language impairment.

    d. The child has no history of hearing impairment.

    e. The child has no history of cognitive impairment.

    f. The child has no history of neurological injury.

    g. The child has normal or corrected to normal vision.

h. The child can isolate one index finger for selecting symbols in Experiment 2.

i. The parent is at least 18 years old.

j. The parent is a fluent English speaker who uses verbal speech.

If all parent-reported inclusion criteria were met, two home visits were scheduled to complete all remaining study tasks on the assumption that the additional inclusion criteria also were met. Visit 1 lasted about two hours. During Visit 1, written parental consent was obtained, additional screening tasks were administered to verify that the inclusion criteria in (13) were met, and Experiment 1 data were collected if all criteria were met.

13. Inclusion criteria verified in person during Visit 1:

a. The child scored within the normal range on a normative test of language development. For children ages 30-36 months, the child was required to score at or above the 10th percentile for boys and girls their age on the *MacArthur Communicative Development Inventory-Level III* (CDI-III; Fenson et al., 2007). For children ages 37-54 months, the child was required to score no more than 1.5 standard deviations below the mean for their age on the *Structured Photographic Expressive Language Test, Preschool 2nd Edition* (SPELT-P2; Dawson et al., 2005).

b. An oral-peripheral examination indicated that the structure and function of the child's speech mechanism was within normal limits.

c. The child passed a bilateral pure-tone hearing screening at 25db at 500, 1,000, 2,000 and 4,000 Hz.

Hearing screenings were repeated at the beginning of Visit 2 for some participants because environmental noise prevented the author from administering a valid hearing screening or because the child was recovering from otitis media at the time of Visit 1. Visit 2 lasted about one hour. Experiment 2 data were collected during Visit 2. Child participants received $10 in compensation at Visit 1 and $15 in compensation at Visit 2. Parent participants were given the option to request a report with their child's individual results from Visit 1 as a child language evaluation.

A power analysis was conducted using published data on spontaneous category productivity scores in typically developing 30 month olds (Rispoli et al., 2012) and pooled data on category productivity scores in typically developing children ages 4;0-4;6 and 5;0-5;6 (Gladfelter & Leonard, 2013) to estimate the sample size necessary to detect true differences in the productivity of a morpheme category between age groups. The power analysis found that a total sample size of 40 participants (20 per age group) would achieve 80% power to detect an effect size of $f = 0.18$ in a mixed ANOVA with $\alpha = 0.05$. This sample size was collected with the original intent of analyzing data using analyses of variance. Regression analyses were used in the final analysis so that age could be treated as a continuous variable.

3.2 EXPERIMENT 1: ASSESSING PRODUCTIVITY IN TYPICALLY DEVELOPING CHILDREN

3.2.1 Purpose and research questions.

The purpose of Experiment 1 was to investigate tense and agreement morpheme emergence and productivity in spontaneous language samples from typically developing 30-54 month olds learning English as a first language. These measures have been reported in separate studies of toddlers (Rispoli et al., 2009, 2012) and preschoolers (Gladfelter & Leonard, 2013), which may not be directly comparable because they used different language sampling methods. Experiment 1 was a confirmatory study that used one sampling method and allowed for a better analysis of development across this full age range. Experiment 1 sought to answer three research questions.

Research question 1A asked if the number of tense markers children use in language samples with a fixed number of multi-morpheme utterances increases with age between 30 and 54 months. This question tested a null hypothesis that the number of tense markers used would not increase with age. In general, there is a consensus among researchers that several tense markers emerge simultaneously (Hadley & Short, 2005; Radford, 1990), and that the tense and agreement system grows over time after the point of initial emergence, suggesting that tense markers appear at a higher frequency in the language of older children (Rice et al., 1998; Rispoli et al., 2009, 2012). The researcher hypothesized that the number of tense markers used would increase with age, and that the null hypothesis would be rejected.

Research question 1B asked if productivity score increases with age between 30 and 54 months in language samples with a fixed number of multi-morpheme utterances. This question tested a null hypothesis that productivity score would not increase with age. Composite measures of accuracy and productivity both show evidence of overall growth in the tense and agreement system during the third year of life. In addition, productivity scores were higher in the 4 year olds studied by Gladfelter and Leonard (2013) than they were in the 2½ year olds studied by Rispoli et al. (2009) even though the patterns of category productivity scores across samples varied. The researcher hypothesized that productivity scores would increase with age and the null hypothesis would be rejected.

Research question 1C asked if morpheme category productivity increases at the same rate across morpheme categories. This question tested the null hypothesis of the OI theory, which predicts that the 5 categories will grow together with similar developmental trajectories. Under this null hypothesis, there should be no significant differences between category productivity scores across morpheme categories. This pattern has not been reported in studies using type-based measures of morpheme productivity.

In contrast, GML theory predicts that morpheme categories will grow in productivity in partial sequence, and that morpheme categories with enough similarity in positional processing will grow in productivity together. In their longitudinal sample, Rispoli et al. (2012) found that COPULA BE grew in productivity faster than any other category, followed by a wave of three categories growing together (AUXILIARY DO, *-3s*, *-ed*), followed by AUXILIARY BE. This pattern is predicted by the GML theory and can be explained by differences in positional processing across categories and unique grammatical features in the AUXILIARY BE category.

The alternative pattern observed by Gladfelter and Leonard (2013) cannot be explained by GML theory. GML theory predicts similar patterns of growth in *-3s* and *-ed* category productivity because these categories have the same positional processing features. A significant difference between *-3s* and *-ed* category productivity would suggest that differences in functional processing features can contribute to different patterns of growth among tense and agreement morphemes.

The researcher hypothesized that the developmental pattern reported by Rispoli et al. (2012) would be replicated in the cross sectional sample of Experiment 1. Differences between category productivity scores may increase with age and still follow this same general pattern. This would be consistent with predictions of GML theory and suggest that the difference between *-3s* category productivity and *-ed* category productivity reported by Gladfelter and Leonard (2013) was an artifact of the authors' language sampling method.

3.2.2 Procedures.

Experiment 1 was completed during Visit 1 after all inclusion criteria were met and the participants were given a bathroom break. Spontaneous, play-based language samples were collected using procedures adapted from Rispoli et al. (2009). Language samples were collected in a play area in the participants' home environments. One language sample was collected in an office board room. A controlled set of age-appropriate toys was provided for the participants to play with to standardize language sampling protocols as much as possible given variation across the participants' home environments. The toys included a ball, a work bench and hammer, wordless picture books to facilitate labeling, a set of Little Red Riding Hood puppet characters to facilitate narratives, a baby doll to facilitate feeding and bathing routines, cars and a play mat featuring a variety of community settings (hospital, police station, etc.) to facilitate pretend play, and a Little Tikes® Cook n' Store Kitchen to facilitate pretend play. These toys were selected for a mobile language sampling lab because they facilitated robust language use about a range of topics and they could be transported and set up with minimal difficulty. Stimulus toys are shown in Figure 3.2.

Following Rispoli et al. (2009), parents were instructed to talk with their child as they normally would at home. Language samples were recorded using a stationary video camera. The author monitored equipment and observed from an adjacent area to minimize interaction with participants during the recording session. The author responded politely any time participants addressed him directly.

Language samples were recorded for one continuous hour. Some samples were ended a few minutes early to accommodate participant schedules or to limit the total time needed to complete study tasks if screening tasks took longer than anticipated.

Figure 3.2. Portable play set for Experiment 1 language sample task.

Language samples were transcribed using C-unit segmentation (Loban, 1976). C-unit segmentation is a transcription process that follows syntactic rules, which can be used to identify utterances in both spoken and text-based language data. A C-unit is structurally defined as “an independent clause and its modifiers” (Loban, 1976, p. 9). Mazes were excluded from analysis.

All language samples were transcribed in CLAN software (MacWhinney, 2000) using the CHAT transcription and coding format (MacWhinney, 2000). The CHAT format is the preferred transcription and coding format used for transcribing and coding language sample data contributed to the CHILDES database (MacWhinney, 2000) after analysis. CLAN software is a program used for transcribing and analyzing spoken language samples in the CHAT format. CLAN software was used to create time-stamped linkages between video recorded language samples data and corresponding text transcripts.

Following Rispoli et al. (2009), non-spontaneous utterances were excluded from morpheme analysis. Rispoli et al. (2009, p. 934) define non-spontaneous utterances as “immediate imitations of adult, self, songs, counting, etc.” or other instances in which there is evidence that a child did not generate an utterance. Rispoli et al. (2009) argued that instances of morpheme use activated through associative mechanisms can inflate measures of language performance above a child’s true developmental level. Their measures of spontaneous tense marker use are intended to capture instances of tense marker use in formulated word combinations and exclude instances of tense marker use that are directly activated through associative connections. The conservative step of excluding non-spontaneous utterances from analysis helps filter out instances of tense marker use activated through associative connections.

Each transcript began at the beginning of the child’s recorded language sample and continued until the child produced 150 spontaneous novel utterances with at least two morphemes. These 150 multi-morpheme child utterances were selected for analysis. Gladfelter and Leonard (2013) found that tense marker total and productivity score measures could differentiate between typically developing 4-5 year olds and those with specific language impairment in language samples with a controlled number of total utterances. They argued that discriminability of these measures may be specific to sample sizes that approximate their samples of 152 utterances because these measures are totals influenced by each new instance of a scorable morpheme. Larger samples may suppress evidence of developmental patterns if participant performance approaches ceiling levels.

The transcribed samples of 150 multi-morpheme child utterances were analyzed using CLAN. The mor program in CLAN was used to automatically add a %mor tier to the transcribed samples with tags on each morpheme. The morpheme tags for tense markers were then manually reviewed to verify the automatically generated tags and make corrections as needed. A series of searches were then run through CLAN to extract all child utterances containing tags corresponding to each of the 15 different tense markers.

Mean syntactic length (MSL) was obtained as a measure of utterance length and used to calculate a predicted mean length of utterance in morphemes (predicted MLUm). MSL is defined as the average number of morphemes per utterance with single-morpheme utterances excluded (Klee & Fitzgerald, 1985). MSL could be obtained from utterances included in analysis even though single-morpheme utterances were excluded (Kovacs & Hill, 2017). Kovacs and Hill (2017) found that MSL predicted MLUm in typically developing children in the single word stage and Brown's Stages I-V. Predicted MLUm was found for each participant using Kovacs and Hill's (2017) regression formula.

Tense marker use was reported for each individual tense marker as a binary variable indicating whether or not the child used each individual tense marker at least once in the language sample. Hadley and Short (2005) considered a single grammatically correct, spontaneous production of a given tense marker in a language sample to be sufficient to be evidence that that tense marker was at least emergent. Overregularizations of *–ed* were counted as instances of correct *–ed* use. Contractions to pronominal subjects were not counted because these high frequency patterns are likely to be directly activated through associative connections (Hadley & Short, 2005). The irregular forms *don't* and *ain't* were not counted because they are specifically negative (Rispoli & Hadley, 2011). *Tense marker total* (Hadley & Short, 2005; Rispoli & Hadley, 2011) was reported as the total number of tense markers the child used at least once in the language sample (range: 0-15).

Category productivity scores were found to characterize the productivity of each morpheme category. The number of sufficiently different uses of tense markers in each morpheme category was counted using Hadley and Short's (2005) operational criteria for sufficiently different uses of each morpheme category. For *-ed* and *-3s*, sufficiently different uses were counted for each lexical verb that was correctly inflected using the tense marked suffix, including overregularizations. For COPULA BE, AUXILIARY BE, and AUXILIARY DO, sufficiently different uses were counted for each different subject-tense marker combination that the child generated. Contractions to pronominal subjects and the specifically negative forms *don't*, *ain't* were not counted (Hadley & Short, 2005; Rispoli & Hadley, 2011). The first 5 sufficiently different uses were counted towards the category productivity score for each morpheme category. Additional uses were not counted. A composite *productivity score* (range: 0-25) was found by calculating the sum of the five category productivity scores.

Inter-judge agreement for language sample transcription was achieved using transcription-by-consensus (Hill, Kovacs, & Shin, 2014). Inter-judge agreement was found for language samples from two randomly selected children ages 30-42 months and two children randomly selected children ages 43-54 months. A two-step process was used to achieve 100% consensus for utterance transcription and tense-marked morpheme tagging. In the first step, consensus was achieved for transcription of utterances in the selected language samples. A second, trained rater independently reviewed video recordings of these language samples and the speaker tier of linked transcripts generated by the author. The second rater independently made corrections on the speaker tier of the author's transcripts when she found discrepancies between the author's transcripts and the words spoken in the actual recordings. Once this was completed, the author and second rater met and resolved all transcription discrepancies by consensus. Morpheme tags in the final transcripts were reviewed using a similar process to achieve a consensus on morpheme tagging.

3.3 EXPERIMENT 2: ASSESSING STIMULABILITY IN TYPICALLY DEVELOPING CHILDREN

3.3.1 Purpose and research questions

The purpose of Experiment 2 was to characterize the effects of age and communication modality on tense marker stimulability and tense and agreement morpheme category stimulability in typically developing children acquiring English as a first language. The children who participated in Experiment 1 were included as participants in Experiment 2. A series of structured tasks were used to assess participants' ability to produce each tense marker following an adult model in either the spoken modality or the graphic symbol modality. In both conditions, the adult model provided a highly invasive maximal prompt demonstrating correct tense marker use in the target communication modality. Tasks were adapted from known procedures for eliciting spoken tense marker productions from verbally speaking children (Dawson et al., 2005; Hughes & Till, 1982; Leonard et al., 2003; Rice & Wexler, 2001; Wilson & Fox, 2015b).

Participants were randomly assigned to complete stimulability tasks for tense markers in each morpheme category in either the spoken modality or the graphic symbol modality. Stimulability was indicated when children demonstrated correct tense marker production in the target communication modality following input demonstrating correct tense marker use in that modality. In the spoken modality, participants used verbal speech to generate responses and the researcher used verbal speech to demonstrate correct morpheme use. In the graphic symbol modality, participants used custom vocabulary pages that isolated target verbs and inflections or subjects and tense markers on a touch-screen based speech generating device (SGD) to generate responses. Explicit modeling in the graphic symbol modality was used to familiarize participants to each graphic symbol in isolation and to demonstrate correct morpheme use in the graphic symbol modality. These tasks were used to obtain measures of tense marker and morpheme category stimulability. Experiment 2 sought to answer three exploratory research questions.

Research question 2A asked if the odds of a child being stimulable for each individual tense marker increased at the same rate across communication modalities. The null hypothesis for any given tense marker was that the odds of a child being stimulable for that tense marker would increase at the same rate across modalities and there would be no significant age-by-modality interaction for tense marker stimulability. In contrast, the alternative hypothesis for any given tense marker was that the odds of a child being stimulable for that tense marker would increase at different rates across modalities and a significant age-by-modality interaction for tense marker stimulability would be found.

Research question 2B asked if morpheme category stimulability increases at the same rate across communication modalities for each morpheme category. The null hypothesis for each morpheme category was that category stimulability would increase at the same rate across modalities and there would be no significant age-by-modality interaction for category stimulability. In contrast, the alternative hypothesis for each morpheme category was that category stimulability would increase at different rates across modalities and a significant age-by-modality interaction would be found.

Findings consistent with the null hypotheses for question 2A would indicate that tense marker stimulability increases at the same rate across communication modalities. Findings consistent with the null hypothesis for question 2B would indicate that morpheme category stimulability grows at the same rate across communication modalities. These findings would provide evidence that growth of morpheme stimulability is not modality-specific or limited to unaided communication strategies. If this is the case, then stimulability testing could potentially be used as a strategy for assessing expressive morphology skills in children who use different communication modalities.

Findings indicating that these null hypotheses should be rejected would indicate that morpheme stimulability grows at a faster rate in one modality than the other. These findings would provide evidence that growth of morpheme stimulability is modality-dependent in some way. If this is the case, it may still be possible to use stimulability testing strategies to assess expressive morphology skills in children who use different communication modalities.

However, it would be necessary to consider factors that contribute to modality-specific patterns of stimulability growth.

The author is not aware of any prior studies testing a null hypothesis that developmental language skills differ across spoken and graphic symbol modalities. Therefore, the author assumed the null hypotheses for questions 2A and 2B because there was no established theoretical basis to motivate the expectation that results would be consistent with an alternative hypothesis or that an age-by-modality interaction would be found for measures of stimulability.

A main effect of age was expected for measures of tense marker stimulability and morpheme category stimulability. Stimulability was expected to increase with age in all analyses.

Main effects of communication modality were considered plausible for measures of tense marker stimulability and morpheme category stimulability. If a main effect of communication modality was found, stimulability was expected to be higher in the spoken modality that children are most familiar with.

Research question 2C asked if morpheme category stimulability increases at the same rate across morpheme categories. This question tested the null hypothesis of the OI theory, which predicted that the 5 categories would grow together with similar developmental trajectories. Under this prediction, there would be no significant main effects of morpheme category on category stimulability and no significant age-by-morpheme category interactions.

In contrast, the alternative hypothesis of GML theory predicted that morpheme categories that differ at the level of positional processing would grow in stimulability at different rates. Specifically, GML theory predicted that category stimulability for COPULA BE would grow faster than any other category, followed by a wave of three categories growing together

(AUXILIARY DO, -*3s*, -*ed*), followed by AUXILIARY BE. This would lead to significant main effects of morpheme category on morpheme category stimulability and significant age-by-morpheme category interactions.

3.3.2 Procedures.

Stimulability was assessed during Visit 2. Tasks for testing stimulability of all 15 tense markers in the spoken and graphic symbol modalities were developed by adapting known procedures for eliciting spoken productions of specific tense markers from verbally speaking children (Dawson et al., 2005; Hughes & Till, 1982; Leonard et al., 2003; Rice & Wexler, 2001; Wilson & Fox, 2015b). These procedures were adapted to assess participants' ability to produce tense markers in a target communication modality following adult models of correct tense marker use in the same modality. In the analyses of tense marker stimulability for individual tense markers, participants were not penalized for a failure to use a tense marker without exposure to at least one model of correct tense marker use in the target communication modality.

At least one task was used for testing stimulability of tense markers in each category, as summarized in Table 3.2. Two separate tasks were used for testing the stimulability of present tense and past tense forms of COPULA BE and AUXILIARY BE because different elicitation tasks are typically used to elicit spoken productions of present and past tense forms of BE (Garrity & Oetting, 2010; Leonard et al., 2003). The tasks for present tense and past tense COPULA BE were completed back-to-back so that all items corresponding to the COPULA BE category would be presented as a block. The tasks for present tense and past tense AUXILIARY BE were completed back-to-back so that all items corresponding to the AUXILIARY BE category would be presented as a block.

Table 3.2. Summary of stimulability tasks

Task	Category	Marker(s)	Probes	Vocabulry Page[a]	Materials	General Procedure	Reference(s)
1	COP BE	*am*, *is*, *are*	6	Present tense BE	picture stimuli, small box	syntactic cloze procedure	Hughes & Till, 1982
2	COP BE	*was*, *were*	4	Past tense COP BE	picture stimuli	Syntactic cloz procedure: Present "after" & "before" pictures. Child describes "before" picture.	Dawson et al., 2005
3	-*3s*	-*3s*	5	-*3s* and –*ed*	picture stimuli, including existing stimuli from Wilson & Fox (2015)	syntactic cloze procedure	Rice & Wexler, 2001
4	-*ed*	-*ed*	5	-*3s* and –*ed*	picture stimuli, including existing stimuli from Wilson & Fox (2015)	syntactic cloze procedure	Rice & Wexler, 2001
5	AUX DO	*does*, *do*, *did*	6	AUX DO	picture stimuli, one puppet (Sidney the Penguin)	Shy puppet activity w/ one puppet & picture stimuli	Rice & Wexler, 2001
6	AUX BE	*am*, *is*, *are*	6	Present tense BE	picture stimuli, small box	syntactic cloze procedure	Hughes & Till, 1982
7	AUX BE	*was*, *were*	4	Past tense AUX BE	video stimuli, one puppet (Sleepy Bunny)	Sleepy puppet activity w/ one puppet & video stimuli	Leonard et al., 2003

Note. COP = COPULA; AUX = AUXILIARY; Probes = number of probe items in task.

[a]Vocabulary Page column identifies SGD vocabulary pages used for each task in the graphic symbol modality.

3.3.2.1 Practice and probe items.

Each task consisted of one practice item, followed by a block of 4-6 probe items as summarized in Table 3.2. General procedures for practice and probe items are outlined in Appendix B. Two consecutive probe items were used to probe each tense marker in the AUXILIARY DO, COPULA BE, and AUXILIARY BE categories. Auxiliary *do, does, did* and copula and auxiliary *is, are, was, were* were each probed in two sufficiently different contexts with different subject-tense marker combinations. Copula and auxiliary *am* only license one subject-tense marker combination (*I am*), which was probed twice. Five consecutive probe items probed *-3s* and *–ed* in sufficiently different contexts requiring inflection of different lexical verbs.

At least 5 items probed each category in sufficiently different contexts. A total of 5 items in one task probed tense markers in the *-3s* and *–ed* categories. A total of 6 items in one task probed tense markers in the AUXILIARY DO category. A total of 10 items across two consecutive tasks probed tense markers in the COPULA BE and AUXILIARY BE categories.

Each practice and probe item included visual stimuli (picture or video) and a verbal cue. Use of manipulatives was minimized. Puppets were used to probe AUXILIARY DO and auxiliary *was* and *were* as they are routinely used in comparable elicitation tasks (Leonard et al., 2003; Rice & Wexler, 2001).

Each practice item provided input demonstrating correct use of a tense marker that would be probed in the upcoming task. Practice items established the communication modality and task for an upcoming set of probe items in a consequence-free environment with no penalty for incorrect responses. Practice items were presented in a similar format to probe items in the upcoming task. However, the author attempted to elicit a repetition of the correct response if a participant did not spontaneously generate a correct response or independently repeat a modeled production during presentation of the practice item. Repetition of the correct response was explicitly reinforced by praising the participant and repeating the participant's utterance, "Good job! I like how you said _____."

Probe items were structured in a way that required participants to spontaneously generate responses in the target modality. A flow chart for the presentation of probe items is shown in Figure 3.3. The author presented the visual stimuli and verbal cue corresponding to an item and gave a 10 second expectant delay waiting for the participant to spontaneously generate a response. Additional wait time was given if a participant was clearly still formulating a response. The cue and wait time were repeated once if the participant did not respond. If the participant generated a verbal response when the target communication modality was the graphic symbol modality, the child was prompted to "show me how you say that with the device." This was followed by a 10 second expectant delay. If the participant generated a correct response, the response was reinforced with positive reinforcement as in the practice item. If the participant generated an incomplete or incorrect response or did not respond after the repeated cue, the cue was repeated and the correct response was produced in the target communication modality. These items were scored as incorrect. For the AUXILIARY DO and auxiliary *was*/*were* tasks, pragmatically appropriate cues were given instead of a repetition of the initial verbal cue, as discussed in the section on stimulability tasks. Participants were allowed to repeat the modeled production of the correct response if their repetition was self-initiated. Repetition of the correct response was reinforced with positive reinforcement.

Items for copula and auxiliary *is, are, was,* and *were* and auxiliary *does, do,* and *did* were designed so participants were required to spontaneously generate a correct response by

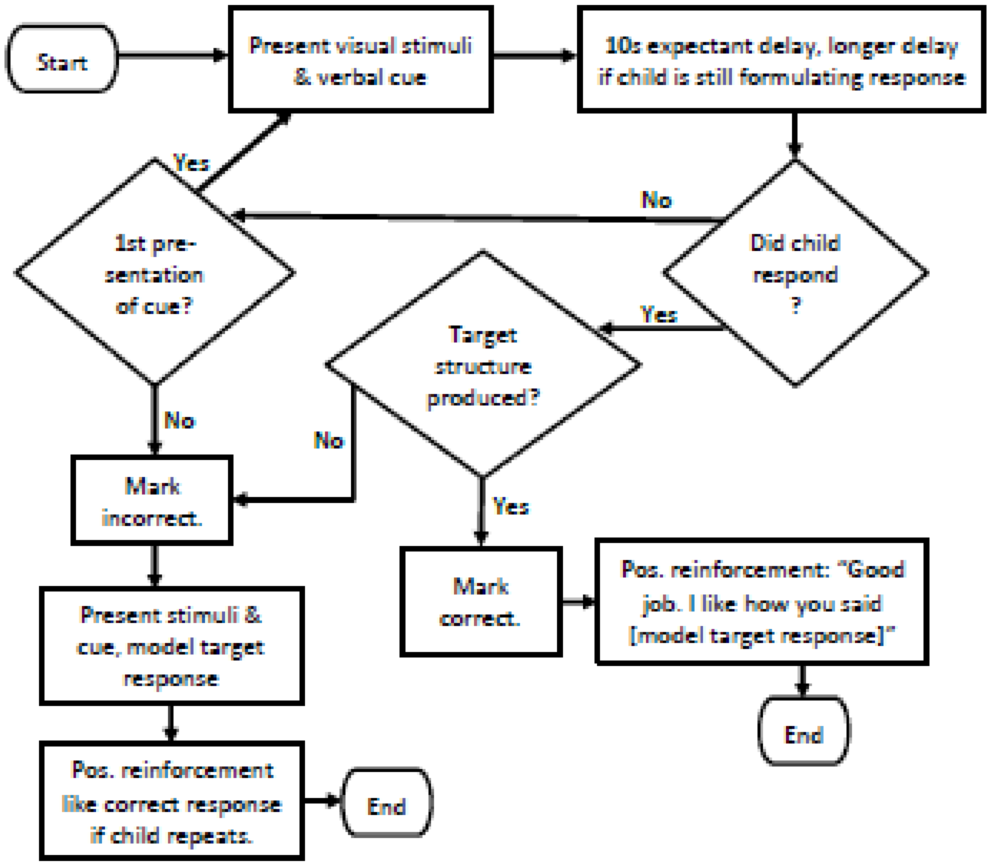


Figure 3.3. Presentation of stimulability probe item in spoken modality.

combining a subject noun phrase with a target tense marker in the correct modality. The subjects were singular animal characters (*the bird*, *the moose*), plural animal characters (*the bears*, *the zebras*), singular inanimate objects (*the wagon*, *the house*) and plural inanimate objects (*the cars*, *the trucks*). Test items were designed to minimize the use of subject pronouns (*he*, *she*, *it*, *they*) in participant responses because high-frequency pronoun-tense marker combinations may be stored in associative memory as a whole unit. For example, the probe items for all forms of COPULA BE and auxiliary *is*, *am*, and *are* used grammatical cloze procedures in which the article *the* was supplied as part of the examiner's verbal cue to prime participants to respond using a noun phrase.

For any given tense marker, each probe item required the participant to produce the target tense marker in a different context. For example, the two probe items for for copula *is* probed copula *is* in the context of two different low-frequency subject noun phrases (*the bird is*, *the moose is*). After the participant's response was scored on the first copula *is* probe item, the author modeled production of the target response (*the bird is*). The second probe item required the participant to spontaneously generate a different response (*the moose is*) with no immediate prior exposure to this exemplar in the target modality. This use of multiple exemplars for morpheme stimulability tasks differs from the elicited imitation procedures used to probe phoneme stimulability in established assessment protocols (e.g., Goldman & Fristoe, 2015).

3.3.2.2 Communication modalities.

Each participant completed the stimulability tasks for all 5 morpheme categories. Each participant was randomly assigned a target communication modality (spoken or graphic symbol) for each morpheme category. Random assignment was completed without replacement for each morpheme category so that 10 participants ages 30-42 months and 10 participants ages 43-54 months were assigned to each communication modality in each morpheme category. All responses in both modalities were recorded using the score sheet in Appendix C. Video recordings were used to verify the accuracy of online data coding as needed.

In the spoken modality, participants used verbal speech to generate responses. The researcher used verbal speech for all cues and for demonstrating correct morpheme use. All tasks for morpheme categories assigned to the spoken modality were completed as a block before any tasks assigned to the graphic symbol modality were completed. Tasks in the spoken modality were completed in the numbered order shown in the Task column of Table 3.2.

In the graphic symbol modality, participants used an experimental SGD application program on an iPAD Mini 4 platform to generate responses. A series of 5 custom-made vocabulary pages with 15 key locations each was used to assess stimulability of all tense markers in the graphic symbol modality. Each task was completed using a single vocabulary page, as summarized in Table 3.2. Participants were not expected to navigate across multiple pages to demonstrate stimulability of any given tense marker. Some vocabulary pages were used for multiple tasks. All vocabulary pages are shown in Appendix D, with text glosses added to each symbol location for readability. The pages presented to the participants did not include text glosses.

Keys on each page were arranged in a grid with 3 rows and 5 columns. All words and morphemes necessary to selectively generate correct responses to each corresponding practice and probe item were represented on these locations by single meaning pictures with no text gloss. On some pages, some symbols were used to represent 2-morpheme subject noun phrase, such as “the bird.” All other symbols will be used to represent 1 free morpheme or 1 bound morpheme. A custom set of single meaning pictures was used to represent each of the 15 tense markers graphically, as shown in Table 3.3. These symbols were designed so that a rationale could be used to explain the association between the symbol and corresponding tense marker. The *-3s* and *–ed* symbols are used in an existing set of sequenced multi-meaning icons (Baker, 1986) with a text gloss. Each page included symbols for multiple contrasting tense markers. Participants were always required to selectively generate a target tense marker as part of a correct response. For example, participants were required to differentiate between three possible forms of AUXILIARY DO.

In the graphic symbol modality, the target response on each item required participants to select at least two of the available graphic symbols, with one symbol representing a subject or verb stem and a second symbol representing a tense marker. The experimental SGD was configured to speak and add text to a message window as symbols were selected. The message window was configured to trigger the SGD to speak all displayed text when it was selected. This allowed participants to produce longer messages after selecting multiple symbols. One location of each vocabulary page was be used as a "clear display" control key, which was used to clear the contents of the message window. Participants were allowed to select the clear display control key if they made an error and then begin generating their intended response a second time. Other locations included single meaning pictures representing language content necessary for generating correct responses on a corresponding task. If all necessary locations for a task were used, any remaining locations were left blank.

The researcher provided explicit models of targets in the graphic symbol modality throughout the administration of stimulability tasks in the graphic symbol modality, consistent with Binger and Light's (2008) suggestion that strategies used for assessing expressive morphology in verbally speaking children can be adapted for AAC speakers by incorporating

Table 3.3. Single meaning pictures corresponding to 15 English tense markers

Category	Marker	Symbol	Rationale
COP BE AUX BE	is is		The bee IS in the honey jar. The bee IS flying in the honey jar.
COP BE AUX BE	am am		I AM in the honey jar. I AM sitting in the honey jar.
COP BE AUX BE	are are		The bees ARE in the honey jar. The bees ARE flying in the honey jar.
COP BE AUX BE	was was		The bee WAS in the honey jar. The bee WAS flying in the honey jar.
COP BE AUX BE	were were		The bees WERE in the honey jar. The bees WERE flying in the honey jar.
AUX DO	does		I DO want to get married.
AUX DO	do		He DOES want to get married.
AUX DO	did		We DID get married.
-3S	-3s		The bucket has [+Agr3S, -Past] water in it.
-ED	-ed		The bucket had [+Past] water in it.

aided language stimulation. Any time models were provided in the graphic symbol modality, the researcher selected graphic symbols and used the experimental SGD to produce speech output. These models were used to familiarize participants with the pre-stored vocabulary on each page at the single word level and to model correct use of tense markers in obligatory contexts.

Each time a new task was started in the graphic symbol modality, aided language stimulation was used to familiarize the participant with the pre-stored vocabulary on the current page at the single-word level. The researcher elicited imitations of the words corresponding to each new symbol in isolation by prompting the participant to select the symbol and generate speech output. The researcher also verified that participants recalled familiar symbols from previous vocabulary pages. The familiarization procedure was not repeated if the same vocabulary page was used for two consecutive tasks.

This is an established familiarization process. Sutton and Morford (1998) used aided language stimulation to model each symbol on an array of 20 single meaning pictures once for typically developing kindergarteners. After this rapid familiarization process, these kindergarteners used their array of 20 single meaning pictures to generate simple sentences in an elicitation task (Sutton & Morford, 1998). Similar procedures have been used to familiarize preschoolers to smaller displays before eliciting repetitions of subject-verb-object sentences (Sutton, Trudeau, Morford, Rios, & Poirier, 2010).

The researcher explained and modeled correct use of the message window and clear display control key while familiarizing the participant to the first page by making and correcting an error during utterance generation. Participants were not evaluated for their accuracy using the clear display control key or message window.

The researcher used verbal speech for scripted verbal cues and explicit models of graphic symbol use for demonstrating correct morpheme use in responses. All tasks for morpheme categories assigned to the graphic symbol modality were completed as a block in the numbered order shown in the Task column of Table 3.2.

Stimulability tests from two children ages 30-42 months and two children ages 43-54 months were randomly selected for inter-rater reliability testing. A second, trained rater independently reviewed video recordings of these selected stimulability tests and independently scored each test item as either correct or incorrect. Interrater reliability for scoring individual test items on the stimulability test was found using the kappa statistic (Cohen, 1960). Excellent agreement was found for scoring individual test items ($\kappa = 0.90$, $p < 0.001$).

3.3.2.3 Stimulability tasks.

This section defines procedures for tasks probing stimulability of all 15 tense markers in both communication modalities. Sample visual stimuli and scripts for one probe item in each task are provided as examples. During the study, target responses were produced in the target communication modality.

All stimulability tasks were completed in a single session during Visit 2. Visual stimuli were presented in a PowerPoint presentation on a laptop computer. The laptop was positioned on a child-sized chair. The participant was asked to sit on a second child-sized chair facing the laptop. The author presented visual stimuli and cues while sitting next to the child and facing the laptop. The session was recorded using a stationary video camera. Some children moved around the testing room during the session and were not within view of the camera for all items. In these cases, audio was still recorded by the camera.

Stimulability tasks for the COPULA BE category.

Different tasks are used to elicit past and present tense forms of BE (Garrity & Oetting, 2010; Leonard et al., 2003). Therefore, two separate tasks and vocabulary pages were used for probing stimulability in the COPULA BE category. These two tasks were presented back-to-back in one communication modality. All items probed stimulability of uncontractible forms, which are typically acquired before contractible forms (Brown, 1973; Kuczaj, 1979). Copula *am*, *is* and *are* are uncontractible in utterances containing grammatical ellipsis (Bloom & Lahey, 1978; Hughes & Till, 1982), such as the sentence-final *is* in (14).

14. The moose is not on the chair, but the bird is

Copula *was* and *were* are always uncontractible. This focus on uncontractible forms is consistent with the measures of spontaneous tense and agreement morpheme use in Experiment 1, which exclude contractions to pronominal subjects.

Present tense forms of COPULA BE (*am, is, are*) were probed in a copula *am, is, are* task consisting of 1 practice item followed by a block of 6 probe items, two for each tense marker. Initial instructions for participants completing this task in the spoken modality were "We are going to talk about some pictures. Let's look at the first picture." Initial instructions for participants completing this task in the graphic symbol modality were "We are going to use the device to talk about some pictures. Let's look at the first picture." Participants completing the copula *am, is, are* task in the graphic symbol modality used the present tense BE vocabulary page shown in Appendix D.

Copula *am*, *is* and *are* were probed using a syntactic cloze procedure adapted from Hughes and Till (1982). For each item probing copula *is* and *are*, participants were shown a photograph of two animal puppets or two groups of animal puppets and an environmental object, such as a chair. The puppets were positioned in contrasting locations relative to the environmental object so that the different locations of the puppets could be described using contrasting statements and one preposition. The author identified the puppets in the picture before presenting a verbal cue in the form of a cloze statement with grammatical ellipsis. A sample copula *is* stimulability probe item is shown in Figure 3.4. Each of these items required the child to use a different nominal subject.

Items probing copula *am* followed the same format. However, the author created a contrast by manipulating the real-world location of a small box in relation to the participant (Hughes & Till, 1982), as in (15).

15. Verbal cue as researcher holds box over child's head: I am not under the box, but

    Target response: I am

Verbal cue: *(Point to moose)* This is the moose. *(Point to bird)* This is the bird. The moose is not on the chair, but the

____

Target response: bird is [on the chair]

Figure 3.4. Sample stimulability probe item for copula *is*.

Past tense forms of COPULA BE (*was, were*) were probed in a copula *was, were* task consisting of 1 practice item followed by a block of 4 probe items. Two items probed each tense marker in the context of different nominal subjects. Copula *was* and *were* were probed in simple declarative sentences using an adaptation of a single-item task used for eliciting copula *was* in the SPELT-P2 (Dawson et al., 2005). In this task, participants were shown a series of photographs of a puppet hippopotamus named Andy, who was a painter and several objects that Andy has painted. Initial instructions established that the task was to say what color the various objects were before Andy painted them. Initial instructions for children tested in the spoken modality were "This is Andy. Andy spent the whole weekend painting. He painted many things. Let's talk about the things Andy painted, and what colors those things were before." Initial instructions for children tested in the graphic symbol modality were "This is Andy. Andy spent the whole weekend painting. He painted many things. Let's use the device to talk about the things Andy painted, and what colors those things were before." Participants completing the copula *was, were* task in the graphic symbol modality used the past tense COPULA BE vocabulary page in Appendix D.

For each copula *was/were* item, the participants were shown a drawing of Andy the painter and an object or group of objects that he has finished painting in a new color. The participants were then shown a picture of the object in its color (or group of objects in their original color). The verbal cue used a syntactic cloze procedure to cue the child to describe what color the object was before (or what color the objects were before). A sample copula *was* stimulability probe item is shown in Figure 3.5.

Verbal cue (Picture 1 on first slide): Andy painted the house red. Let's see what color the house was before.

Verbal cue (Picture 2 on second slide):
The ______________

Target response: [The] house was yellow

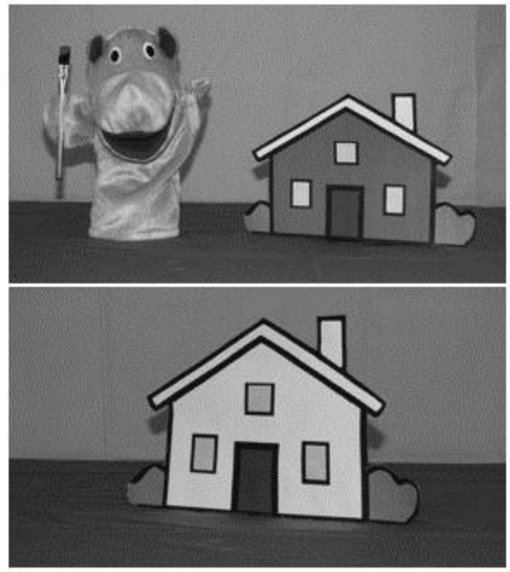


Figure 3.5. Sample stimulability probe item for copula *was*.

Stimulability tasks for the -3s category.

The task for the *-3s* category included one practice item followed by a block of 5 consecutive probe items. Each item required participants to generate a different lexical verb inflected with the *-3s* suffix.

Initial instructions for the *-3s* task and verbal cues for practice and probe items in this category were adapted from the Third Person Singular Probe in the *Rice/Wexler Test of Early Grammatical Impairment* (TEGI; Rice & Wexler, 2001). Initial instructions were "I am going to show you some pictures and ask you to help me say what each person does." Participants completing this task in the graphic symbol modality were be directed to "Use the communication device to answer." Participants completing this task in the graphic symbol modality used the *-3s* and *-ed* vocabulary page shown in Appendix D.

Stimulus pictures for the *-3s* task consisted of an illustrated picture of a person completing an action. The picture stimuli included a mix of existing pictures from the *LanguageLinks to Literacy* program (Wilson & Fox, 2015b) and original illustrations created for this task. Verbal cues used a syntactic cloze procedure to cue the child to produce the *-3s* form of a lexical verb illustrated in a corresponding stimulus picture. A sample *-3s* stimulability probe item is shown in Figure 3.6.

Verbal cue: The man is a painter. What does the man do? The man ________

Target response: [The man] paints

Figure 3.6. Sample stimulability probe item for *-3s*.

Note. Stimulus picture used with permission from Wilson and Fox (2015a).

Stimulability tasks for the –ed category.

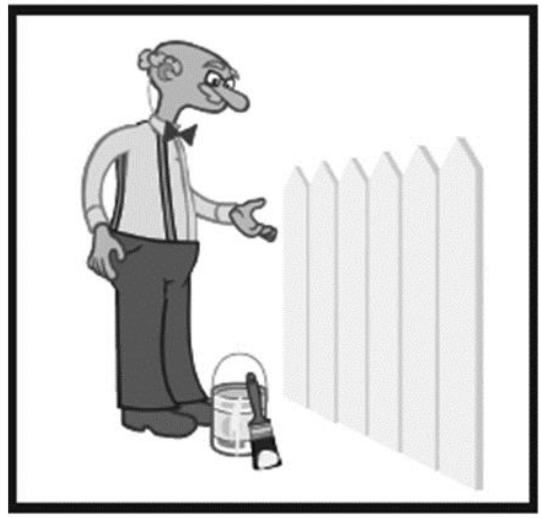
The task for the *–ed* category included one practice item followed by a block of 5 consecutive probe items. Each item required participants to generate a different lexical verb inflected with the *–ed* suffix. All target lexical verbs had regular past tense forms.

Initial instructions the *-ed* task and verbal cues for the corresponding practice and probe items were adapted from the Past Tense Probe of the TEGI. Initial instructions were “I have two pictures. I will describe the first one and you help me describe the second one.” Participants completing this task in the graphic symbol modality also were directed to “Use the device to answer.” Participants completing this task in the graphic symbol modality used the *-3s* and *-ed* vocabulary page shown in Appendix D.

Stimulus pictures for the *–ed* task consisted of a pair of illustrated pictures presented side-by-side. The first picture showed a person completing an action. The second picture showed the same person after the action was completed. The picture stimuli included a mix of existing pictures from the *LanguageLinks to Literacy* program (Wilson & Fox, 2015b) and original illustrations created for this task. Verbal cues used a syntactic cloze procedure to cue the child to produce the *-ed* form of a lexical verb illustrated in a corresponding pair of stimulus pictures. A sample *-ed* stimulability probe item is shown in Figure 3.7.

Stimulability tasks for the AUXILIARY DO category.

The AUXILIARY DO task included 1 practice item, followed by a block of 6 probe items. Two items probed each form of DO in contexts requiring different nominal subjects. All items probed stimulability of forms of DO in yes-no questions that require *do*-support. No items required negation.



Verbal cue (Pointing at first picture):
Here the man is painting.

Verbal cue (Pointing at second picture):
Now he is done. Tell me what he did. He _____.

Target response: [He] painted

Figure 3.7. Sample stimulability probe item for *-ed*.

Note. Stimulus pictures used with permission from Wilson and Fox (2015a).

The AUXILIARY DO task was adapted from the BE/DO probe on the TEGI. This task presented a pragmatic context that encourages participants to ask questions in the target communication modality instead of answering questions verbally. In this task, participants were introduced to a penguin puppet named Sidney and presented with a narrative scenario that encouraged them to ask Sidney questions: "His animal friends are visiting for dinner. His friends are a bird, a moose, two zebras, and two bears. Sidney's friends are all very shy. They will not speak to you or me because they only understand animal talk. They do not understand people talk. Fortunately, Sidney understands people talk. You can talk to Sidney and ask him questions." Participants tested in the graphic symbol modality also were told that "he will answer you if you use the communication device." Participants completing this task in the graphic symbol modality used the AUXILIARY DO vocabulary page shown in Appending D.

The researcher familiarized the participant to the task by asking Sidney a "do you" question using the target communication modality before the practice item was administered. The Sidney puppet was used to answer the researcher's question. This demonstration reinforced the pragmatic context for asking Sidney questions with *do* support. The participant then completed one practice item by asking Sidney a "do you" question and waiting for an answer.

After this practice item, a series of six probe items was presented. In each probe item, a stimulus picture showing one or two of Sidney's animal friends and a food item was presented. The child was then verbally cued to ask Sidney if the animal friend(s) in the picture wanted the food item shown. After the child asked the question, the Sidney puppet relayed the question to the animal friend(s) using nonsense words. An answer to the child's question was relayed back to the child using English. A sample probe item for auxiliary *does* is shown in Figure 3.8.

Verbal cue: The moose looks hungry. I wonder if the moose wants pizza. Ask Sidney if the moose wants pizza.

Target response: Does the moose want pizza?

Respond as Sidney: I'll find out. *(Sidney whispers to the moose, using whispering sounds, not words).* Yes. The moose wants pizza.

Figure 3.8. Sample auxiliary *does* stimulability item.

Similar “shy puppet” activities are frequently used to elicit productions of questions from young children (e.g. Ambridge, Rowland, Theakston, & Tomasello, 2006; Crain & Thornton, 2000; Rice & Wexler, 2001; Rowland & Theakston, 2009). Typically, the child is introduced to one or more shy puppets who will not speak to the child and one puppet who will answer the child’s questions about the shy puppet(s). The child is then cued to ask the puppet a series of questions. This classic formulation of the shy puppet task requires coordinated use of several manipulatives, which may be cumbersome. In Experiment 2, all manipulatives except the Sidney puppet were replaced with photographs so the researcher could focus on providing models of correct production.

If providing a model of correct production of the target for any AUXILIARY DO item was necessary, the researcher told the child “I will ask Sidney if _____” before modeling the target response. This was considered pragmatically appropriate for a task simulating a group conversation. This strategy was implemented after the first four children in the study disengaged from the task when the researcher repeated full verbal cues before modeling target responses.

Stimulability tasks for the AUXILIARY BE category.

As with the COPULA BE category, different tasks are needed to probe past and present tense forms of AUXILIARY BE (Garrity & Oetting, 2010; Leonard et al., 2003). Therefore, two separate tasks and vocabulary pages were used for probing stimulability in the AUXILIARY BE category. These tasks were presented back-to-back in one communication modality. All items probed stimulability of uncontractible forms, which are typically acquired before contractible forms (Brown, 1973; Kuczaj, 1979). Auxiliary *am*, *is* and *are* are uncontractible in utterances containing grammatical ellipsis (Bloom & Lahey, 1978; Hughes & Till, 1982), such as the sentence-final *is* in (16).

16. The bird is not sleeping, but the moose is

Auxiliary *was* and *were* are always uncontractible.

Present tense forms of AUXILIARY BE (*am*, *is*, *are*) were probed in an auxiliary *am*, *is*, *are* task consisting of 1 practice item followed by a block of 6 probe items, two for each tense marker. Initial instructions for this task were the same as the initial instructions for the copula *am*, *is*, *are* task. Participants completing the auxiliary *am, is, are* task in the graphic symbol modality used the present tense BE vocabulary page shown in Appendix D.

Auxiliary *am*, *is* and *are* were probed using a syntactic cloze procedure adapted from Hughes and Till (1982). For each item probing auxiliary *is* and *are*, participants were shown a photograph of two animal puppets or two groups of animal puppets. In each picture, one puppet or group of puppets was performing an action with a progressive aspect, such as flying. The second puppet or group of puppets was standing in a neutral position and doing nothing in particular. The author identified the puppets in the picture before presenting a verbal cue in the form of a cloze statement with grammatical ellipsis. An example of an auxiliary *is* probe item is shown in Figure 3.9. Each of these items required the participant to respond using a different subject noun phrase.

Verbal cue: *(Point to bird)* This is the bird. *(Point to moose)* This is the moose. The bird is not sleeping, but the ____________

Target response: moose is [sleeping]

Figure 3.9. Sample stimulability probe item for auxiliary *is*.

Items probing auxiliary *am* followed the same format. However, the author created a real-world contrast between what he was doing and what the participant was doing before presenting a verbal cue (Hughes & Till, 1982), as in (17).

17. Verbal cue as researcher stands next to sitting child: I am not sitting, but

Target response: I am

Past tense forms of AUXILIARY BE (*was*, *were*) were probed in an auxiliary *was*, *were* task consisting of 1 practice item followed by a block of 4 probe items. Two items probed each of these forms in contexts that require children to use different subject noun phrases in their responses.

A modified sleepy puppet activity (Leonard et al., 2003) was used to probe auxiliary *was* and *were*. In this task, participants were introduced to a Sleepy Bunny puppet. They were told that "Sleepy Bunny wants to watch some movies. Sometimes Sleepy Bunny falls asleep during movies. If he falls asleep, he needs you to tell him what was happening." The Sleepy Bunny puppet also directly asked the participant to remember the task in advance, with modality-specific information. In the spoken modality, the Sleepy Bunny puppet asked, "Will you tell me what was happening if I fall asleep?" In the graphic symbol modality, the Sleepy Bunny puppet asked, "Will you use the device to tell me what was happening if I fall asleep?"

For each item in the auxiliary *was*, *were* task, the participants and the Sleepy Bunny puppet watched a short movie together. The first 5 seconds of each movie featured a black screen. During this time, the Sleepy Bunny puppet fell asleep. This was followed by a short movie of one or two puppets performing an action with a progressive aspect (*drawing*, *singing*, *dancing*). The movie faded to a black screen before the action was completed, indicating that the action was still ongoing when the movie ended. After the movie ended the Sleepy Bunny puppet woke up and said “I fell asleep. Tell me what was happening.” A sample auxiliary *were* probe item is shown in Figure 3.10.

If providing a model of correct production of auxiliary *was* or *were* was necessary, the researcher asked the child if he could tell Sleepy Bunny what was happening before modeling the target response. This was considered pragmatically appropriate for a task simulating a group conversation with a series of events. This strategy was implemented after the first four children in the study disengaged from the task when the author replayed movies and repeated full verbal cues before modeling target responses.

Verbal cue (Speaking as Sleepy Bunny):
Let’s watch a movie about the bears.

(Watch movie of the bears dancing to instrumental music. Sleepy Bunny falls asleep at the beginning of the movie and does not wake up until the movie is finished).

Verbal cue (Speaking as Sleepy Bunny): I fell asleep. Tell me what was happening.

Target response: The bears were dancing.

Speaking as Sleepy Bunny: Oh. Thank you for telling me.

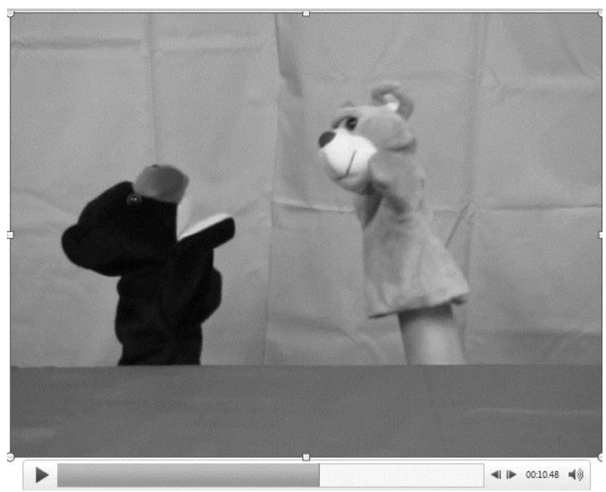


Figure 3.10. Sample stimulability probe item for auxiliary *were*.

3.4 DATA ANALYSIS FOR EXPERIMENTS 1 AND 2

Binomial logistic regression analyses and mixed effect generalized linear regression models were used for all analyses. Mixed effect generalized linear regression models with logistic link functions were used to answer questions 1A, 1B, 1C, 2B, and 2C. Binomial logistic regression models were used to answer question 2A. These analyses were used because they have several advantages over analyses of variance. First, they are useful for explaining the effects of predictor variables on a binary (categorical) regressor variable, including count data and binary response data such as “correct” or “incorrect” responses on test items (Barr, 2008; Jaeger, 2008). Second, they allow predictor variables related to time, such as age to be treated as continuous variables (Barr, 2008). Finally, mixed-effect modeling can be used to account for random effects of participants in analyses with repeated measures to improve the accuracy of inferences (Baayen, Davidson, & Bates, 2008).

These models used use the cumulative distribution function of the logistic distribution to estimate the probability that the response of a binary regressor variable will be positive as a function of a set of predictor variables (Myers, Montgomery, Vining, & Robinson, 2010). For analyses of individual tense marker stimulability (Question 2A), models estimated the probability that a given child would be stimulable for each tense marker. For all other analyses, models estimated the probability that a given child would give a positive response for any given observation and treated each tense marker, productivity point or stimulability probe item as a separate observation. Predicted probabilities of a positive result are a sigmoid function limited between 0 and1. Predicted probabilities asymptotically approach 0 as the probability of a positive result decreases and asymptotically approach 1 as the probability of a positive response increases. The odds of a positive response for any given observation are a ratio of the probability of a positive response to probability of a negative response for that observation. The odds of a positive response are multiplicative, unbounded, and vary as a function of the predictor variables.

Prior studies following Hadley and Short (2005) have treated each language sample as a single observation. In these prior analyses, one tense marker total (range 0-15), one composite productivity score (range 0-25) and one category productivity score for each category (ranges 0-5) was reported for each language sample. In the analyses for questions 1A and 2A, each tense marker was treated as a separate, binary observation. In the analyses for questions 1B and 1C, each productivity point was treated as a separate, binary observation. In the analyses for questions 2B and 2C, each probe item was treated as a separate, binary observation.

3.4.1 Question 1A.

Research question 1A asked if the number of tense markers children use in language samples with a fixed number of multi-morpheme utterances increases with age between 30 and 54 months. In order to answer this question, tense marker use was modeled using a mixed-effect generalized linear regression with a logistic link function. This model estimated the proportion of tense markers used in a sample of 150 multi-morpheme utterances as a function of age (at Visit 1). Age was entered into the model as a fixed covariate. Age was centered at 21 months following Rispoli et al.'s (2009) finding that the true zero for tense marker development occurred at 21 months. Random effects intercepts were entered for children.

The exponential function of the estimated coefficient (e^b) was used to interpret the relationship between age and the odds that a child would use any given tense marker. e^b is the odds ratio associated with a one-unit increase in the value of a predictor variable (Szumilas, 2010). For age, e^b is the odds ratio associated with a one-month increase in age.

3.4.2 Question 1B.

Research question 1B asked if productivity scores increase with age between 30 and 54 months in language samples with a fixed number of multi-morpheme utterances. In order to answer this question, productivity point level data were modeled using a mixed-effect generalized linear regression with a logistic link function. This model estimated the proportion of 25 possible productivity points that children received in a sample of 150 multi-morpheme utterances as a function of age (at Visit 1). Age was entered into the model as a fixed covariate. Age was centered at 21 months following Rispoli et al.'s (2009) finding that the true zero for tense marker development occurred at 21 months. Random effects intercepts were entered for children. e^b was used to interpret the relationship between age and the odds that a child would receive any given productivity point.

3.4.3 Question 1C.

Research question 1C asked if morpheme category productivity increases at the same rate across morpheme categories. In order to answer this question, productivity point-level data were analyzed using a mixed-effect generalized linear regression with a logistic link function. The model estimated the proportion of productivity points obtained in each morpheme category as a function of three fixed effects (morpheme category, age, and age-by-category interaction) and random effects intercepts for children. Morpheme categories were entered into the model using dummy variables. Age was entered into the model as a continuous fixed effect variable. The model was formed three times with age centered at different ages to examine differences between morpheme category productivity scores at different ages. Age was centered at 30 months, 42 months, and 54 months.

When significant fixed effects were found, the exponential function of the estimated coefficient (e^b) was used to interpret the relationship between the average odds that a child would receive any given productivity point in a morpheme category and values of the fixed effect variables. For age, e^b is the odds ratio associated with a one-month increase in age. For morpheme category, e^b is the odds ratio associated with a change between two morpheme categories.

3.4.4 Question 2A.

Research question 2A asked if the odds of a child being stimulable for each individual tense marker increased at the same rate across communication modalities. In order to answer this question, a series of multivariate binomial logistic regression analyses were used to estimate the effects of age and communication modality on the probability that a child will be stimulable for each individual tense marker. A separate set of regression models was estimated for each of the 15 individual tense markers. The set of models included a full model with 3 parameters (main effect of age, main effect of communication modality, and age-by-modality interaction), a two-parameter subset model with main effects of age and communication modality, age-only and modality-only subset models, and an intercept-only model. For each of these analyses, the unit of observation was the individual child participant. The regressor variable in each of these analyses was tense marker stimulability. Stimulable tense markers were coded as 1. Non-stimulable tense markers were coded as 0. The predictor variables in each analysis were age and communication modality. Age was centered at 42 months and entered into each model as a continuous variable. Communication modality was a categorical variable with two categories (graphic symbol, spoken). The model with the lowest *AICc* was selected as the best fitting model for each tense marker. Random effects of children were not included in these models because these models only used one observation per child.

e^b was used to interpret the relationship between the odds that a child would be stimulable for a given tense marker and values of the significant predictor variables in the final models. For age, e^b is the odds ratio associated with a one-month increase in age. For modality, e^b is the odds ratio associated with a change from the graphic symbol modality to the spoken modality.

Children were excluded from individual tense marker analyses casewise if they did not complete all probe items on the stimulability test corresponding to a given tense marker. One child was excluded from the analyses for copula *was*, *were*, *-ed*, auxiliary *does*, *do*, *did*, auxiliary *is*, *am*, and *are*. Three children were excluded from the analyses for auxiliary *was* and *were*.

3.4.5 Question 2B.

Research question 2B asked if morpheme category stimulability increases at the same rate across communication modalities for each morpheme category. In order to answer this question, probe item-level data were analyzed using a series of mixed-effect generalized linear regressions with a logistic link function. Each model estimated the proportion of correct responses on stimulability probe items corresponding to all tense markers in one morpheme category as a function of three fixed effects (communication modality, age, and age-by-modality interaction). A separate regression model was estimated for each morpheme category. Communication modality was entered into each model as a fixed effect with two categories (graphic symbol, spoken). Age was centered at 42 months and entered into each model as a continuous fixed effect variable. When significant fixed effects were found, e^{β} was used to interpret the relationship between the average odds that a child would give a correct response on any given probe item in a morpheme category and values of the fixed effect variables.

Each model also included random effects intercepts for children. Since communication modality was randomly assigned at the level of morpheme categories, any random effects of communication modality were nested within child participants.

Children were excluded from analyses casewise if they did not complete all probe items on the stimulability test corresponding to a given morpheme category. Three children were excluded from the AUXILIARY BE category analysis. One child was excluded from each of the other four category analyses.

3.4.6 Question 2C.

Research question 2C asked if morpheme category stimulability increases at the same rate across morpheme categories. In order to answer this question, probe item-level data were analyzed using mixed-effect generalized linear regression models with logistic link functions. A separate model was formed for each communication modality. The models estimated the proportion of correct responses on stimulability probe items corresponding to all tense markers in each morpheme category as a function of three fixed effects (morpheme category, age, and age-by-category interaction). Morpheme categories were entered into the models using dummy variables. Each model was formed three times with age centered at different ages to examine differences between morpheme category stimulability scores at different ages. Age was centered at 30 months, 42 months, and 54 months. When significant fixed effects were found, e^{β} was used to interpret the relationship between the average odds that a child would give a correct response on any given probe item in a morpheme category and values of the fixed effect variables.

Each model also included random effects intercepts for children. Since communication modality was randomly assigned at the level of morpheme categories, and each child was tested in at least one morpheme category in each modality, each child's data were split between models. The pattern of this split varied randomly from child to child.

Children were excluded from analyses pairwise if they did not complete all probe items on the stimulability test corresponding to a given morpheme category. Three children were excluded from the AUXILIARY BE category analysis in the graphic symbol model. One child was excluded from each of the other four category analyses in the graphic symbol model. No children were excluded from any analyses in the spoken model.

3.5 EXPERIMENT 3: ASSESSING PRODUCTIVITY AND STIMULABILITY IN PEDIATRIC AAC SPEAKERS

3.5.1 Purpose and research questions

The purpose of Experiment 3 was to use direct evidence of tense and agreement morpheme use and stimulability in pediatric AAC speakers with cerebral palsy to accomplish four tasks (18-21):

18. To establish pretest levels of tense and agreement morpheme use and stimulability prior to intervention

19. To ensure that each participant has access to an SGD with potential to support a fully productive tense and agreement system prior to intervention
20. To use these pretest measures to inform selection of a specific tense marker to target in an intervention program for each participant, and
21. To compare patterns of cross-morpheme generalization to predictions made by OI theory, GML theory, and the hypothesis that tense marker stimulability is a prognostic indicator of generalization.

Experiment 3 sought to answer the following research questions: Does intervention focused on increased production of a target tense marker in pediatric AAC speakers generalize across tense markers? If so, does stimulability testing predict patterns of cross-morpheme generalization? These research questions were addressed using a pretest posttest design to test for cross-morpheme generalization in the full set of fifteen tense markers. This design tested a null hypothesis that there would be no evidence of cross-morpheme generalization.

The assessment process used in Experiment 3 followed the basic structure of the process used by Powell et al. (1991) for selecting specific phonemes to target in intervention based on a structured inventory of pretest skills and task-specific measures of stimulability. Measures from Experiments 1 and 2 were obtained to establish pretest levels of tense marker use and stimulability. Spontaneous language samples and an elicitation probe were used to establish pretest levels of independent tense and agreement morpheme use. Then, the graphic symbol-based stimulability tasks from Experiment 2 were administered to establish pretest levels of tense and agreement morpheme stimulability. In addition, a grammaticality judgment task was used to assess pretest tense marker comprehension. The combined results from these tasks were used to rate the productivity and stimulability of each morpheme category and differentiate between tense markers that were at least emergent, tense markers that were stimulable, and non-stimulable tense markers.

The pre-stored vocabulary on each participant's SGD was reviewed to assess the SGD's potential for supporting productive tense and agreement morpheme use. If necessary, modifications would have been made so the SGD could support use of all tense markers and a fully productive tense and agreement morpheme system.

The researcher traveled to visit participants and collect all pretest, intervention, and posttest data during an intensive ten-day period. A comprehensive pretest assessment was conducted to characterize the participants' pretest development of all tense markers and tense marker categories. At the end of this assessment, a target tense-marker was identified and the author developed stimulus materials for intervention. Then, an intensive course of intervention focused on increasing target tense marker production was provided for up to five days with a frequency of two treatment sessions per day. Finally, a series of posttest generalization measures were collected to assess change in the participant's tense and agreement system.

An analysis of within-participant patterns of cross-morpheme generalization tested three theoretically-driven hypotheses about cross-morpheme generalization (22-24):

22. H_0: Cross-morpheme generalization is not possible.

23. H_1: Cross-morpheme generalization occurs between morphemes with similar features.

24. H_2: Individual patterns of cross-morpheme generalization can be predicted by morpheme stimulability tests.

The OI theory assumes null hypothesis H_0. Cross-morpheme generalization of learned skills cannot occur in the purely maturational system of OI theory. Under this null hypothesis, cross-morpheme generalization was not expected, even if increased production of the target tense marker was observed. Under the OI theory tense and agreement system growth happens as the unique checking constraint withers. In strong interpretations of OI theory, this withering is driven by a purely maturational process. If treatment encourages the unique checking constraint to wither at a faster rate, production of all tense markers should improve. Unbounded generalization across all tense markers would be consistent with a null hypothesis and this weaker interpretation of OI theory.

In contrast, GML theory assumes alternative hypothesis H_1. In a GML system, treatment effects were expected to generalize across morphemes with similar features. For example, treatment of copula *is* may generalize to other third-person present tense singular tense markers (*-3s*, auxiliary *is*) (Rispoli, 2016).

Alternative hypothesis H_2 assumes that morpheme stimulability is a prognostic indicator of generalization. Phoneme stimulability is a known prognostic indicator of generalization: treatment effects generalize to stimulable phonemes more than non-stimulable ones (Rvachew, 2005; Rvachew & Nowak, 2001). If morpheme stimulability is prognostic, then stimulability tests administered during the pretest assessment process would predict within-participant patterns of cross-morpheme generalization.

3.5.2 Participants and recruitment.

Two pediatric AAC speakers who met the inclusion criteria shown in (25) participated in Experiment 3.

25. Pediatric AAC speaker participant inclusion criteria:

    a. The participant is a child or adolescent with a diagnosis of cerebral palsy who uses an SGD as a primary modality of expressive communication.

b. The referring speech-language pathologist reports that the participant's expressive language skills are in the range of Brown's Stage III, IV or V.

c. The participant has had a hearing screening or hearing evaluation within the last 24 months, which indicated that the child had normal hearing or no more than mild hearing loss.

d. The participant has adequate central vision to identify graphic symbols on the display of their own AAC system and to focus on objects in a picture book.

e. The participant has no known difficulties with color perception.

f. The participant uses a language application program that includes sequenced multi-meaning icons and provides access to grammatical keys for production of grammatical morphemes.

g. The participant's SGD includes an automated datalogging feature, such as LAM, that can be used to automatically record language sample data.

h. The participant's parents agree to withdraw their child from any other speech-language therapy services for the duration of the study.

i. The participant's parents report that he or she understands at least 306 words on the *MacArthur-Bates Communicative Development Inventory-Words and Sentences* (CDI Words and Sentences; Fenson et al., 2007). This is consistent with the average number of words produced by typically developing 2 year olds in a normative sample (Fenson et al., 2007).

j. Item-analysis of the Early LAMbaseline indicates that the participant's parents believe they have pretest comprehension of tense-marked forms of BE and DO.

k. The present levels of communication skills in the participant's most current IFSP or IEP indicate that he or she produces at least simple subject-verb-object sentences using spontaneous novel utterance generation.

Children with cerebral palsy were selected as participants for two reasons. First, a clinical need for language-based AAC intervention programs to assist children with cerebral palsy exists. Many of these children present with severe motor speech disorders and require an AAC system to help them meet their expressive communication needs. Second, children with cerebral palsy often present with relative strengths in language or a relatively intact language system. For example, the AAC speakers with cerebral palsy assessed by Redmond and Johnston (2001) demonstrated robust receptive knowledge of grammatical morphology in a grammaticality judgment task.

Although a child with cerebral palsy may present with relative strengths in language, it would be inappropriate to assume that such a child’s expressive language skills are fully intact without directly measuring the child’s performance using language skills expressively. The assessment process used in Experiment 3 was designed to inform selection of intervention goals based on a model of typical language development; it focused on a specific morphosyntactic system that is known to be selectively disordered in verbally speaking children with developmental language impairments (e.g., Rice et al., 1998). If participants presented with a similar weakness, this assessment process would provide a fine-grained approach for identifying problematic tense markers to target in intervention.

A broad search was conducted to identify potential participants. Recruitment letters were distributed using three different strategies. First, recruitment letters were distributed to speech-language pathologists in the author’s professional network who work with pediatric AAC speakers. The speech-language pathologists were asked to distribute these letters to the parents (or guardians) of potential participants. Second, recruitment letters were sent to individuals who participated in the Pittsburgh AAC Language Seminar Series, a monthly training seminar for parents, speech-language pathologists, and other service providers working with pediatric AAC speakers hosted by Semantic Compaction Systems in Pittsburgh. These letters invited the parents of potential participants who may meet the inclusion criteria to contact the author if they were interested in participating.

Finally, an internal chart review was conducted to identify potential participants who attended an intensive summer camp for pediatric AAC speakers and their families and met the criteria. When potential participants were identified in the internal chart review, recruitment letters will be sent directly to their families by mail. This letter asked each potential participant's family to consult with their current speech-language pathologist to verify that the inclusion criteria were met and then contact the researcher to indicate that they were interested in participating.

The researcher responded directly to all parents who expressed interest in having their child participate. Parental consent to complete the initial screening process was obtained via email. Parents of potential participants were encouraged to contact the researcher by telephone or email if they had questions about the initial screening process.

Once consent was obtained, the researcher and parents scheduled a telephone meeting to discuss the project in more detail. The researcher summarized the study goals, summarized the structure of the assessment and intervention process, and asked the parents a series of initial screening questions. The initial screening questions asked the parents to verify (to the extent of their knowledge) that inclusion criteria a-g were met. The initial screening questions also asked the parent's age to verify that they were at least 18 years old. If the potential participant's hearing had not been screened or evaluated within the last 24 months, the parents were asked to have their child's hearing screened by a local audiologist.

At the time of the telephone interview, a packet containing a parent report form for the CDI Words and Sentences and a parent report form for the *Early LAMbaseline* were mailed to the parents. Parents were asked to identify items on the CDI and *Early LAMbaseline* that their child understands. The parents were asked to return both completed forms with a copy of their child's most current individualized family service plan (IFSP) or individualized education program (IEP) and the results of their child's most recent hearing evaluation or screening. These documents were reviewed to verify that all inclusion criteria were met. Once this was verified, the child was invited to participate in the main study.

3.5.3 Comprehensive pretest assessment.

A series of assessment tasks was administered as a comprehensive pretest assessment of each participant's tense and agreement system. Pretest assessment tasks were distributed across a two-day period as outlined in (26):

26. Two-day schedule for pretest assessment:

   a. Day 1: spontaneous language sample, TACL-3 (Carrow-Woolfolk, 1999), grammaticality judgment task, review language on SGD between Day 1 and Day 2 to assess SGD productivity.

   b. Day 2: elicitation probe, stimulability tasks, make changes if necessary to ensure SGD productivity.

3.5.3.1 Spontaneous language sample.

Spontaneous language samples were collected to obtain the measures of spontaneous tense and agreement morpheme use from Experiment 1. These same language samples were used to obtain measures of MSL (Klee & Fitzgerald, 1985) and predicted MLUm (Kovacs & Hill, 2017). Language samples were collected using a language activity monitor (LAM) tool (Hill, 2004; Hill & Romich, 1999, 2001) and a stationary video camera as participants communicated with their SGD in their daily environments. LAM tools automatically record a log of language produced using SGDs in a text-based logfile. A new language event is added to the logfile in a standard logfile format each time the user generates one or more text characters using the SGD. Formatted language events include three components, a timestamp in 24-hour clock format, a 3-letter mnemonic indicating the language representation method used to generate the language event, and the output of the language event. For example: 18:24:30 SEM "How" (Kovacs & Hill, 2015). In this example, an AAC speaker produced *how* at 18:24:30 using sequenced multi-meaning icons on his SGD.

The spontaneous language samples were collected during an extended observation period in the home environment (Participant A) or a speech-language pathologist's office (Participant B) where the participant was known to be highly interactive. The participant and a parent interacted with each other during this task. They were instructed to talk like they normally do at home. An adult babysitter also participated in Participant A's pretest language sample as a communication partner. As much language data was collected as possible within the intensive data collection schedule. The initial intent was to allot three hours for online data collection and collect as many utterances as possible for language sample analysis. AAC speakers typically generate fewer utterances than their verbally speaking peers in any given time window because they have severely reduced communication rates. All spontaneous novel utterances with at least 2 morphemes and at least 2 language events were selected for analysis as part of the spontaneous language sample.

Spontaneous language samples were transcribed using C-unit segmentation (Loban, 1976). C-units were segmented using PeRT software (Hill & Romich, 2003), a program for transcribing and analyzing LAM logfiles. A temporal utterance boundary based on gaps between timestamps also was used to define boundaries between utterances when clear, consistent evidence of utterance terminators could not be identified (Kovacs & Hill, 2015; Ortloff, 2010). The temporal utterance boundary was defined as a two-minute pause between timestamps, which was unlikely to occur within an utterance. Video data and notes taken online during language sample collection were reviewed as needed for information on conversational context.

3.5.3.2 TACL-3.

The TACL-3 was administered as a standardized test of receptive language skills.

3.5.3.3 Grammaticality judgment.

Participants' receptive knowledge of each tense marker was tested using a grammaticality judgment task based on Blockberger and Johnston's (2003) procedure for eliciting grammaticality judgments from AAC speakers, which was adapted from Stromswold's (as cited in Blockberger & Johnston, 1993) procedure for eliciting grammaticality judgments from young typically developing children. Two grammatical sentences and two ungrammatical sentences were presented for each tense marker. The ungrammatical sentences included one sentence where the tense marker was omitted (e.g., The children playing), and one sentence where the tense marker was used in conflict with other information in the sentence (e.g., The children is playing). Sentences were presented in a randomized order. The researcher introduced the participant to a dog puppet from outer space that was “just learning to talk.” The participants were asked to assist by listening to the dog say some sentences. A plate of doggie treats and a plate of rocks were set in front of the participant. The participant was given the initial instructions in (27).

27. The dog will practice saying some sentences one at a time. Listen to the sentence. If it sounds okay, feed him a treat. If the dog gets mixed up or makes a mistake, feed him a rock.

Four practice sentences that did not contain tense markers (two grammatical, two ungrammatical) were presented to familiarize the child participant to this task.

3.5.3.4 Assessing and ensuring SGD productivity.

The pre-stored vocabularies on participants’ SGDs was reviewed to assess the SGD’s potential to support productive tense and agreement morpheme use (Redmond & Johnston, 2001) and identify cases where failure to use a morpheme could relate to limitations in the SGD’s pre-stored vocabularies. Measurements of potential tense and agreement morpheme use based on Hadley and Short’s (2005) measures of spontaneous tense and agreement morpheme use were obtained from the sentences generated during this review. See Appendix A for details.

The researcher attempted to generate grammatical sentences with the pre-stored vocabularies on the participants’ SGDs to demonstrate selective use of tense and agreement morphemes. Spelling mode was not used during this process.

Operational criteria for demonstrating selective use of each tense marker were modeled after Binger and Light’s (2008) examples showing that AAC systems supporting selective use of -*ed* allow for production of inflected and uninflected forms of the same lexical verb. The criteria for *–ed* and *-3s* required production of one sentence that includes the inflected form of a lexical verb and one sentence that includes the uninflected stem form of the same lexical verb.

The criteria for copula and auxiliary *is*, *are*, *was*, *were*, and auxiliary *do*, *does*, *did* required production of one sentence using the tense marker with a nominal subject and a second sentence using the same nominal subject with no copula or auxiliary forms. Some SGDs allow these tense markers to be produced with pronominal subjects in a single multi-word language event (*he is*, *he does*, etc.) (e.g., Baker, 1982). Producing a sentence with a tense marker and a nominal subject requires generation of a subject-tense marker combination with at least two language events. Hadley and Short (2005) and Rispoli and Hadley (2011) exclude instances of tense markers contracted to pronominal subjects from their analysis because these high frequency contractions may be stored in memory as unanalyzed wholes. Requiring generation of sentences with nominal subjects served a similar purpose in this review process. An SGD could not be unduly credited with supporting selective use of a tense marker that could only be used with pronominal subjects as part of a multi-word language event.

Copula and auxiliary *am* only take one possible subject, *I*. The criteria for demonstrating selective use of copula and auxiliary *am* required production of two sentences with *I* as the subject: one with correct use of *am* and one that does not include *am*.

Potential tense marker use was reported for each tense marker as a binary variable indicating whether or not the tense marker could be selectively used on the participant's SGD. After the pretest assessment was completed, tense markers were be added as needed so that selective use of all 15 tense markers was supported.

The researcher selectively used morphemes from each morpheme category in sufficiently different contexts to demonstrate the potential productivity of each morpheme category on the participant's SGD. *Potential category productivity scores* were measures of the potential productivity of each morpheme category, with up to 5 sufficiently different uses counted towards each category. A composite *potential productivity score* was found by calculating will be the sum of the five potential category productivity scores.

Operational criteria for demonstrating sufficiently different uses of each morpheme category were consistent with the definitions used by Hadley and Short (2005). For *–ed* and *-3s*, sufficiently different uses were demonstrated by demonstrating selective use of the inflected and uninflected forms of different lexical verbs. For COPULA BE, AUXILIARY BE, and AUXILIARY DO, sufficiently different uses were demonstrated by generating pairs of sentences that met criteria for selective tense marker use with different subject-tense marker combinations. After all pretest assessment tasks were completed, the pre-stored vocabulary on the participant's SGD was modified if necessary to allow the SGD to support productive use of all 5 morpheme categories.

3.5.3.5 Elicitation probes.

Single-item elicitation probes were administered as an attempt to elicit a production of each tense marker that participants did not produce at least once in their pretest language sample. The elicitation probe was used to identify tense markers that the child participant has acquired but did not use spontaneously. Elicitation probes were administered using similar prompts to corresponding items in one of the stimulability tasks. However, the participants used their personal SGD to respond and no models of correct production or corrective feedback were provided. Elicitation probes for copula and auxiliary *is*, *am*, and *are* used visual stimuli with human characters so children could respond using high-frequency pronominal subjects.

3.5.3.6 Tense marker stimulability test.

The stimulability tasks from Experiment 2 were administered in the graphic symbol modality to assess tense and agreement morpheme stimulability. These tasks were used to obtain the measures of tense and agreement morpheme stimulability from Experiment 2.

Both participants used direct selection with their fingers to complete the stimulability tasks because this was the access method they routinely used to access their SGDs. Participant Participant A made selections on the iPAD mini 4 used in Experiment 2. Participant B used his own SGD after the custom vocabulary pages for the stimulability task were uploaded onto his device. These custom vocabulary pages were only used in the stimulability testing tasks defined in Experiment 2. Participants used their own SGDs for all other tasks.

3.5.3.7 Dependent variables obtained during pretest assessment.

The combined results from the spontaneous language sample and the elicitation probe were used to measure tense marker use, tense marker total, productivity score, and the five different category productivity scores. All of these measures were found from the pretest language samples using the operational criteria defined in Experiment 1. If a participant produced a tense marker during the elicitation probe, this was considered evidence that the participant produced the tense marker independently and counted towards the participant's tense marker total. This was counted towards the category productivity score and productivity score to avoid categories with tense marker use and zero productivity.

MSL was calculated as a measure of utterance length. MSL also was used to predict MLUm following Kovacs and Hill's (2017) procedure for measuring utterance length in language samples collected using LAM.

Comprehension of each tense marker was rated on a scale of 0-4 using the total number of correct responses on corresponding items of the grammaticality judgment task.

Tense marker stimulability was measured for all 15 tense markers using the procedures defined in Experiment 2. Category stimulability scores were obtained for all five tense marker categories using the procedures defined in Experiment 2.

3.5.4 Intervention

One target tense marker was selected as an intervention target for each participant. The combined results from the language sample, elicitation probe, grammaticality judgment, and stimulability tasks were used to guide target tense marker selection. The target tense markers met the selection criteria in (28).

28. Target tense marker selection criteria:

a. The participant demonstrated pretest comprehension of the target tense marker by responding correctly on at least 3/4 corresponding grammaticality judgment items.
b. The participant did not use the target tense marker in the pretest language sample or the pretest elicitation probe.
c. The participant demonstrated stimulability for the target tense marker by producing the target tense marker at least once in the stimulability task.

A five-day course of intensive intervention focused on increasing target tense marker production was provided. The planned course of intervention consisted of ten pairs of treatment and probe sessions, with two pairs of treatment and probe sessions administered per day.

3.5.4.1 Treatment sessions.

Each treatment session consisted of a shared reading activity focused on increasing production of the target tense marker. These activities were designed to adhere to Fey et al.'s (2003) principles for principles for grammatical intervention procedures and activities (29).

29. Principles for grammatical intervention activities.

   a. Create frequent opportunities to use targets.
   b. Use stimulus materials that provide appropriate contexts for using specific targets.
   c. Manipulate discourse so that targets are rendered more salient.
   d. Contrast the child's forms with more mature adult forms using recasts.
   e. Always present grammatical models in well-formed phrases and sentences.
   f. Use elicited imitation to make target forms more salient.

Storybooks that provided robust opportunities for using the target tense marker were developed following the general guidelines of Tönsing (Tönsing, 2012; Tönsing et al., 2014). Tönsing used an operationalized set of design principles (Tönsing, 2012) to create custom storybooks that provided scripted opportunities for production of targeted semantic relations in the graphic symbol modality during shared reading activities. These same principles, listed in (30), were used to guide the development of storybooks that provide opportunities for target tense marker use during shared reading activities.

30. Design principles for storybook activities, adapted from Tönsing (2012).

   a. Use simple sentences.
   b. Use developmentally appropriate vocabulary words that are consistent with the participants' receptive language age. Target vocabulary used during elicitations should be included in the pre-stored vocabulary of the participant's AAC system.

c. Use a story grammar with one or more simple episodes consisting of an initiating event, an overt attempt by the main character, and a direct consequence (Peterson & McCabe, 1983).
d. The story should include 10 opportunities for producing the target tense marker in a variety of different semantic contexts (one opportunity per illustrated page).
e. The story should include 10 examples of a contrasting form in order to increase the conceptual saliency of the target tense marker (Fey et al., 2003).

During each treatment session, the author and participant read the storybook together. The author and participant read the story, looked at illustrations, and discussed the story with comments and discourse as necessary. Participants always had access to their personal SGD during these activities. All participant contributions were acknowledged and responded to. To the extent possible, the author provided aided language stimulation by responding to unscripted participant comments with contextual responses that used the target tense marker. Some of these responses were contrastive recasts that presented a more mature form of the participants' own utterances in the graphic symbol modality.

Following Tönsing (Tönsing, 2012; Tönsing et al., 2014), a least-to-most prompting hierarchy was used when possible to elicit a production of the target tense marker during each scripted opportunity. During the assessment, the stimulability task provided evidence that participants could produce target tense markers with maximal assistance following an adult model. The least-to-most prompt hierarchy gave participants opportunities to produce target tense markers with greater independence and lower levels of support. A maximal prompt (full verbal prompt with aided language stimulation) was given as a contrastive recast if a participant generated a partial/incomplete response that did not include the target tense marker. Positive verbal reinforcement was provided any time a participant correctly used a target tense marker. When possible, the author explicitly reinforced target tense marker use by praising the participant and repeating the utterance containing the target tense marker.

Parents were encouraged to provide opportunities for target tense marker use at home between sessions and taught about grammatical contexts for target tense marker use. Parents also were taught how to produce target tense markers on the participants' SGDs and were encouraged to model target tense marker use with aided language stimulation when possible. Each family was given copies of storybooks highlighting target tense markers to use at home as scripted practice activities.

3.5.4.2 Probe sessions.

A 10-item probe task was developed for probing independent production of target tense markers. Probe items included picture stimuli with a variety of human (male and female) and animal characters and flexible verbal cues that allowed participants to respond using either a pronominal or nominal subject. Participants were credited with a correct response on a probe item if they used the correct tense marker in a grammatically correct context. Other responses were be scored as incorrect.

The probe task was administered after each treatment session. Participants were expected to respond using their personal SGDs. An expectant delay of 10 seconds was given after the stimuli are presented. The delay was extended if the participant was still clearly formulating a response. No further prompts were provided.

3.5.5 Posttest Generalization

A posttest assessment was conducted after the intervention phase to assess generalization of treatment effects to non-target tense markers. Generalization tasks replicated tasks from the pretest assessment and measure the same dependent variables. The original intent was to collect all posttest data in a single day. However, posttest tasks were distributed over multiple days because participants were fatigued, and it was not possible to analyze language samples online in preparation for the elicitation probe task.

A spontaneous language sample was collected using the same procedures that were used to collect the pretest language sample. Single-item elicitation probes were administered following the procedures used during pretest assessment as an attempt to elicit a production of each tense marker that was not produced in the posttest language sample. The combined results from the spontaneous language sample and the elicitation probe were used to measure tense marker use, tense marker total, productivity score, and the five different category productivity scores. MSL and predicted MLUm were computed from the language sample. A descriptive comparison of pretest and posttest measures was used to measure changes in tense marker use over the course of intervention.

4.0 RESULTS

4.1 EXPERIMENT 1 AND 2 PARTICIPANTS

The child participants in Experiments 1 and 2 included a total of 40 children (18 boys and 22 girls) who met all inclusion criteria and completed both experiments. Experiments 1 and 2 were completed in two separate home visits on different days. Experiment 2 was completed 1-14 days after Experiment 1 (M = 5.35, SD = 4.35).

Parents completed a brief questionnaire with screening questions before Visit 1 was conducted with each child. By parent report, all child participants were monolingual English speakers. Parents reported that 6 children spoke words, but not phrases in a second language. These languages included Spanish (2), Arabic (1), Lebanese Arabic (1), Hebrew (1), and Portuguese (1). Parents reported that their child had no history of speech, language, hearing, or cognitive impairments and no history of neurological injury. Parents also reported that their child had normal or corrected to normal vision and the ability to move an index finger in isolation.

Three additional screening tasks were completed during Visit 1: an oral-peripheral examination, a pure-tone hearing screening, and a standardized developmental language test (the CDI-III or SPELT-P, depending on the child's age). All children performed within normal limits on all three of these tasks.

Twenty child participants were toddlers. At Visit 1, toddlers ages ranged from 30 to 42 months (*M* = 35.95, *SD* = 4.03). At Visit 2, toddlers ages ranged from 30 to 42 months (*M* = 36.00, *SD* = 4.10). The toddlers' MSL ranged from 3.27 to 6.67 morphemes (*M* = 4.58, *SD* = 0.81). Their predicted MLUm ranged from 2.38 to 5.32 morphemes (*M* = 3.52, *SD* = 0.71).

Twenty child participants were preschoolers. At Visit 1, preschoolers ages ranged from 43 to 54 months (*M* = 48.00, *SD* = 3.51). At Visit 2, preschoolers ages ranged from 43 to 54 months (*M* = 48.20, *SD* = 3.50). The preschoolers MSL ranged from 4.01 to 6.57 morphemes (*M* = 5.30, *SD* = 0.69). Their predicted ranged from 4.01 to 6.57 morphemes (*M* = 4.14, *SD* = 0.60).

Parents of 11 toddlers age 30-36 months completed the CDI-III. One child scored in the 8th percentile for 36-37 month old boys and girls on the Vocabulary Checklist section of the CDI-III. All other children scored above the 10th percentile for boys and girls their age on the Vocabulary Checklist section. All children scored at or above the 10th percentile on the Sentences and Using Language sections of the CDI-III. CDI-III scores are summarized in Table 4.1.

Table 4.1. Summary of *CDI-III* Test Scores

Subtest	N	Raw Score		Percentile[a]	
		M	*SD*	*M*	*SD*
Vocabulary	11	61.09	16.28	37.45	21.71
Sentences	11	9.64	2.66	61.00	31.46
Using Language	11	9.27	1.27	65.00	16.58
M3L[b]	10	9.75	2.36		

Note. *M* = mean; *SD* = Standard deviation; M3L = mean length of three longest sentences.

[a]Percentile rank relative to pooled norms for age-matched boys and girls in the normative sample (Fenson et al., 2007). [b]This section left blank on one parent response form.

Nine toddlers ages 37-42 months and all preschoolers completed the SPELT-P2 using Standard American English (28) or African American Vernacular English (1). SPELT-P2 scores are summarized in Table 4.2. All children achieved standard scores of 87 or higher on the SPELT-P2. It should be noted that the mean SPELT-P2 score for preschoolers fell within the upper quartile of the normative sample (Dawson et al., 2005). The children in this study may have had relatively advanced expressive morphological and syntactic skills for their age. If this is the case, then the children may have performed on a level that would be more consistent with the average performance of slightly older children on some tasks. Alternatively, they may have achieved higher scores than expected because they were tested in the comfort of their home environments under relatively optimal conditions.

Table 4.2. Summary of *SPELT-Ps* Test Scores

	Toddlers[a]		Preschoolers[b]	
	M	*SD*	*M*	*SD*
Raw Score	20.33	5.12	30.00	5.34
Standard Score[c]	101.78	9.31	114.20	11.88
Percentile[c]	55.22	20.87	79.65	24.87

Note. *M* = mean; *SD* = Standard deviation.

[a]N = 9. [b]N = 20. [c]Standard score and percentile rank relative to age-matched peers in the normative sample (Dawson et al., 2005).

Completed Early-LAMBaseline forms were returned for 32 children during Visit 2. Twenty parents completed the Word Inventory section by checking off both words that their child understands and words that their child uses. These parents indicated that their child understands at least 95 of the high frequency words on the word inventory and use at least 75 of the high frequency words on the word inventory. The other 12 parents checked off words on one list only or checked of different words on each list as if the lists are mutually exclusive. All parent respondents indicated that their child understands more than 75 nouns. One parent respondent indicated that their child uses 50-75 nouns. All others indicated that their child uses more than 75 nouns.

A total of 39 parents participated as communication partners in the language sample task. Parent participants included 35 mothers and 4 fathers. Parent ages ranged from 26 to 50 years (M = 35.64, SD = 5.43). All parents spoke fluent English.

One pair of siblings participated. Their father participated as the adult communication partner with both children. All other children were the only child in their family to participate. All other parents participated as communication partners with one child.

Five additional children and their parents were excluded after consent was obtained. One child was excluded after Visit 1 because inclusion criteria were not met. Three children were excluded after Visit 1 because they did not consent to completing all screening tasks. One child was excluded after Visit 2 because the stimulability tasks were invalidated.

4.2 EXPERIMENT 1 RESULTS

4.2.1 Question 1A.

On average, the children used 6.43 out of 15 English tense markers at least once in their language samples (SD = 2.15, range = 3-11). The main effect of age was a significant predictor of tense marker use, b = 0.03, SE = 0.01, p = 0.01. Increasing age by one month increased the odds of a child using any given tense marker by a factor of 1.031. Predicted proportions of tense markers used at least once are plotted as a function of age in Figure 4.1.

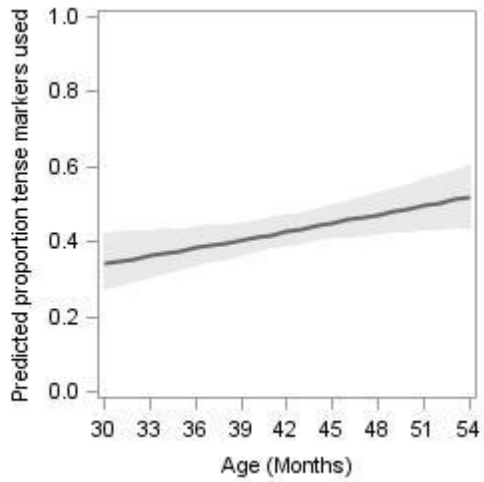


Figure 4.1. Predicted proportions of tense markers used out of 15 English tense markers.

4.2.2 Question 1B.

Children's average productivity scores were 12.38 (*SD* = 4.87, range = 3-24). The main effect of age was a significant predictor of productivity, $b = 0.05$, $SE = 0.02$, $p = 0.01$. Increasing age by one month increased the odds of a child receiving any given productivity point by a factor of 1.051. Predicted productivity scores are plotted in Figure 4.2 as proportions of the maximum possible score of 25 productivity points.

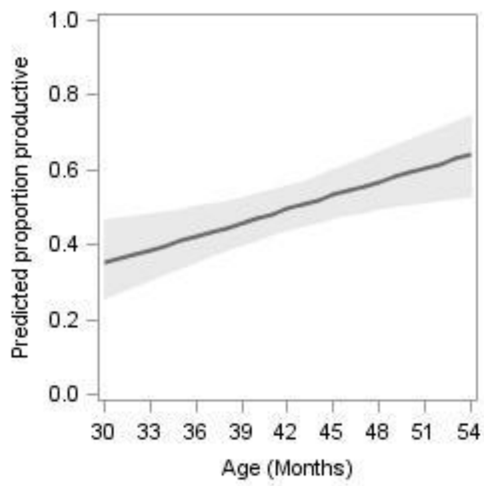


Figure 4.2. Predicted productivity scores as proportions of the maximum possible score (of 25).

4.2.3 Question 1C.

Predicted probabilities of morpheme category productivity as a function of age and morpheme category are shown in Figure 4.3.

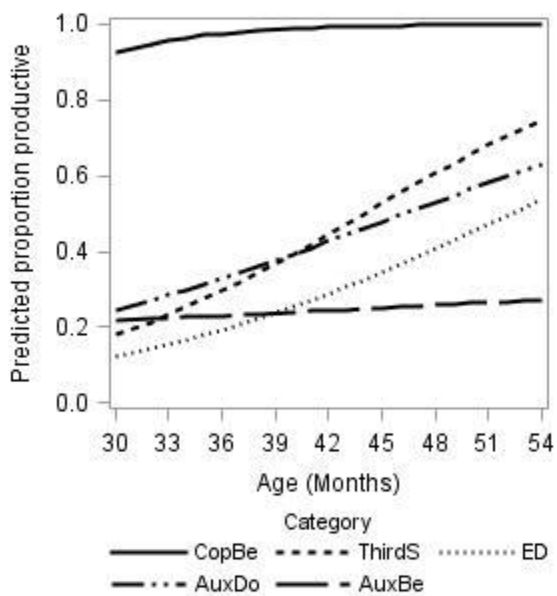


Figure 4.3. Proportional predicted morpheme category productivity scores.

Morpheme category productivity was modulated by a significant age-by-morpheme category interaction, indicating that productivity grew at different rates across morpheme categories. Both *-ed* productivity, $b = 0.08$, $SE = 0.04$, $p = 0.04$ and *-3s* productivity, $b = 0.10$, $SE = 0.04$, $p = 0.01$ increased at significantly faster rates than AUXILIARY BE productivity. No other significant interactions were found, *p*s > 0.05.

The main effect of age was a significant predictor of morpheme category productivity for COPULA BE, $b = 0.20$, $SE = 0.10$, $p = 0.05$, *-3s*, $b = 0.11$, $SE = 0.03$, $p < 0.01$, *-ed*, $b = 0.09$, $SE = 0.03$, $p = 0.01$, and AUXILIARY DO, $b = 0.07$, $SE = 0.03$, $p = 0.04$. The odds of a child receiving any given COPULA BE productivity point increased by a factor of 1.222 per month. The odds of a child receiving any given *-3s* productivity point increased by a factor of 1.114 per month. The odds of a child receiving any given *-ed* productivity point increased by a factor of 1.092 per month. The odds of a child receiving any given AUXILIARY DO productivity point increased by a factor of 1.072 per month. The main effect of age was not a significant predictor of morpheme category productivity for AUXILIARY BE, $b = 0.01$, $SE = 0.03$, $p = 0.72$.

The main effect of morpheme category was a significant predictor of morpheme category productivity at 30 months, 42 months, and 54 months. A ceiling effect was found for COPULA BE productivity. All children received COPULA BE productivity scores of 3 or more, and 37 children (92.5%) were at ceiling for COPULA BE productivity scores. At 30 months, 42 months, and 54 months, COPULA BE productivity was significantly higher than *-3s* productivity, *p*s < 0.01, *-ed* productivity, *p*s < 0.01, AUXILIARY DO productivity, *p*s < 0.01 and AUXILIARY BE productivity, *p*s < 0.01. At 30 months, the odds of a child receiving a given COPULA BE productivity point were at least 37.72 times as high as the odds of a child receiving a given productivity point in any other morpheme category. At 42 months, the odds of a child receiving a given COPULA BE productivity point were at least 168.2 times as high as the odds of a child receiving a given productivity point in any other morpheme category. At 54 months, the odds of a child receiving a given COPULA BE productivity point were at least 511.8 times as high as the odds of a child receiving a given productivity point in any other morpheme category. This ceiling effect was consistent with GML theory's prediction that COPULA BE would grow in productivity faster than all other categories, and indicates that children reach high levels of COPULA BE productivity before the age of 30 months.

Patterns of productivity in the four remaining categories varied as a function of age. At 30 months, there were no significant differences between AUXILIARY DO productivity and AUXILIARY BE productivity, $b = 0.16$, $SE = 0.48$, $p = 0.75$, *-3s* productivity, $b = 0.38$, $SE = 0.48$, $p = 0.42$ or *–ed* productivity, $b = 0.84$, $SE = 0.51$, $p = 0.10$. At 30 months, there also were no significant differences between AUXILIARY BE productivity and *-3s* productivity, $b = 0.23$, $SE = 0.50$, $p = 0.65$ or *–ed* productivity, $b = 0.69$, $SE = 0.52$, $p = 0.19$. Finally, no significant differences were found between *-3s* productivity and *-ed* productivity at 30 months, $b = 0.46$, $SE = 0.52$, $p = 0.38$.

At 42 months, there were no significant differences between *-3s* productivity and AUXILIARY DO productivity, $b = 0.08$, $SE = 0.23$, $p = 0.73$ and no significant differences between *-ed* productivity and AUXILIARY BE productivity, $b = 0.22$, $SE = 0.24$, $p = 0.36$ at 42 months. However, *-3s* productivity was significantly higher than both *-ed* productivity, $b = 0.69$, $SE = 0.24$, $p < 0.01$ and AUXILIARY BE productivity, $b = 0.92$, $SE = 0.24$, $p < 0.01$ at 42 months. AUXILIARY DO productivity also was significantly higher than –ED productivity, $b = 0.61$, $SE = 0.23$, $p = 0.01$ and AUXILIARY BE productivity, $b = 0.84$, $SE = 0.23$, $p < 0.01$ at 42 months. The odds of a child receiving a given *-3s* productivity point were 2.000 times as high as the odds of a child receiving a given *-ed* productivity point and 2.498 times as high as the odds of a child receiving a given AUXILIARY BE productivity point at 42 months. The odds of a child receiving a given AUXILIARY DO productivity point were 1.848 times as high as the odds of a child receiving a given *-ed* productivity point and 2.308 times as high as the odds of a child receiving a given AUXILIARY BE productivity point at 42 months.

At 54 months, AUXILIARY BE productivity was significantly lower than *-3s* productivity, $b = 2.06$, $SE = 0.48$, $p < 0.01$, AUXILIARY DO productivity, $b = 1.52$, $SE = 0.47$, $p < 0.01$, and *–ed* productivity, $b = 1.13$, $SE = 0.48$, $p = 0.02$. The odds of a child receiving a given *-3s* productivity point were 7.831 times as high as the odds of a child receiving a given AUXILIARY BE productivity point at 54 months. The odds of a child receiving a given AUXILIARY DO productivity point were 4.561 times as high as the odds of a child receiving a given AUXILIARY BE productivity point at 54 months. The odds of a child receiving a given *–ed* productivity point were 3.097 times as high as the odds of a child receiving a given AUXILIARY BE productivity point at 54 months.

At 54 months, there were no significant differences between *-3s* productivity and AUXILIARY DO productivity, $b = 0.54$, $SE = 0.45$, $p = 0.23$ and no significant differences between AUXILIARY DO productivity and *-ed* productivity, $b = 0.39$, $SE = 0.45$, $p = 0.39$. However, *-3s* productivity was significantly higher than *-ed* productivity at 54 months, $b = 0.93$, $SE = 0.46$, $p < 0.04$. The odds of a child receiving a given *-3s* productivity point were 2.529 times as high as the odds of a child receiving a given *-ed* productivity point at 54 months.

4.3 EXPERIMENT 2 RESULTS

4.3.1 Question 2A.

Both types of 1-parameter subset model (age-only, modality-only), the 2-parameter subset model (age and modality), and the full model were selected as the best fitting model for at least one tense marker. Full models were selected for auxiliary *is*, *am*, *are* and *were* because they had the lowest *AICc* values in their respective sets of models. The full models for auxiliary *is* and *were* failed to converge. The age-by-modality interaction effect was not significant in the full model for auxiliary *am*, b = -1.25, SE = 1.00, p = 0.21 or the full model for auxiliary *are*, b = -0.59, SE = 0.34, p = 0.08. The lack of significant interaction effects indicates that tense marker stimulability grows at the same rate across communication modalities for all tense markers. Full models were removed and the model selection process was repeated using the two-parameter subset model with main effects of age and communication modality, the age-only and modality-only subset models, and the intercept-only model as candidate models.

At least one main effect was a significant predictor of tense marker stimulability in the selected models for 13 tense markers. The main effect of age was the only significant predictor in the selected models for 7 tense markers (copula *is*, *am* and *was*, auxiliary *did* and auxiliary *is*, *am* and *are*). The main effect of communication modality was the only significant predictor in the selected models for 3 tense markers (auxiliary *does* and *do* and auxiliary *were*). The main effects of both age and communication modality were significant predictors in the selected models for 3 tense markers (copula *are*, *-ed* and auxiliary *was*). No significant predictors were found in the selected models for the 2 remaining tense markers (copula *were*, and *-3s*).

4.3.1.2 Age-only subset models.

An age-only subset model was the best fitting model for predicting copula *am* and *was*, auxiliary *did*, and auxiliary *is* stimulability. Age-only subset models were significantly better than intercept models for copula *am* stimulability, likelihood ratio = 8.34, $p < 0.01$, copula *was* stimulability, Likelihood ratio = 4.40, $p = 0.04$, auxiliary *did* stimulability, likelihood ratio = 5.84, $p = 0.02$, auxiliary *is* stimulability, likelihood ratio = 12.91, $p < 0.01$, auxiliary *am* stimulability, likelihood ratio = 8.92, $p < 0.01$, and auxiliary *are* stimulability, likelihood ratio = 8.62, $p < 0.01$. Predicted probabilities of stimulability for these tense markers are plotted in Figure 4.4, with 95% confidence bands.

The main effect of age was a significant predictor of copula *am* stimulability, $b = 0.14$, $SE = 0.06$, $p = 0.01$, copula *was* stimulability, $b = 0.11$, $SE = 0.06$, $p = 0.05$, auxiliary *did* stimulability, $b = 0.12$, $SE = 0.05$, $p = 0.03$, auxiliary *is* stimulability, $b = 0.20$, $SE = 0.07$, $p < 0.01$, auxiliary *am* stimulability, $b = 0.15$, $SE = 0.06$, $p < 0.01$, and auxiliary *are* stimulability, $b = 0.15$, $SE = 0.06$, $p = 0.01$. Age coefficients were positive in these models, indicating that the odds of a child being stimulable for these tense markers increased with age regardless of the communication modality the child was tested in. Increasing age by one month increased the odds of a child being stimulable for copula *am* by a factor of 1.154. Increasing age by one month increased the odds of a child being stimulable for copula *was* by a factor of 1.116. Increasing age by one month increased the odds of a child being stimulable for auxiliary *did* by a factor of 1.128. Increasing age by one month increased the odds of a child being stimulable for auxiliary *is* by a factor of 1.220. Increasing age by one month increased the odds of a child being stimulable for auxiliary *am* by a factor of 1.162. Increasing age by one month increased the odds of a child being stimulable for auxiliary *are* by a factor of 1.159.

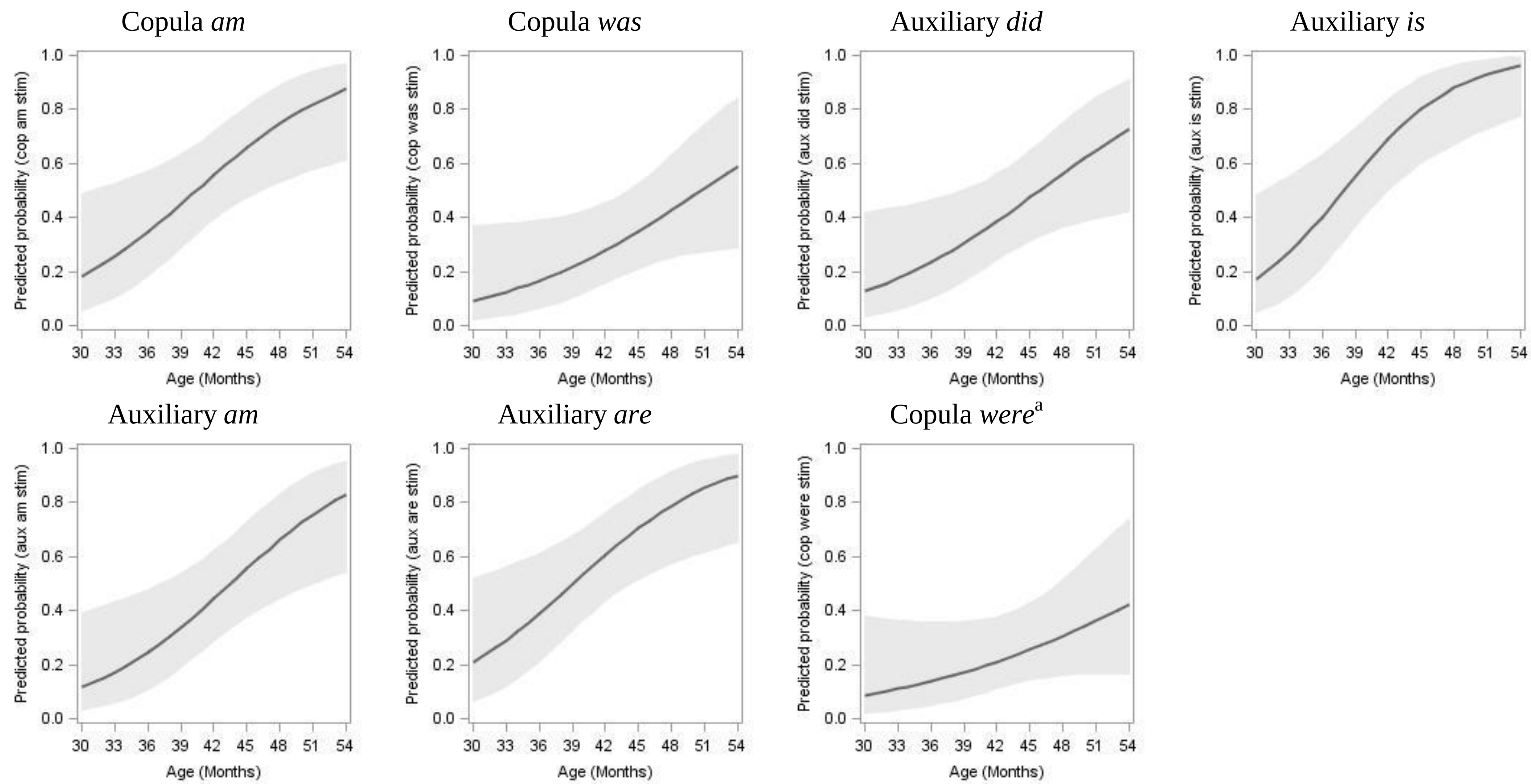


Figure 4.4. Predicted probabilities of tense marker stimulability in age-only subset models.

Note. [a]Copula *were* model is not significantly better than intercept-only model.

An age-only model also was the best fitting model for predicting copula *were* stimulability. However, the copula *were* model was not significantly better than an intercept model, likelihood ratio = 2.27, p = 0.13. The significance of the age parameter for this model was not considered because the model itself was not significant. Predicted probabilities of copula *were* stimulability are plotted in Figure 4.4, with a 95% confidence band.

4.3.1.3 Modality-only subset model.

The modality-only model was only selected as the best fitting model for predicting stimulability of an individual tense marker one time. A modality-only subset model was the best fitting model for predicting auxiliary *does* stimulability. The modality-only subset model of auxiliary *does* stimulability was significantly better than an intercept model, likelihood ratio = 5.35, p = 0.02. The main effect of communication modality was a significant predictor of auxiliary *does* stimulability, b = 1.67, SE = 0.77, p = 0.03. Testing a child in the spoken modality instead of the graphic symbol modality increased the odds of that child being stimulable for auxiliary *does* by a factor of 5.333 regardless of age. Predicted probabilities of auxiliary *does* stimulability are plotted in Figure 4.5, with a 95% confidence band.

Auxiliary *does*

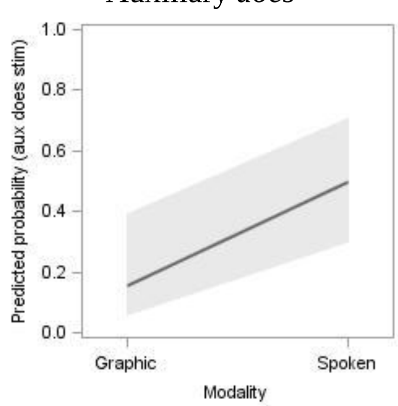


Figure 4.5. Predicted probability of aux *does* stimulability in modality-only subset model.

4.3.1.4 Age and modality subset models.

An age and modality subset model with two predictors was the best model for predicting tense marker stimulability for copula *is* and *are*, *-3s*, *-ed*, auxiliary *do*, and auxiliary *was* and *were*. Age coefficients were positive for all seven of these tense markers. In models where the main effect of age was a significant predictor of tense marker stimulability, the odds of a child being stimulable for that tense marker increased with age when communication modality was held constant. Modality coefficients also were positive for all seven of these tense markers. In models where the main effect of communication modality was a significant predictor of tense marker stimulability, the odds of a child being stimulable for that tense marker were higher in the spoken modality than the graphic symbol modality when age was held constant. Predicted probabilities of stimulability for all six of these tense markers are plotted in Figure 4.6, with 95% confidence bands.

For copula *is* stimulability, the age and modality subset model was significantly better than an intercept model, likelihood ratio = 10.92, $p < 0.01$. The main effect of age was a significant predictor of copula *is* stimulability, $b = 0.16$, $SE = 0.06$, $p = 0.01$. Increasing age by one month increased the odds of a child being stimulable for copula *is* by a factor of 1.175 when communication modality was held constant. The main effect of communication modality was not a significant predictor of copula *is* stimulability even though it was included in the model, $b = 1.59$, $SE = 0.86$, $p = 0.06$.

For auxiliary *do* stimulability, the age and modality subset model was significantly better than an intercept model, likelihood ratio = 7.85, $p = 0.02$. The main effect of communication modality was a significant predictor of auxiliary *do* stimulability, $b = 1.77$, $SE = 0.82$, $p = 0.03$. When age was held constant, testing a child in the spoken modality instead of the graphic symbol

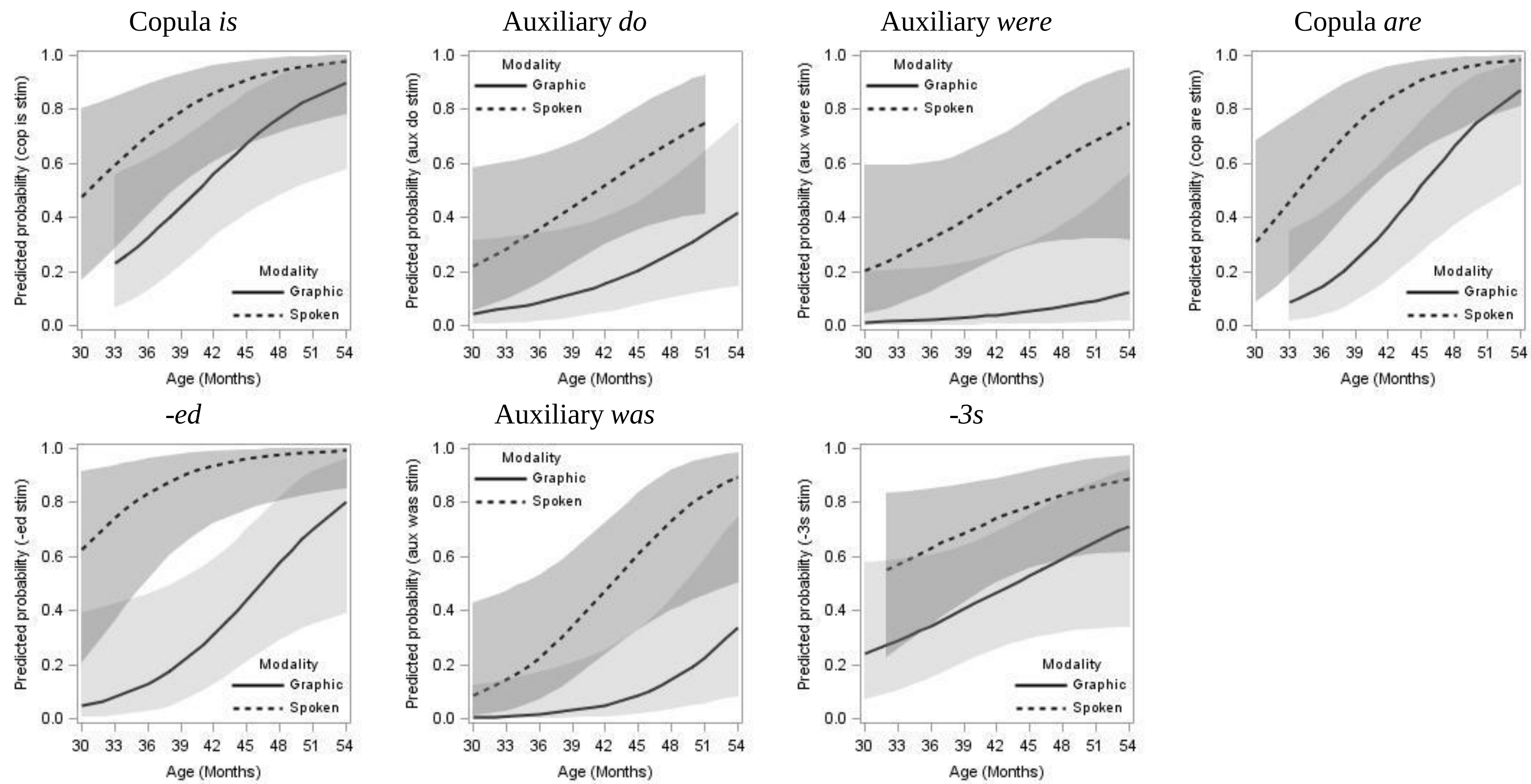


Figure 4.6. Predicted probability of tense marker stimulability in models with age and modality parameters.

modality increased the odds of that child being stimulable for auxiliary *do* by a factor of 5.875. Although age was included in the model, the main effect of age was not a significant predictor of auxiliary *do* stimulability, $b = 0.11$, $SE = 0.06$, $p = 0.06$.

For auxiliary *were* stimulability, the age and modality subset model was significantly better than an intercept model, likelihood ratio = 10.78, $p < 0.01$. The main effect of communication modality was a significant predictor of auxiliary *were* stimulability, $b = 3.05$, $SE = 1.21$, $p = 0.01$. When age was held constant, testing a child in the spoken modality instead of the graphic symbol modality increased the odds of that child being stimulable for auxiliary *were* by a factor of 21.148. Although age was included in the model, the main effect of age was not a significant predictor of auxiliary *were* stimulability, $b = 0.10$, $SE = 0.07$, $p = 0.12$.

For copula *are* stimulability, the age and modality subset model was significantly better than an intercept model, likelihood ratio = 17.41, $p < 0.01$. The main effects of age, $b = 0.20$, $SE = 0.07$, $p < 0.01$ and communication modality, $b = 2.21$, $SE = 0.94$, $p = 0.02$ were both significant predictors of copula *are* stimulability. Increasing age by one month increased the odds of a child being stimulable for copula *are* by a factor of 1.227 when communication modality was held constant. When age was held constant, testing a child in the spoken modality instead of the graphic symbol modality increased the odds of that child being stimulable for copula *are* by a factor of 9.131.

For -*ed* stimulability, the age and modality subset model was significantly better than an intercept model, likelihood ratio = 20.36, $p < 0.01$. The main effects of age, $b = 0.18$, $SE = 0.08$, $p = 0.02$ and communication modality, $b = 3.51$, $SE = 1.16$, $p < 0.01$ were both significant predictors of -*ed* stimulability. Increasing age by one month increased the odds of a child being stimulable for -*ed* by a factor of 1.201 when communication modality was held constant. When age was held constant, testing a child in the spoken modality instead of the graphic symbol modality increased the odds of that child being stimulable for *-ed* by a factor of 33.506.

For auxiliary *was*, the age and modality subset model was significantly better than an intercept model, likelihood ratio = 13.23, $p < 0.01$. The main effects of age, $b = 0.19$, $SE = 0.08$, $p = 0.02$ and communication modality, $b = 2.80$, $SE = 1.09$, $p = 0.01$ were both significant predictors of auxiliary *was* stimulability. Increasing age by one month increased the odds of a child being stimulable for auxiliary *was* by a factor of 1.206 when communication modality was held constant. When age was held constant, testing a child in the spoken modality instead of the graphic symbol modality increased the odds of that child being stimulable for auxiliary *was* by a factor of 16.453.

For *-3s* stimulability, the age and modality subset model was significantly better than an intercept model, likelihood ratio = 6.85, $p = 0.03$. Although both of these predictors were included in the model, neither the main effect of age, $b = 0.09$. $SE = 0.05$, $p = 0.09$ nor the main effect of communication modality, $b = 1.18$, $SE = 0.71$, $p = 0.10$ were significant predictors of *-3s* stimulability.

4.3.2 Question 2B.

Figure 4.7 shows the predicted proportion of correct probe item responses on stimulability tests for each morpheme category as a function of the significant fixed effects. These plots include 95% confidence bands. Fixed effects for each model are discussed in the following paragraphs.

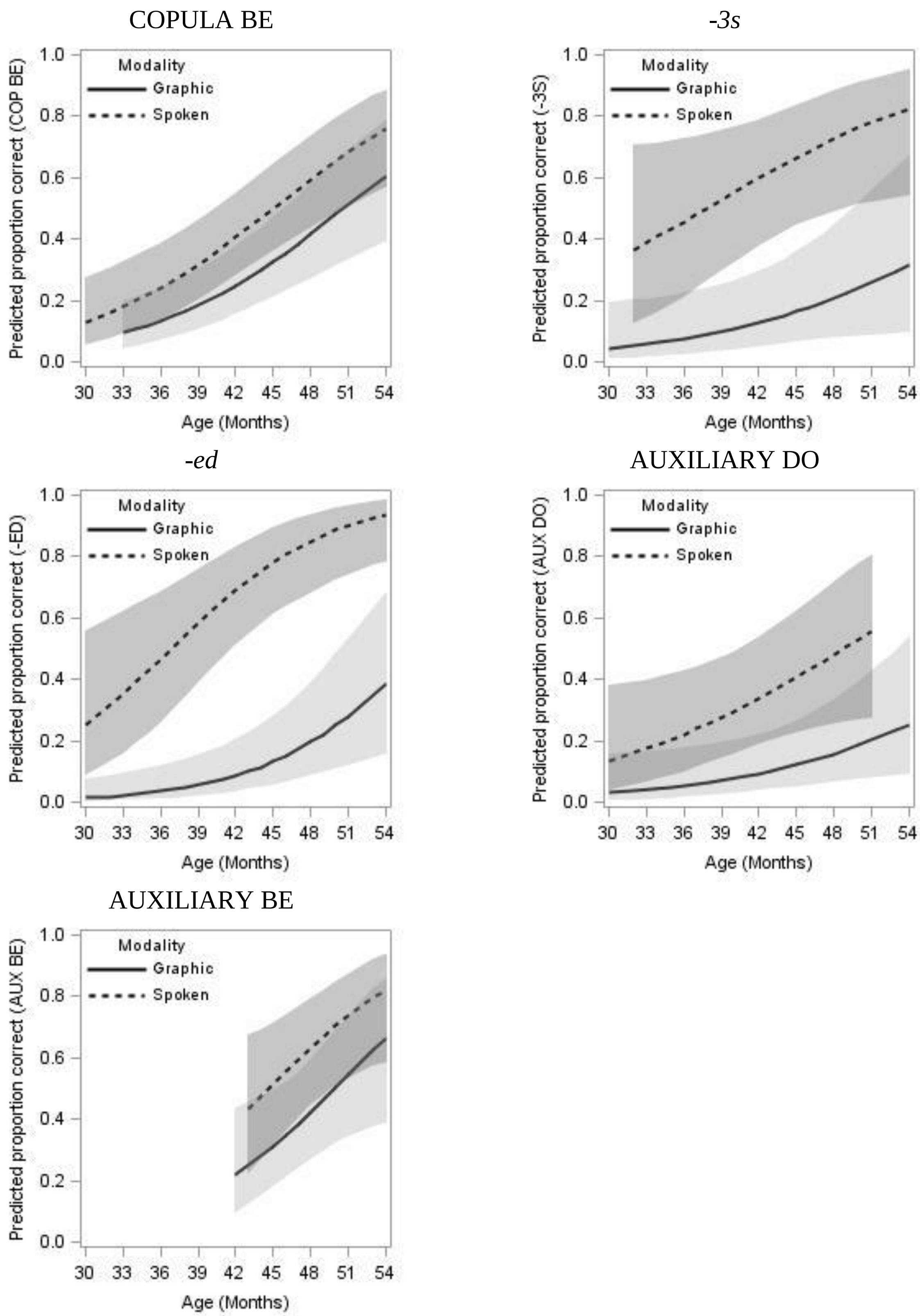


Figure 4.7. Predicted proportions of correct responses on category stimulability tests.

No significant age-by-modality interaction effects were found for COPULA BE category stimulability, $b = -0.04$, $SE = 0.07$, $p = 0.58$, *-3s* category stimulability, $b = -0.05$, $SE = 0.10$, $p = 0.63$, *-ed* category stimulability, $b = -0.03$, $SE = 0.11$, $p = 0.79$, or AUXILIARY DO category stimulability, $b = -0.12$, $SE = 0.10$, $p = 0.26$. This indicates that category stimulability grows at the same rate across communication modalities for each of these morpheme categories. The interactions were removed, and a second model was formed for each of these morpheme categories using fixed effects for the main effects of age and communication modality.

For AUXILIARY BE, morpheme category stimulability was modulated by a significant age-by-modality interaction, $b = -0.24$, $SE = 0.09$, $p = 0.01$. However, no children younger than 42 months on any AUXILIARY BE stimulability probe items. All children younger than 42 months were removed from the analysis. A second model was formed using data from the 23 children who were at least 42 months old (13 in the graphic symbol modality, 10 in the spoken modality). In this second model, the age-by-modality interaction was not significant, $b = -0.23$, $SE = 0.14$, $p = 0.08$, indicating that AUXILIARY BE category stimulability grows at the same rate across communication modalities after the age of 42 months. The interaction was removed, and a third model was formed using fixed effects for the main effects of age and communication modality.

The main effects of age were significant predictors of morpheme category stimulability in main effects models of COPULA BE category stimulability, $b = 0.13$, $SE = 0.03$, $p < 0.01$, *-ed* category stimulability, $b = 0.16$, $SE = 0.05$, $p < 0.01$, AUXILIARY DO category stimulability, $b = 0.10$, $SE = 0.05$, $p = 0.04$, and the main effects model of AUXILIARY BE category stimulability in children at least 42 months old, $b = 0.16$, $SE = 0.07$, $p = 0.02$. Age coefficients were positive in these models, indicating that the odds of a child giving a correct response on any corresponding COPULA BE, *-ed*, AUXILIARY DO, or AUXILIARY BE stimulability probe item increased with age when communication modality was held constant. Increasing age by one month increased the odds of a child giving a correct response on any given COPULA BE stimulability probe item by a factor of 1.137 when modality was held constant. Increasing age by one month increased the odds of a child giving a correct response on any given *-ed* stimulability probe item by a factor of 1.169 when modality was held constant. Increasing age by one month increased the odds of a child giving a correct response on any given AUXILIARY DO stimulability probe item by a factor of 1.103 when modality was held constant. For children at least 42 months old, increasing age by one month increased the odds of a child giving a correct response on any given AUXILIARY BE stimulability probe item by a factor of 1.176 when modality was held constant. The main effect of age was not a significant predictor of morpheme category stimulability in the main effects model of *-3s* category stimulability, $b = 0.10$, $SE = 0.05$, $p = 0.06$.

The main effects of communication modality were significant predictors of morpheme category stimulability in main effects models of *-3s* category stimulability, $b = 2.31$, $SE = 0.69$, $p < 0.01$, *-ed* category stimulability, $b = 3.15$, $SE = 0.64$, $p < 0.01$, and AUXILIARY DO category stimulability, $b = 1.60$, $SE = 0.66$, $p = 0.02$. Testing a child in the spoken modality instead of the graphic symbol modality increased the odds of that child giving a correct response on any given *-3s* stimulability probe item by a factor of 10.107 when age was held constant. Testing a child in the spoken modality instead of the graphic symbol modality increased the odds of that child giving a correct response on any given *-ed* stimulability probe item by a factor of 23.278 when age was held constant. Testing a child in the spoken modality instead of the graphic symbol modality increased the odds of that child giving a correct response on any given AUXILIARY

DO stimulability probe item by a factor of 4.954 when age was held constant. The main effects of communication modality were not significant predictors of morpheme category stimulability in the main effects model of COPULA BE category stimulability, $b = 0.72$, $SE = 0.42$, $p = 0.08$ or the main effects model AUXILIARY BE category stimulability in children at least 42 months old, $b = 0.85$, $SE = 0.52$, $p = 0.11$.

4.3.3 Question 2C.

Predicted probabilities of morpheme category stimulability as a function of age and morpheme category in each modality are shown in Figure 4.8.

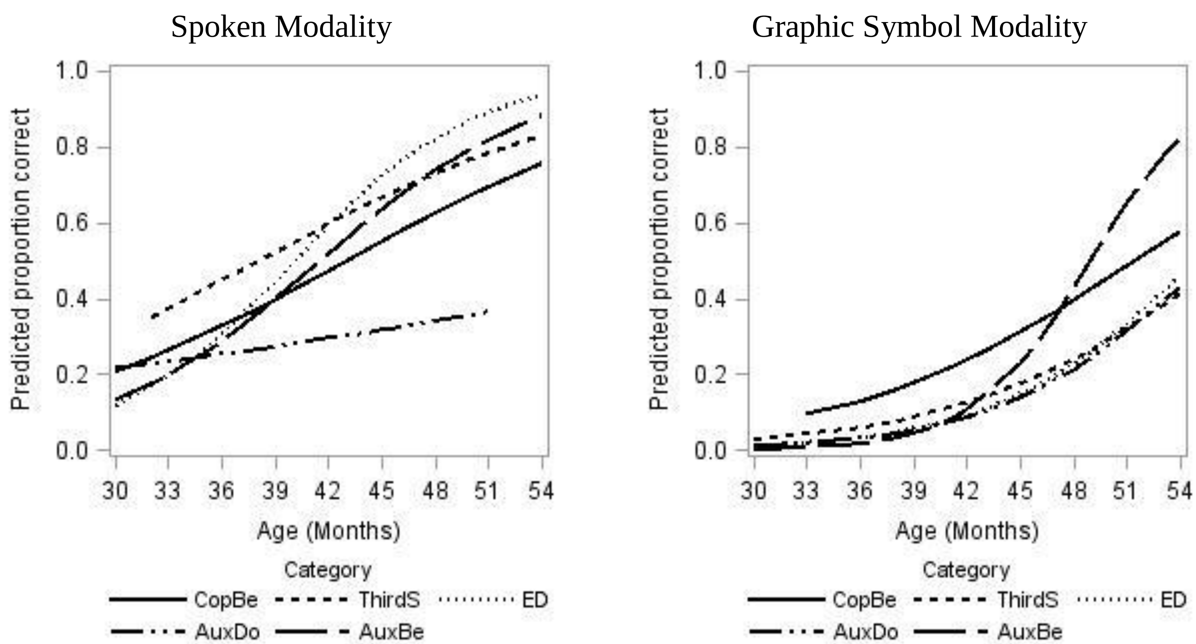


Figure 4.8. Predicted proportions of correct responses on stimulability across morpheme categories in the spoken and graphic symbol modalities.

4.3.3.1 Model for spoken modality.

Morpheme category stimulability in the spoken modality was modulated by a significant age-by-morpheme category interaction, indicating that morpheme category stimulability grew at different rates across categories. AUXILIARY BE category stimulability, $b = 0.13$, $SE = 0.05$, $p = 0.01$ and –ED category stimulability, $b = 0.16$, $SE = 0.06$, $p < 0.01$ both increased at significantly faster rates than AUXILIARY DO category stimulability. No other significant age-by-morpheme category interactions were found in the spoken modality , *p*s > 0.05.

The main effect of age was a significant predictor of –ED category stimulability, $b = 0.20$, $SE = 0.05$, $p < 0.01$, AUXILIARY BE category stimulability, $b = 0.16$, $SE = 0.04$, $p < 0.01$ and COPULA BE category stimulability, $b = 0.10$, $SE = 0.04$, $p = 0.02$ in the spoken modality. The odds of a child giving a correct response on any given *-ed* stimulability probe item increased by a factor of 1.218 per month in the spoken modality. The odds of a child giving a correct response on any given AUXILIARY BE stimulability probe item increased by a factor of 1.176 per month in the spoken modality. The odds of a child giving a correct response on any given COPULA BE stimulability probe item increased by a factor of 1.109 per month in the spoken modality. The main effect of age was not a significant predictor of *-3s* category stimulability, $b = 0.10$, $SE = 0.06$, $p = 0.08$ or AUXILIARY DO category stimulability, $b = 0.03$, $SE = 0.05$, $p = 0.50$ in the spoken modality.

The main effect of morpheme category was not a significant predictor of morpheme category stimulability in the spoken modality at 30 months. No significant differences were found between category stimulability scores in the spoken modality at 30 months, *p*s > 0.05.

The main effect of morpheme category was a significant predictor of morpheme category stimulability in the spoken modality at 42 months and at 54 months. At both 42 months and 54 months, AUXILIARY DO was the least stimulable category in the spoken modality. At 42 months, AUXILIARY DO category stimulability was significantly lower than COPULA BE category stimulability, $b = 0.77$, $SE = 0.33$, $p = 0.02$, AUXILIARY BE category stimulability, $b = 0.94$, $SE = 0.32$, $p < 0.01$, *-ed* category stimulability, $b = 1.24$, $SE = 0.37$, $p < 0.01$, and *-3s* category stimulability, $b = 1.27$, $SE = 0.37$, $p < 0.01$ in the spoken modality. At 54 months, AUXILIARY DO category stimulability was significantly lower than COPULA BE category stimulability, $b = 1.59$, $SE = 0.67$, $p = 0.02$, AUXILIARY BE category stimulability, $b = 2.46$, $SE = 0.70$, $p < 0.01$, *-ed* category stimulability, $b = 3.19$, $SE = 0.82$, $p < 0.01$, and *-3s* category stimulability, $b = 2.06$, $SE = 0.80$, $p = 0.01$ in the spoken modality. The odds of a child giving a correct response on any given COPULA BE stimulability probe item in the spoken modality were 2.151 times as high as the odds of a child giving a correct response on any given AUXILIARY DO stimulability probe item in the spoken modality at 42 months and 4.887 times as high at 54 months. The odds of a child giving a correct response on any given AUXILIARY BE stimulability probe item in the spoken modality were 2.557 times as high as the odds of a child giving a correct response on any given AUXILIARY DO stimulability probe item in the spoken modality at 42 months and 11.76 times as high at 54 months. The odds of a child giving a correct response on any given *-ed* stimulability probe item in the spoken modality were 3.471 times as high as the odds of a child giving a correct response on any given AUXILIARY DO stimulability probe item in the spoken modality at 42 months and 24.32 times as high at 54 months. The odds of a child giving a correct response on any given *-3s* stimulability probe item in the spoken modality were 3.546 times as high as the odds of a child giving a correct response on any given AUXILIARY DO stimulability probe item in the spoken modality at 42 months and 7.810 times as high at 54 months.

At 54 months, *-ed* category stimulability was significantly higher than COPULA BE category stimulability, $b = 1.60$, $SE = 0.67$, $p = 0.02$. The odds of a child giving a correct response on any given *-ed* stimulability probe item in the spoken modality were 1.218 times as high as the odds of a child giving a correct response on any given COPULA BE stimulability probe item in the spoken modality at 54 months. No other significant differences were found between category stimulability scores in the spoken modality at 42 months or at 54 months, *p*s > 0.05.

4.3.3.2 Model for graphic symbol modality.

Morpheme category stimulability in the graphic symbol modality was modulated by a significant age-by-morpheme category interaction, indicating that morpheme category stimulability grew at different rates across categories. AUXILIARY BE category stimulability increased at a significantly faster rate than *-3s* Category stimulability, $b = 0.17$, $SE = 0.08$, $p = 0.02$ and COPULA BE category stimulability, $b = 0.18$, $SE = 0.07$, $p = 0.01$. No other significant age-by-morpheme category interactions were found in the graphic symbol modality , *p*s > 0.05.

The main effect of age was a significant predictor of AUXILIARY BE category stimulability, $b = 0.31$, $SE = 0.06$, $p < 0.01$, *-ed* category stimulability, $b = 0.18$, $SE = 0.06$, $p = 0.01$, AUXILIARY DO category stimulability, $b = 0.17$, $SE = 0.05$, $p < 0.01$, *-3s* category stimulability, $b = 0.13$, $SE = 0.05$, $p = 0.01$, and COPULA BE category stimulability, $b = 0.12$, $SE = 0.04$, $p = 0.01$ in the graphic symbol modality. The odds of a child giving a correct response on any given AUXILIARY BE stimulability probe item increased by a factor of 1.357 per month in the graphic symbol modality. The odds of a child giving a correct response on any given *-ed* stimulability probe item increased by a factor of 1.194 per month in the graphic symbol modality. The odds of a child giving a correct response on any given AUXILIARY DO stimulability probe item increased by a factor of 1.190 per month in the graphic symbol modality. The odds of a child giving a correct response on any given *-3s* stimulability probe item increased by a factor of 1.142 per month in the graphic symbol modality. The odds of a child giving a correct response on any given *-3s* stimulability probe item increased by a factor of 1.131 per month in the graphic symbol modality.

The main effect of morpheme category was a significant predictor of morpheme category stimulability in the graphic symbol modality at 30 months. COPULA BE category stimulability was significantly higher than AUXILIARY BE category stimulability at 30 months, $b = 3.16$, $SE = 1.18$, $p = 0.01$. The odds of a child giving a correct response on any given COPULA BE stimulability probe item in the spoken modality were 23.54 times as high as the odds of a child giving a correct response on any given AUXILIARY BE stimulability probe item in the graphic symbol modality at 30 months. No other significant differences were found between category stimulability scores in the graphic symbol modality at 30 months, ps > 0.05.

At 42 months, COPULA BE was the most stimulable category in the graphic symbol modality. COPULA BE category stimulability was not significantly higher than *-3s* category stimulability in the graphic symbol modality at 42 months, $b = 0.78$, $SE = 0.42$, $p = 0.06$. However, COPULA BE category stimulability was significantly higher than AUXILIARY BE category stimulability, $b = 0.97$, $SE = 0.45$, $p = 0.03$, *-ed* category stimulability, $b = 1.12$, $SE = 0.47$, $p = 0.02$, and AUXILIARY DO category stimulability, $b = 1.21$, $SE = 0.47$, p 0.02 in the graphic symbol modality at 42 months. At 42 months, the odds of a child giving a correct response on any given COPULA BE stimulability probe item in the graphic symbol modality were 2.629 times as high as the odds of a child giving a correct response on any given AUXILIARY BE stimulability probe item in the graphic symbol modality, 3.069 times as high as the odds of a child giving a correct response on any given *-ed* stimulability probe item in the graphic symbol modality, and 3.345 times as high as the odds of a child giving a correct response on any given AUXILIARY DO stimulability probe item in the graphic symbol modality. No other significant differences were found between measures of morpheme category stimulability in the graphic symbol modality at 42 months, *p*s > 0.05.

At 54 months, AUXILIARY BE was the most stimulable category in the graphic symbol modality. AUXILIARY BE category stimulability was significantly higher than COPULA BE category stimulability, $b = 1.23$, $SE = 0.57$, $p = 0.03$, *-ed* category stimulability, $b = 1.69$, $SE = 0.68$, $p = 0.01$, AUXILIARY DO category stimulability, $b = 1.82$, $SE = 0.60$, $p < 0.01$, and *-3s* category stimulability, $b = 1.89$, $SE = 0.70$, $p = 0.01$. At 54 months, the odds of a child giving a correct response on any given AUXILIARY BE stimulability probe item in the graphic symbol modality were 3.406 times as high as the odds of a child giving a correct response on any given COPULA BE stimulability probe item in the graphic symbol modality, 5.431 times as high as the odds of a child giving a correct response on any given *–ed* stimulability probe item in the graphic symbol modality, 6.158 times as high as the odds of a child giving a correct response on any given AUXILIARY DO stimulability probe item in the graphic symbol modality, and 6.596 times as high as the odds of a child giving a correct response on any given *-3s* stimulability probe item in the graphic symbol modality. No other significant differences were found between measures of morpheme category stimulability in the graphic symbol modality at 54 months, *p*s $>$ 0.05.

4.4 EXPERIMENT 3 RESULTS

4.4.1 Participants.

Two pediatric AAC speakers participated in Experiment 3. Participant A was an 8 year, 5 month old male. Participant B was an 8 year, 9 month old male. Both participants had severe motor speech impairments secondary to cerebral palsy, which made it difficult to produce intelligible verbal speech. Participant demographics, screening results and TACL-3 raw scores are shown in Table 4.3. Both participants used a high-tech SGD as their primary means of expressive language production and communicated with a variety of other communication strategies, including signs, gestures, and vocalizations.

Both participants used high-tech Accent 1000 devices (Prentke Romich Company) with the Unity™ 84-sequenced Minspeak™ Application Program (Semantic Compaction Systems) as their personal SGD. Given that both participants used the same language application program and the same SGD hardware, primary, secondary, and tertiary features (Hill, 2010) were effectively matched across participants. Detailed definitions of device features discussed in the following paragraphs are provided in the glossary for reference.

Primary device features relate to the availability and organization of language content on the participants SGDs. The participants had access to a full range of language representation methods, including alphabet-based methods (spelling and word prediction), single meaning pictures, and semantic compaction. Spelling and word prediction were available at any time to produce an infinite set of spelled words. Single meaning pictures and semantic compaction were available at any time to produce pre-stored vocabulary words. The dictionaries of pre-stored vocabulary words on the participants SGSs included high frequency words, such as forms of BE, forms of DO, personal pronouns, and irregular verbs. The dictionaries also included many low frequency words, such as a variety of regular verbs, nouns, and adjectives. In many cases, the dictionaries on the participants SGSs included multiple inflected forms of the same lemma (e.g., *want*, *wants*, *wanted*, *wanting*). The participants used all language representation methods for spontaneous novel utterance generation. They also used single meaning pictures and Semantic Compaction to produce a few pre-programmed utterances, which were excluded from analysis.

Secondary device features relate to control interfaces, user interfaces, and outputs of the participants' SGDs. The participants' devices included a 10.1" touch-screen display. Symbols were displayed in a grid pattern with 84 fixed locations. Both participants used an Icon Prediction feature, which provided visual feedback differentiating multi-meaning icon sequences that can be selected to produce pre-stored vocabulary words from multi-meaning icon sequences that do not correspond to any stored content. Both participants made direct selections on their devices using their fingers to activate the devices' touch-screen control interface. Both participants' devices produced synthesized speech output, displayed generated utterances as text in a message window, and included a LAM feature that supported automated language sample collection.

Tertiary device features included peripheral technologies to support device use. Both participants had mounting systems for mounting their SGDs to their power wheelchairs. Participant B also used a keyguard.

Measures of sample size and utterance length from the participants' language samples are shown in Table 4.4. Language sample duration ranged from 2 hours, 12 minutes in Participant B's posttest sample to 2 hours, 39 minutes in Participant A's posttest sample. Sample size ranged from 78 multi-morpheme utterances in Participant B's posttest sample to 122 multi-morpheme utterances in Participant A's pretest sample. Measures of utterance length for Participant A were within the range of Brown's Stage III in both of his language samples. Measures of utterance length for Participant B were within the range of Brown's Stage IV in his pretest sample and within the range of Brown's Stage V in his posttest sample.

Table 4.3. Experiment 3 participant demographics.

	Participant A	Participant B
Age (Years;Months)	8;5	8;9
Sex	Male	Male
Hearing	Mild right hearing loss	WNL
CDI Words & Sentences Vocab Checklist[a]	676	576
EarlyLAMBaseline Word Inventory		
100 core words (understood, used)	100, 100	100, 97
Nouns (understood, used)	> 75, > 75	> 75, > 75
TACL-3 subtests (raw score)		
Vocabulary	32	36
Grammatical Morphemes	21	27
Elaborated Phrases & Sentences	19	26
Language application program on SGD	Unity 84	Unity 84
Selection method	Direct selection w/fingers	Direct selection w/fingers

Note. WNL = within normal limits.

[a]Parents given modified instructions to "Please go through the list and mark the words you think your child understands."

Table 4.4. Measures of sample size and utterance length.

	Participant A		Participant B	
	Pretest	Posttest	Pretest	Posttest
Communication Partners	Dad Baby Sitter	Dad Dr. Hill (via Skype) Researcher	Mom Researcher	Baby Sitter Researcher
Total duration (hours:minutes)	2:14	2:39	2:20	2:12
Multi-morpheme utterances	122	110	97	78
Shortest multi-morpheme utterance	2	2	2	2
Longest multi-morpheme utterance	10	9	13	13
MSL	3.85	3.90	5.45	6.05
Predicted MLUm	2.88	2.92	4.27	4.79

Note. MSL = mean syntactic length (Klee & Fitzgerald, 1985); Predicted MLUm = predicted mean length of utterance in morphemes (Kovacs & Hill, 2017).

Table 4.5. Pretest and posttest category productivity scores.

Morpheme Category	Participant A		Participant B	
	Pretest	Posttest	Pretest	Posttest
COPULA BE	3	5[b]	5[b]	5
-3s	0	1	1[a]	3
-ed	1	4	1	1
AUXILIARY DO	3[a]	0	3	5
AUXILIARY BE	4[a]	5	5[a]	5[a]
Total Productivity Score	11[a]	15	15[a]	19[a]

Note. [a]Includes added points for tense markers produced during elicitation probe. [b]Production of tense marker during elicitation probe does not increase score because score is at ceiling.

4.4.2 Pretest Assessment Results

Category productivity scores and productivity scores from both participants' language samples are shown in Table 4.5. Participant A's pretest and posttest tense marker assessment results are summarized in Table 4.6. Participant B's pretest and posttest tense marker assessment results are summarized in Table 4.7. Spontaneous novel utterances containing tense markers in the participants' pretest and posttest language samples are provided for reference in Appendix E.

In his pretest language sample, Participant A used a total of 6 different tense markers in a total of 9 sufficiently different contexts: copula *is* and *are*, one *–ed* verb produced as a single-word utterance, auxiliary *did* and auxiliary *are* and *was*. Participant A produced two additional tense markers in one context each during the pretest elicitation probe task: auxiliary *do* and auxiliary *is*. He produced copula and auxiliary *am* in the pronominal contraction *I'm*, which Hadley and Short (2005) do not count as a context of tense marker use. He did not produce copula *was* or *were*, *-3s* verbs, auxiliary *does*, or auxiliary *were* in his pre-test language sample or elicitation probe.

Participant A was stimulable for a total of 10 different tense markers. He was stimulable for 4/6 of the tense markers that he produced in his pretest language sample (copula *is*, *-ed*, auxiliary *did* and auxiliary *are*), one of the additional tense markers produced during his pretest elicitation probe task (auxiliary *is*), and both tense markers that were only produced in pronominal contractions (copula and auxiliary *am*). He also was stimulable for three additional tense markers that were not produced in either his pretest language sample or his pretest elicitation probe (copula *was* and *were* and *-3s*).

Table 4.6. Summary of Participant A's pretest and posttest tense marker assessment

TM	Contexts of Use		Gram. Judgment[c,d]	Stimulable[d]	Potential Use[e]
	Pre-Test	Post-Test			
cop *is*	2	4	4	Yes	Yes
cop *am*	0[a]	0[a]	2	Yes	Yes
cop are	1	1[b]	3	No	Yes
cop was	0	9	3	Yes	Yes
cop were	0	0	4	Yes	Yes
-3s	0	1	1	Yes	Yes
-ed	1	4	2	Yes	Yes
aux does	0	0	2	No	Yes
aux *do*	1[b]	0	3	No	Yes
aux did	2	0	1	Yes	Yes
aux *is*	1[b]	4	2	Yes	Yes
aux *am*	0[a]	0[a]	3	Yes	Yes
aux are	2	1	1	Yes	Yes
aux was	1	0	3	No	Yes
aux were	0	0	3	No	Yes
Total	11	24			

Note. TM = tense marker. [a]Produced in pronominal contraction "*I'm*," but not as an uncontracted form "*I am*." [b]Produced in elicitation probe, but not language sample. [c]Number of correct responses out of four items on grammaticality judgment task. [d]Measured at pre-test only. [e]Potential use is reported if a tense marker can be selectively produced on Participant A's SGD.

Table 4.7. Summary of Participant B's pretest and posttest tense marker assessment

TM	Contexts of Use		Gram. Judgment[e,f]	Stimulable[f]	Potential Use[g]
	Pre-Test	Post-Test			
cop *is*	4	6	3	Yes	Yes
cop *am*	0[a]	1[c]	4	Yes	Yes
cop *are*	3	1	2	Yes	Yes
cop *was*	2	1	4	Yes	Yes
cop *were*	1[b]	1	3	Yes	Yes
-3s	1[b]	3	4	Yes	Yes
-ed	1	1	3	Yes	Yes
aux *does*	0	2	3	Yes	Yes
aux *do*	2	3	3	No	Yes
aux *did*	1	0[d]	3	Yes	Yes
aux *is*	1	1	3	Yes	Yes
aux *am*	0[a]	1[c]	3	Yes	Yes
aux *are*	2	3	4	No	Yes
aux *was*	1[b]	1[b]	3	Yes	Yes
aux *were*	1[b]	1[b]	3	No	Yes
Total	20	26			

Note. TM = tense marker. [a]Produced in pronominal contraction "I'm," but not as an uncontracted form "I am." [b]Produced in elicitation probe, but not language sample. [c]Approximated in elicitation probe using phonologically similar form "I an." [d]Produced with over marking error in elicitation probe. [e]Number of correct responses out of four items on grammaticality judgment task. [f]Measured at pre-test only. [g]Potential use is reported if a tense marker can be selectively produced on Participant B's SGD.

In his pretest language sample, Participant B used eight different tense markers in a total of 16 sufficiently different contexts: copula *is*, *are*, and *was*, one *-ed* verb, auxiliary *do* and *did*, and auxiliary *is* and *are*. Participant B produced four additional tense markers one context each during the elicitation probe: copula *were*, *-3s*, auxiliary *was* and auxiliary *were*. He produced copula and auxiliary *am* in the pronominal contraction *I'm*, which Hadley and Short (2005) do not count as a context of tense marker use.

The only tense marker that Participant B did not produce in any form during the pretest language sample and elicitation probe tasks was auxiliary *does*. Auxiliary *does* is one of four tense markers that are used exclusively to mark the present tense in utterances with third person singular subjects such as the personal pronouns *he*, *she*, and *it*. A follow-up analysis of Participant B's use of personal pronouns found that he used second person (*you*) and first person (*I*) pronouns as subjects in a majority of his utterances. He also produced several utterances containing *it* or a subject noun phrase. Zero productions of *he* or *she* were found. The absence of these third person singular pronouns in the pretest language sample was consistent with the absence of a corresponding third person present tense singular tense marker.

Participant B was stimulable for a total of 12 different tense markers. He was stimulable for 6/8 tense markers that he produced in his pretest language sample (copula *is*, *are*, and *was*, one *-ed* verb, auxiliary *did*, and auxiliary *is*), three of the additional tense markers produced during his pretest elicitation probe (copula *were*, *-3s*, and auxiliary *was*), and both tense markers that were only produced in pronominal contractions (copula and auxiliary *am*). He also was stimulable for auxiliary *does*.

The review of the pre-stored vocabulary on the participants' language application programs found that the participants' SGDs supported selective production of all 15 tense markers and had potential to support productive use of the tense and agreement system. A potential productivity score of 5 was obtained in all categories, for a total potential productivity score of 25. Utterances generated in the participants' language application programs during this review process are shown in Appendix F.

4.4.3 Intervention Sessions

4.4.3.1 Participant A

Copula *was* was selected as a target tense marker for the intervention phase with Participant A because it met the selection criteria in (28). Copula *is* was selected as a contrasting non-target tense marker to practice during treatment sessions because of the reasons in (31).

31. Reasons for selection of copula *is* as a non-target contrasting tense marker.

   a. Copula *is* and copula *was* are a minimal pair that share many grammatical features and are part of the same morpheme category (COPULA BE). In sentences containing third person singular subjects, tense is the only grammatical feature differentiating copula *is* from copula *was*.

   b. Participant A demonstrated comprehension of copula *is* by responding correctly on all four corresponding items of the grammaticality judgment task.

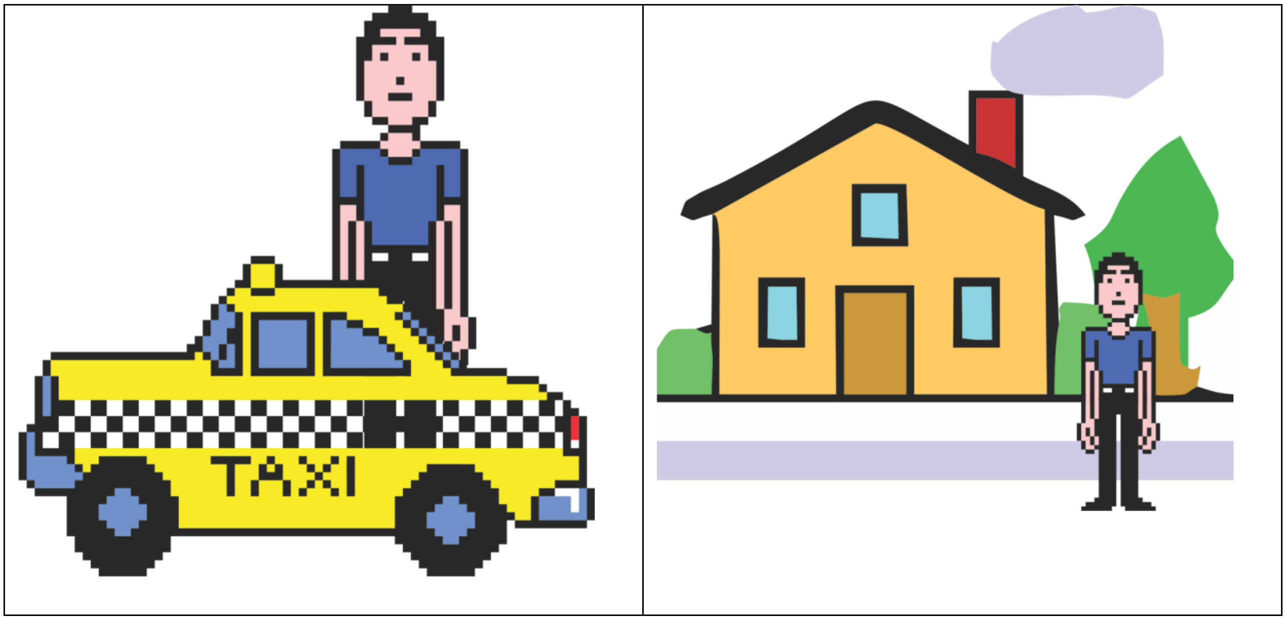
   c. Participant A used copula *is* correctly twice during his pretest language sample.

   d. Participant A used copula *is* instead of copula *was* in the pretest elicitation probe, indicating potential confusion between present tense and past tense copula forms.

Each treatment session consisted of a shared reading activity focused on increasing production of copula *was*. A custom storybook providing robust scripted opportunities for using copula *was* and contrasting opportunities for using copula *is* was developed for the treatment sessions. This storybook consisted of illustrated pages about people and things at the airport. This topic was selected because Participant A and his babysitter role-played an elaborate story about traveling to the airport during his pretest language sample. Each page used two pictures to illustrate a simple episode about a subject (person or thing) at the airport. The first picture illustrated the present status of the subject. The second picture illustrated the past status of the subject. Participant A and the author worked together to discuss each picture. Discussing both pictures provided Participant A with one highly contextualized opportunity to produce copula *is* in a sentence describing the current status of the subject and one highly contextualized opportunity to produce copula *was* in a sentence describing the prior status of the subject. A sample page and script with target responses is shown in Figure 4.9.

Participant A's storybook initially had 10 illustrated pages, and provided 10 scripted opportunities to produce copula *is* and *was* in a variety of subject-tense marker combinations. Participant A became familiar with the ten pages of his storybook after two sessions and began working through his storybook very quickly. At times, he produced familiar target responses before the researcher finished presenting a sentence prompt. New pages were added progressively to provide diverse opportunities to produce the target form and help maintain interest in the task. Added stimuli included supplemental pages for the airport storybook and photographs from family vacations.

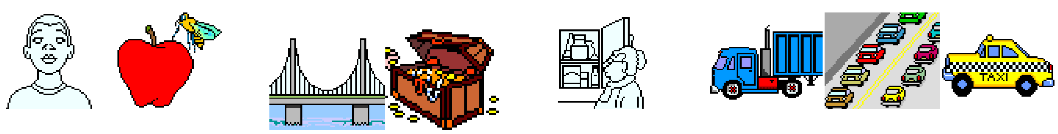
In some sessions, Participant A initiated spontaneous conversations about unrelated topics. These conversations were shaped to provide additional opportunities for Participant A to produce copula *was* and *is* when possible. To the extent possible, the author provided prompts in spontaneous conversations that followed the format of sentence prompts in the airport storybook



Sentence Prompt (Author points to Picture 1): Let's talk about this man. Where is he now?

Target Response and Icons:

He is | In | A | Taxi



Sentence Prompt (Author points to Picture 2): Where was he before?

Target Response and Icons:

He | Was | At | home

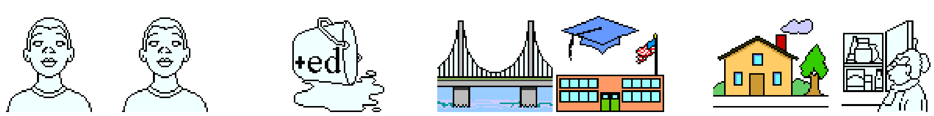


Figure 4.9. Sample storybook page for Participant A

reading to elicit spontaneous productions of the target form and highlight distinctions between past and present tense scenarios in real-time.

Participant A's parents were encouraged to provide opportunities for him to use copula *was* and *is* between sessions and provided with copies of his airport storybook. They were alsoencouraged to model correct productions of copula *was* and *is* using both verbal productions and aided language stimulation.

A 10-item probe task was administered after each treatment session to measure Participant A's independent use of copula *was* in the absence of prompting and reinforcement. Probe items were designed following the template of storybook pages in Participant A's airport storybook. The probe items presented brief scenarios about a variety of human characters (male and female), animals (cat, dog, horse), and inanimate objects (car, airplane) to probe use of copula *was* in a wide range of semantic contexts. For each probe item, the researcher presented one verbal cue in an attempt to elicit a production of a sentence containing copula *was*. No additional prompts or reinforcements were provided. Participant A was credited with correct responses for sentences containing a third person singular subject and copula *was*, even if other words, such as a preposition or article were omitted.

Participant A responded to at least one probe item using auxiliary *was* (e.g. The lady was driving) in each session. Three probe items were replaced after session 4 because Participant A consistently responded to these items using auxiliary *was*.

4.4.3.2 Participant B.

Auxiliary *does* was selected as a target tense marker for the intervention phase with Participant B because it met the selection criteria in (28). Copula *is* was selected as a contrasting non-target tense marker to practice during treatment sessions because of the reasons in (32).

32. Reasons for selection of copula *is* as a non-target contrasting tense marker.

a. Copula *is* and auxiliary *does* have the same functional processing features. Both forms are used to mark the present tense in sentences containing third person singular subjects.
b. Participant B demonstrated comprehension of copula *is* by responding correctly on three of the four corresponding items of the pretest grammaticality judgment task.
c. Participant B used copula *is* in ten utterances during his pretest language sample.
d. In his pretest language sample, Participant B used copula *is* in a variety of yes/no, where, and when questions.

Each treatment session consisted of a shared reading activity focused on production of auxiliary *does*. Two custom storybooks providing robust scripted opportunities for using auxiliary *does* and copula *is* were developed and used for treatment sessions. Participant B's custom storybooks consisted of illustrated pages about a TV chef from The Food Network preparing a meal. This topic was selected because Participant B and his mother had an elaborate conversation about cooking meals during his baseline language sample.

Two storybooks were used to provide Participant B with robust opportunities to ask questions using both feminine and masculine pronouns. One book was used in Sessions 1-4 and a second book was used in later sessions. The first book featured pages about a female chef baking macaroni and cheese and provided robust opportunities to ask questions using *does she*. The second book featured pages about a male chef grilling Swiss and mushroom burgers provided robust opportunities to ask questions using *does he*.

Each page in each storybook included a picture of the chef and ingredients or equipment used in one step of the recipe. A sample page and prompts for a question using auxiliary *does* is shown in Figure 4.10. Participant B and the researcher worked together to discuss each picture using a modified shy puppet activity (Crain & Thornton, 2000; Rice & Wexler, 2001) that incorporated positive reinforcement and modeled productions of target forms using aided language stimulation. In this task, a puppet was used as a third communication partner. Participant B was prompted to ask the puppet questions about the chef or the recipe. The puppet then relayed the question to the chef using nonsense words and relayed the chef's answer to Participant B. The puppet relayed all questions using third person singular forms to the chef, including several silly un-scripted questions generated by Participant B. This task provided Participant B with highly contextualized opportunities to ask a wide range of questions using both auxiliary *does* and copula *is*.

Either the author or the puppet acknowledged and responded to each of the participant's utterances. When Participant B produced statements or asked questions using second person tense markers, the author responded directly, recasting Participant B's utterances into prompts for questions about the chef or recipe whenever possible. An example of a transcribed interaction from Session 5 showing a recast of Participant B's utterance and Participant B asking a puppet a question is transcribed in (33).

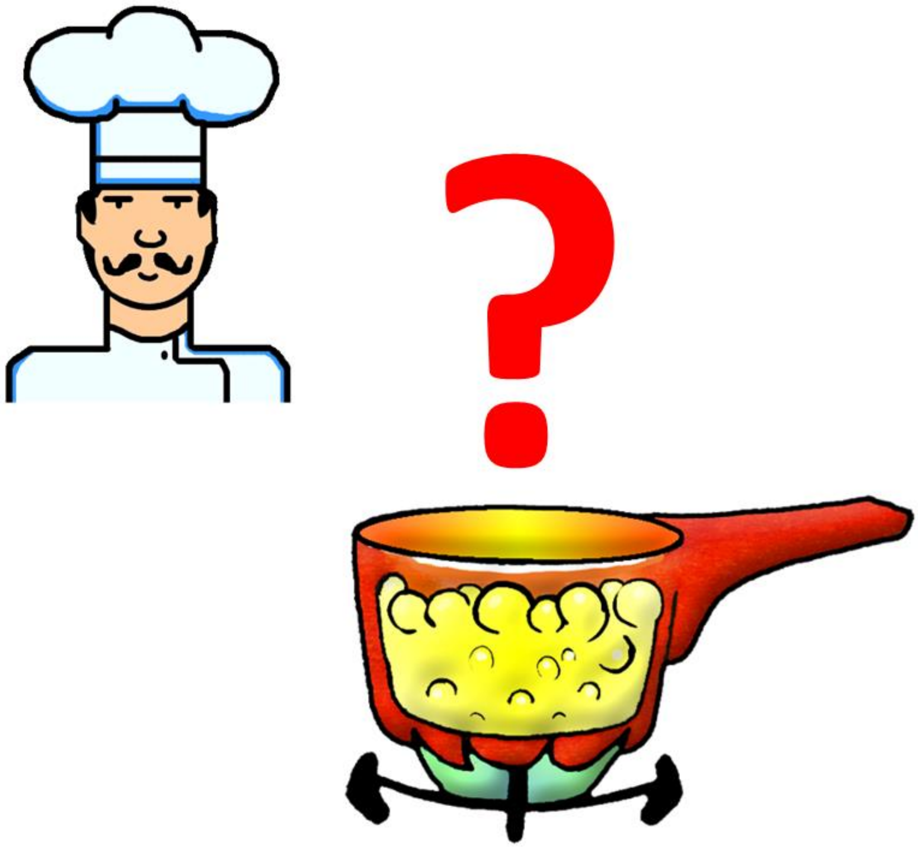
33. **Researcher:** He says he puts some Worcestershire sauce in there so he can cook some other ingredients. What do you think some of those other ingredients are?

Participant B: Dog food cat food and a person.

Researcher: You think he puts dog food cat food and a person in there? Ask Sleepy Bunny if he puts dog food in the pan.

Participant B: Does he put dog food in the pan?

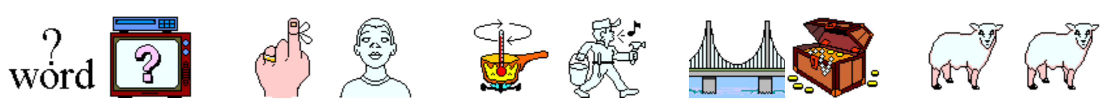
Sleepy Bunny puppet: I'll find out. (Sleepy Bunny relays question to chef Bobby using nonsense words) He says no. Not dog food.



Sentence Prompt: Bobby has a pan. I wonder what he cooks in the pan. Ask Sidney what he cooks in it.

Target Response and Icons:

What does he cook In It



Sidney (puppet): I'll find out. (Relays question to chef using nonsense words).

Sidney's Answer: Bobby says he cooks mushrooms and onions in the pan.

Figure 4.10. Sample storybook page for Participant B

At times, Participant B spontaneously prompted the author to ask the puppet questions. When this happened, the author modeled production of questions on Participant B's SGD using aided language stimulation.

A 10-item probe task was administered after each treatment session to measure Participant B's independent use of auxiliary *does* in the absence of prompting and reinforcement. The probe task was presented in the form of a shy puppet activity (Crain & Thornton, 2000; Rice & Wexler, 2001). The author introduced the probe task by inviting Participant B to help ask a puppet questions about new friends in the neighborhood. For each probe item, the author presented a picture of a character and directed Participant B to ask the puppet a question about the character in the picture. The puppet then relayed Participant B's question to the character in the picture using nonsense words and relayed an answer to Participant B. The characters in the probe task included a variety of male and female characters, animal characters, and inanimate objects to give Participant B opportunities to ask questions using *does he, does she* and *does it* or a variety of *does* + subject noun phrases. Two probe items were used to probe use of auxiliary *does* in each of five different question types (yes/no, what, when, where, and how).

4.4.4 Session by session results.

4.4.4.1 Participant A.

Participant A made rapid progress increasing his use of copula *was* in both treatment sessions and probe sessions. Based on his progress, the author, dissertation advisor, and Participant A's family decided to end the intervention phase after 8 pairs of treatment and probe sessions instead of continuing to complete the planned course of 10 pairs of treatment and probe sessions.

Participant A's use of copula *was* in treatment sessions is reported in Figure 4.11. The solid line reports the number of times he produced sentences containing copula *was* in each session, regardless of context. Participant A produced copula *was* seven times in his first session. This increased to a peak of 24 tokens in his fifth session. The dashed line reports the cumulative number of sufficiently different contexts in which he correctly produced copula *was* across sessions. Participant A used copula *was* in a wide range of contexts that became increasingly diverse over the course of eight sessions, demonstrating use of the target morpheme in a wide range of novel contexts. In his first session, he used copula *was* with two different subject pronouns (he was, she was). Over the course of 8 sessions, Participant A used copula *was* with four subject pronouns (he, she, it, I) and 23 different subject noun phrases for a total of 27 sufficiently different contexts. In sessions 4-8, Participant A used copula *was* in 7-10 contexts per session. Direct repetitions of the author's spoken utterances are not counted as either tokens of use or new contexts.

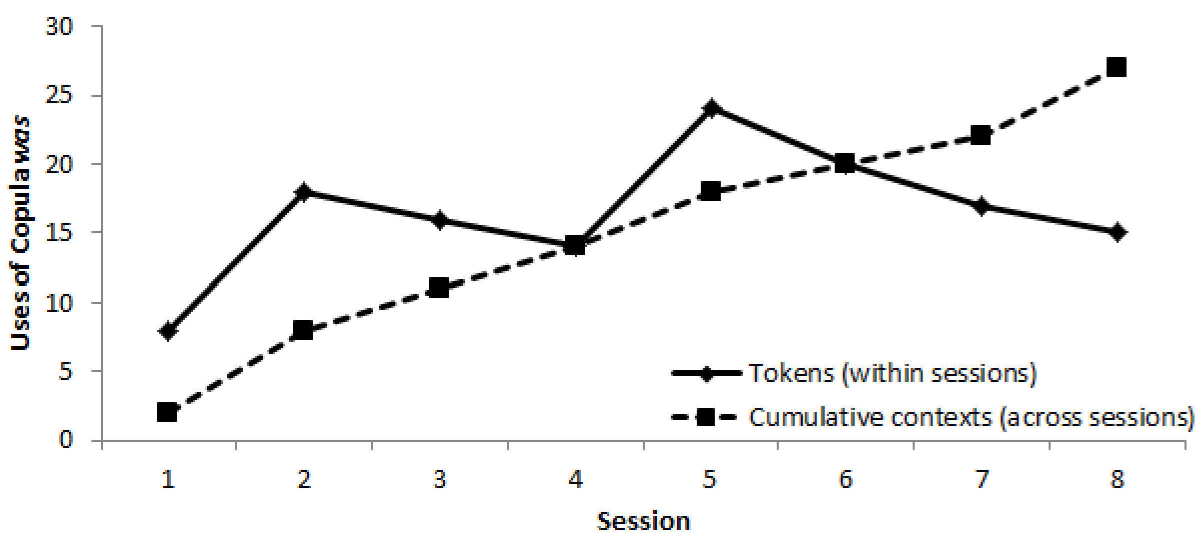


Figure 4.11. Participant A's use of copula *was* in treatment sessions

Participant A’s use of *was* in probe sessions is reported in Figure 4.12. The solid line reports Participant A’s responses using copula *was*. The dashed line reports Participant A’s responses using both copula *was* and auxiliary *was*. Participant A’s use of copula *was* in probe items increased from 5/10 items (50%) to 9/10 items (90%) over the course of 8 sessions.

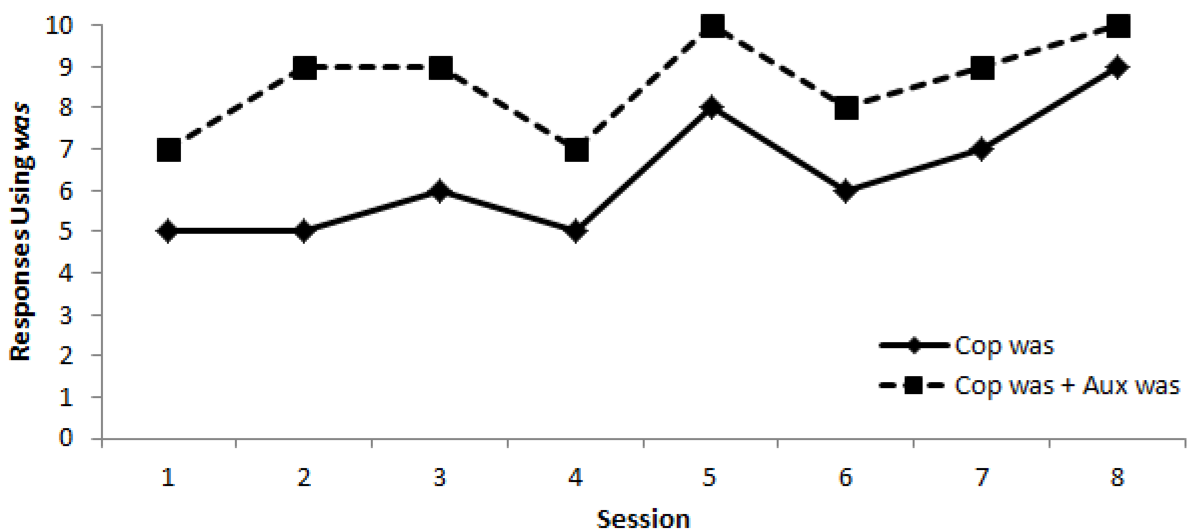


Figure 4.12. Participant A's use of copula *was* in probe sessions

4.4.4.2 Participant B.

Participant B made rapid progress increasing his use of auxiliary *does* in both treatment sessions and probe sessions. Based on his progress, the author, dissertation advisor, and Participant B’s family decided to end the intervention phase after 7 pairs of treatment and probe sessions instead of continuing to complete the planned course of 10 pairs of treatment and probe sessions.

Participant B’s use of auxiliary *does* in treatment sessions is reported in Figure 4.13. Direct repetitions of the author’s utterances are not counted towards any of the lines presented.

The solid line reports the number of times he produced grammatically correct sentences containing auxiliary *does* in each session, regardless of context. Participant B produced five grammatically correct sentences using auxiliary *does* in his first session. This increased to a peak of 13 grammatically correct sentences using auxiliary *does* in his fifth session.

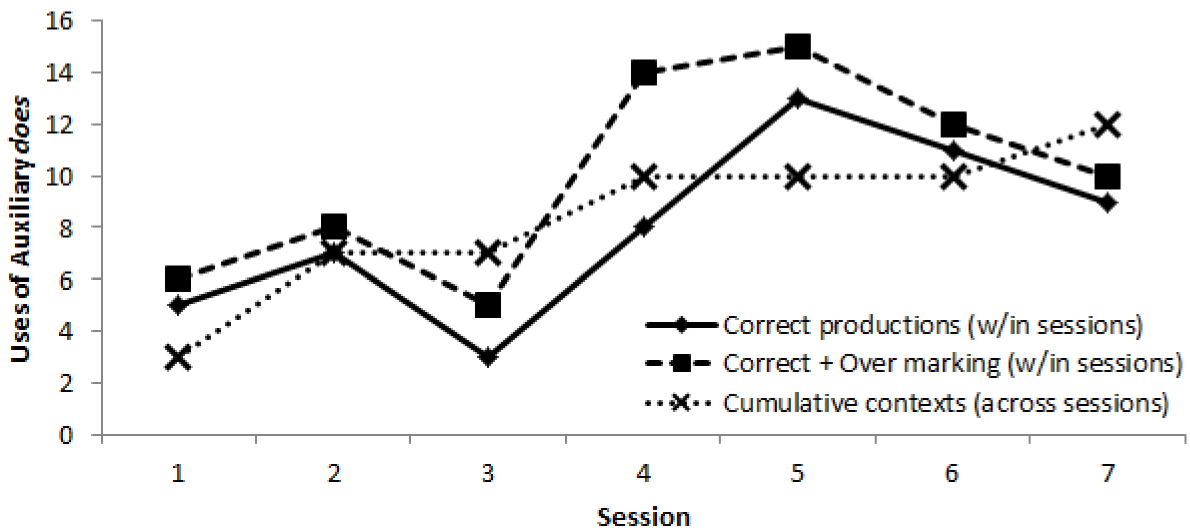


Figure 4.13. Participant B's use of auxiliary *does* in treatment sessions

Over the course of seven treatment sessions, Participant B used auxiliary *does* in all of the question types observed in his pretest language sample that require use of auxiliary *does*. Across treatment sessions, He correctly used auxiliary *does* in 29 yes/no questions, 16 *what* questions, 6 *how* questions, 1 *where* question and 1 *when* question. He did not produce any *why* questions in his pretest language sample or in any treatment sessions.

Across treatment sessions, Participant B used the negative form of auxiliary *does* (*doesn't*) in two grammatically correct statements and one grammatically correct *how* question. His earliest use of *doesn't* was observed in Session 4. Aided language stimulation demonstrating correct use of *doesn't* and opportunities to for Participant B to use *doesn't* were incorporated into sessions after the first production of *doesn't* was observed.

Participant B produced at least one ungrammatical sentence containing auxiliary *does* and an over marking error in each treatment session. The dashed line on Figure 4.13 reports a combined total of Participant B's grammatically correct sentences using auxiliary *does* and sentences containing auxiliary *does* and an over marking error in each treatment session. In these sentences, such as (34), Participant B marked tense twice in the same clause; once using auxiliary *does* and once using the third-person present tense singular form of the main verb.

34. Does the cat wants fish

Over marking errors are a type of normal developmental error (Maratsos & Kuczaj, 1978), which many children produce when they are learning how to formulate questions. Over marking errors provide evidence that a child understands the grammatical rules used for tense marking and occasionally over-applies those rules in sentences that follow a different pattern. Most children gradually out-grow over marking errors without intervention.

Participant B used auxiliary *does* correctly in a range of sufficiently different contexts that became increasingly diverse over the course of seven treatment sessions, demonstrating use of the target morpheme in a growing range of novel contexts. The dotted line on Figure 4.13 reports the cumulative number of sufficiently different contexts in which he correctly produced auxiliary *does* across treatment sessions. In his first session, Participant B used auxiliary *does* with two different subject pronouns (*does it*, *does she*) and one subject noun phrase (*does the macaroni*). Over the course of 7 sessions, Participant B used auxiliary *does* with three subject pronouns (*he*, *she*, *it*) and 9 different subject noun phrases for a total of 12 sufficiently different contexts.

Participant B's use of auxiliary *does* in probe sessions is reported in Figure 4.14. Participant B gave some correct responses and produced some over marking errors in every session. The solid line reports Participant B's correct responses using auxiliary *does*. Participant B's correct responses increase from 2/10 items (20%) to 9/10 items (90%) over the course of 7 sessions. The dashed line reports all of Participant B's responses using auxiliary *does*, including correct responses and responses with over marking errors. When responses containing over marking errors are included, Participant B used auxiliary *does* in 5/10 (50%) items in Session 1 and in 59/60 (98%) of the combined probe items across Sessions 2-7.

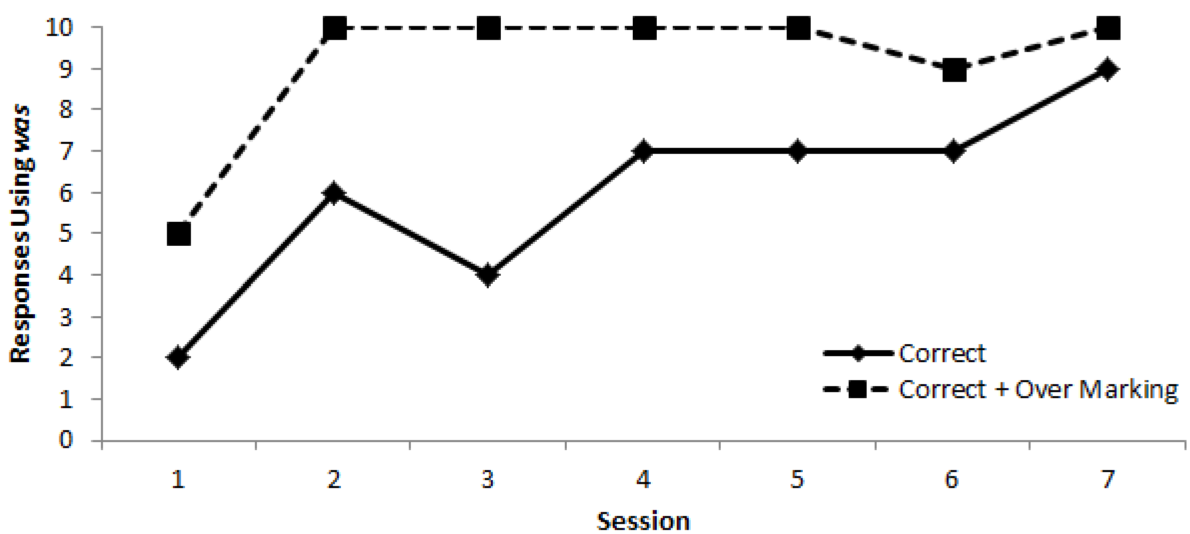


Figure 4.14. Participant B's use of auxiliary *does* in probe sessions

4.4.5 Posttest generalization.

4.4.5.1 Participant A.

In his posttest language sample, Participant A used six different tense markers in a total of twenty-three sufficiently different contexts: copula *is* and *was*, *-3s* verbs, *-ed* verbs, and auxiliary *is* and *are*. He also produced copula *are* in one context during the posttest elicitation probe.

Participant A used copula *was* in nine different contexts in his posttest language sample, and zero contexts in his pretest language sample. He also used copula *is* in four sufficiently different contexts in his post-test language sample, and only two contexts in his pre-test language sample. His new productions of copula *was* followed a short course of intensive intervention focused on the specific goal of increasing Participant A's production copula *was*. His increased productivity of copula *is* followed a period of high-intensity practice using copula *is* as a contrasting form to provide contextual information about copula *was*.

Participant A produced four *-ed* verbs in sentences during his posttest language sample. In his pretest language sample, Participant A only produced one *-ed* verb as a single word utterance. This increased productivity of *-ed* verbs is evidence of cross-morpheme generalization, or generalization of treatment effects for intervention focused on copula *was* to another past-tense morpheme with different positional processing features that was not targeted in intervention.

Evidence of cross-morpheme generalization to third person present-tense singular forms was also found. Participant A produced one *-3s* verb in a sentence during his posttest language sample and zero *-3s* verbs during his pretest language sample. Participant A used auxiliary *is* in four sufficiently different contexts in his posttest language sample. In his pretest assessment, he only used auxiliary *is* during the elicitation probe task. These examples of increased productivity are evidence of cross-morpheme generalization from the non-target contrasting tense marker (copula *is*) to other tense markers with shared functional processing features that were not addressed in intervention.

4.4.5.2 Participant B.

In his posttest language sample, Participant B used ten different tense markers in a total of twenty-two sufficiently different contexts: copula *is*, *are*, *was* and *were*, one *–ed* verb, three *-3s* verbs, auxiliary *does* and *do*, and auxiliary *is* and *are*. He also produced auxiliary *was* and *were* in one context each during the elicitation probe task.

Participant B used auxiliary *does* in two sufficiently different contexts in his posttest language sample, and zero contexts in his pretest language sample. He also used copula *is* in six sufficiently different contexts in his posttest language sample, and only four contexts in his pretest language sample. His new production of auxiliary *does* followed a short course of intensive intervention focused on the specific goal of increasing Participant B's production auxiliary *does*. His increased productivity of copula *is* followed a period of high-intensity practice using copula *is* as a contrasting form to provide contextual information about auxiliary *does*.

Evidence of cross-morpheme generalization to other third person present-tense singular forms was also found. Participant B produced the *-3s* form of three different verbs in spontaneous novel utterances during his post-test language sample. He only used the *-3s* form of a verb in the elicitation probe during his pre-test assessment.

The personal pronouns used in treatment activities (*she* and *he*) were produced throughout the post-test language sample. Participant B used *she* as the subject of one utterance and *he* as the subject of six utterances.

Participant B did not correctly produce copula *am*, auxiliary *am* or auxiliary *did* in his posttest language sample or elicitation probe. In the elicitation probe, he produced a sentence containing auxiliary *did* and an over marking error. In the elicitation probe, he produced a phonological approximation for both copula and auxiliary *am* (*I an*), suggesting that he may not have been familiar with the icon sequences used to produce *I am*/*am I*.

For Participant B, both the target and contrasting non-target tense marker were third person present tense singular forms. All cases of cross morpheme generalization in Participant B's data were to tense markers that shared the [-Past] feature in functional processing. The most robust examples of cross-morpheme generalization observed in Participant B's data were to tense markers that shared all functional processing features with auxiliary *does* and copula *is*.

5.0 DISCUSSION

5.1 COMPARISONS ACROSS MORPHEME CATEGORIES

In Experiment 1, the productivity scores of the different morpheme categories diverged as age increased, indicating that different tense-marked morpheme categories grow in productivity with different developmental trajectories. This rejects the null hypothesis of OI theory, which assumes that all five categories will grow together with similar developmental trajectories (Rice et al., 1998). COPULA BE productivity was at ceiling and consistently higher than the productivity of all four lexical verb categories. *-3s*, AUXILIARY DO, and *–ed* category productivity increased with age. AUXILIARY BE productivity increased at a slower, non-significant rate as age increased. At 30 months, the four lexical verb categories were equally productive. By 54 months, the productivity of AUXILIARY BE was significantly lower than the productivity of all other categories.

The partial developmental sequence observed across morpheme categories was partially consistent with GML theory's prediction that morpheme categories would increase in productivity together, provided that they are similar enough in positional processing (Rispoli et al., 2012). As expected, the [-V] COPULA BE category was consistently more productive than all four of the lexical verb categories that include a [+V] feature in positional processing. This replicates findings from prior studies of children ages 27 months (Rispoli et al., 2012) to 5 years

(Gladfelter & Leonard, 2013). The finding that COPULA BE productivity was at ceiling was not surprising. Mean COPULA BE productivity scores of 4.11 have been reported for a group of typically developing children at 33 months (Rispoli et al., 2012). It should be noted that COPULA BE productivity scores were at ceiling when the number of utterances in participant language samples was controlled. Patterns of productivity across morpheme categories may be obscured if productivity of other categories approach ceiling levels in larger language samples, as suggested by Gladfelter and Leonard (2013).

The finding that AUXILIARY BE was the least productive category at 54 months also was consistent with the predictions of GML theory. Rispoli et al. (2012) argued that the addition of a [+Prog] feature for progressive aspect makes AUXILIARY BE the most complex category in functional processing. The slower developmental trajectory for AUXILIARY BE reflects this added complexity, and the need to learn an additional grammatical feature.

The reason for the lack of significant differences between the four lexical verb categories at 30 months is not clear, given Rispoli et al.'s (2012) data showing that AUXILIARY BE is the least productive tense-marked morpheme category in children as young as 27 months. One possible explanation relates to differences in language sampling methods. Experiment 1 collected language samples with a controlled number of multi-morpheme utterances, while Rispoli et al. (2012) collected language samples over a controlled time period. The language samples analyzed by Rispoli et al. (2012) contained more utterances than the language samples in Experiment 1. These children may have produced tense markers in additional contexts in these additional utterances.

Another possible explanation relates to different methods of statistical analysis. Rispoli et al. (2012) used a mixed analysis of variance and a normal distribution to analyze morpheme category productivity at each time point in their study. Experiment 1 used generalized linear mixed modeling and a logistic distribution to model the growth of morpheme category productivity over time. In Rispoli et al.'s (2012) analysis, one morpheme category productivity score (range: 0-5) was entered into the analysis for each child in each category. In Experiment 1, each point that would be applied to a morpheme category productivity score was entered into the model as a separate binary observation and random intercepts were used to account for within-child clustering effects. The use of random intercepts may have accounted for clustering effects that cannot be explained by a mixed analysis of variance. Unfortunately, Rispoli et al.'s (2012) analysis cannot be replicated with the Experiment 1 data to verify this possible explanation because assumptions of normality and sphericity are violated.

The growth of the *-3s*, *-ed*, and AUXILIARY DO categories observed in Experiment 1 was not consistent with the prediction that the three lexical verb categories with nonspecific aspect would grow in productivity together (Rispoli & Hadley, 2011; Rispoli et al., 2012). At 30 months, *-3s*, *-ed*, AUXILIARY DO, and AUXILIARY BE had similar levels of productivity. After 30 months, *-3s* productivity increased at a faster rate than *–ed* productivity. Although the difference between growth rates was non-significant, *-3s* productivity and *–ed* productivity diverged as age increased. At 42 months, *-3s* and AUXILIARY DO significantly were more productive than *–ed* and AUXILIARY BE, replicating the pattern observed in Gladfelter and Leonard's (2013) sample of typically developing 4 and 5 year olds. At 54 months, *-ed* was significantly more productive than AUXILIARY BE and significantly less productive than *-3s*.

This divergence between *-3s* productivity and *–ed* productivity is particularly notable because *-3s* and *–ed* are the only morpheme categories differentiated by tense in the category system reported in the literature. When past and present tense forms of COPULA BE are collapsed into a single category, the composite COPULA BE productivity score does not differentiate between the productivity of present tense forms and the productivity of past tense forms. The same can be said for composite measures of AUXILIARY DO and AUXILIARY BE productivity (Hadley & Short, 2005) and composite measures of COPULA BE, AUXILIARY DO, and AUXILIARY BE accuracy (Rice et al., 1998). In these categories, children can potentially reach ceiling levels of productivity using only present tense forms or only past tense forms.

Some evidence suggests that the growth of productivity in tense markers with shared functional processing features is connected across morpheme categories. For example, Rispoli (2016) found that early productivity of copula *is* predicted later productivity of *-3s*, but not later productivity of *–ed*. Based on this finding, Rispoli (2016) argued that early development of one form can facilitate later development of other forms with shared functional processing features. A study that systematically differentiates between growth of past tense productivity and growth of present tense productivity may find that present tense markers grow in productivity faster than past tense markers. This would indicate that differences in functional processing features and differences in positional processing features can both contribute to different growth rates across morphemes.

The composite productivity scores of typically developing children also increased with age between 30 and 54 months as expected in language samples with a controlled number of multi-morpheme utterances. The children in the study used tense markers in a growing range of productive contexts even though sample size was held constant.

Some composite productivity scores were near ceiling levels. Increasing the number of utterances in the Experiment 1 language samples would likely lead to robust ceiling effects. This is consistent with Gladfelter and Leonard's (2013) argument that the discriminability of productivity measures may be specific to sample size, and that samples with too many utterances could suppress evidence of developmental or disordered patterns as participant performance approaches ceiling levels.

It may be possible to establish developmental norms for measures of productivity in language samples with a controlled number of utterances. Prior authors have modeled growth of tense productivity in toddlers using language samples with a controlled duration and a variable number of utterances (Hadley et al., 2014; Rispoli et al., 2009). Experiment 1 showed that it is possible to model growth of productivity in older children using language samples with a controlled number of utterances. Further research is needed to define the optimal number of utterances for language samples in a normative study and identify other factors that may need to be controlled in a normative study.

In Experiment 2, the patterns of category stimulability scores were more consistent across morpheme categories than expected, given the divergent category productivity scores observed in the language samples of Experiment 1. By comparison, only a few significant differences between category stimulability scores were found in either model.

In the model for the spoken modality, stimulability grew at a slower rate for AUXILIARY DO than for any other category. At 42 months and 54 months, AUXILIARY DO was significantly less stimulable than any other category in the spoken modality. At 54 months, -*ed* also was significantly more stimulable than COPULA BE.

In the model for the graphic symbol modality, stimulability grew at a faster rate for AUXILIARY BE than for any other category. At 30 and 42 months, COPULA BE was significantly more stimulable than AUXILIARY BE. At 54 months, AUXILIARY BE was significantly more stimulable than all other categories. COPULA BE also was significantly more stimulable than *–ed* and AUXILIARY DO at 42 months.

Neither of these patterns was consistent with the partial developmental sequence predicted by GML theory. The low AUXILIARY DO stimulability scores in the spoken modality are inconsistent with the prediction that stimulability of the three lexical verb categories with nonspecific aspect would grow together. The unexpectedly high AUXILIARY BE stimulability scores at 54 months in the graphic symbol modality were inconsistent with the prediction that stimulability of AUXILIARY BE would grow at a slower rate than stimulability of all other categories.

The unexpectedly low growth rate for AUXILIARY DO stimulability in the spoken modality may be related to a high-frequency error pattern that children tested in both modalities produced during the AUXILIARY DO stimulability task. In AUXILIARY DO stimulability probe items, children were expected to ask a penguin puppet (Sidney) yes/no questions about his animal friends using a target form of DO. The probe items were presented after a practice item in which children were expected to ask the penguin puppet questions about himself using *do you*. During the probe items, several children continued to ask the penguin puppet a question about himself using *do you* instead of asking questions about the penguin's animal friends. In the first AUXILIARY DO probe item, children were expected to ask the penguin a question about "the bird," a bald eagle in a photograph. Children may have continued asking *do you* questions after the practice item because the puppet they were expected to talk to and the animal friend they were expected to ask about were members of the same semantic category. This same error pattern may have contributed to a non-significant reduction in AUXILIARY DO stimulability in the graphic symbol modality.

The unexpectedly fast growth rate for AUXILIARY BE stimulability in the graphic symbol modality reflects a relatively abrupt age-related change in the children's responses on the AUXILIARY BE task. Seven of the children who completed the AUXILIARY BE stimulability tasks in the graphic symbol modality were younger than 42 months. These children did not give a single correct response on any AUXILIARY BE stimulability probe items. Eight of the children who completed the AUXILIARY BE stimulability tasks in the graphic symbol modality were at least 46 months old. Seven of these eight children gave correct responses on at least 5/6 of the stimulability probe items for present tense AUXILIARY BE forms. This abrupt change could potentially be an artifact of the strategy used to test AUXILIARY BE stimulability in the graphic symbol modality, sampling error, or a true developmental change. Additional studies with larger samples are needed to determine whether or not this abrupt change is evidence that children become stimulable for auxiliary *is*, *am* and *are* in the graphic symbol modality when they are between 42 and 46 months.

Other than the notable exceptions discussed above, stimulability scores grew with highly similar developmental trajectories across morpheme categories. This high similarity is most consistent with the null hypothesis of OI theory.

Although the model of category productivity was more consistent with a hypothesis that tense marker categories grow in partial sequence, the models of category stimulability were more consistent with a hypothesis that all tense marker categories grow together with highly similar developmental trajectories. The reasons for this are not clear. One possible explanation may be related to the representation of tense markers in these two measures.

The measures of productivity were obtained from play-based language samples, which may not have provided children with representative opportunities to produce the full set of 15 tense markers. Each category productivity score was a composite score reflecting the total number of sufficiently different uses across all tense markers within a category. Given the structure of these composite measures, it was possible for children to achieve high levels of productivity in the COPULA BE, AUXILIARY DO, and AUXILIARY BE categories using only a subset of the tense markers within each category. This phenomenon was observed in the data. Several children used copula *is* in at least five sufficiently different contexts, effectively reaching the ceiling for COPULA BE productivity by using copula *is* alone. The absence of copula *am*, *are*, *was*, or *were* in these children's language samples would have had no influence on their COPULA BE productivity scores. Similar limitations are encountered when measures of tense marker accuracy are collapsed across categories (e.g., Rice et al., 1998).

In contrast, measures of category stimulability were obtained during a series of structured tasks that provided children with representative opportunities to produce each of the 15 tense markers. Each category stimulability score reflected the number of times a child correctly produced a target tense marker across a fixed number of stimulability probe items. Every tense marker within a category was probed the same number of times so that all tense markers within a category would be equally weighted.

Productivity measures obtained from spontaneous language samples provide crucial information about children's use of morphosyntax in conversation, but may over-estimate the development of a morphosyntactic system if evidence of high productivity in a subset of forms obscures evidence of low productivity in other forms. These measures can be supplemented by structured tasks that ensure the child is provided with representative opportunities to produce all forms in a morphosyntactic system. More research is needed to characterize the development of morphosyntactic systems using synthesized approaches that combine spontaneous language samples and structured tasks.

5.2 USE AND STIMULABILITY OF INDIVIDUAL TENSE MARKERS

In Experiment 1, The number of different tense markers typically developing children produced increased with age between 30 and 54 months as expected in language samples with a controlled number of multi-morpheme utterances. Although the number of different tense markers produced increased, no children in the sample used all 15 tense markers within 150 multi-morpheme utterances. Only one child produced 11 different tense markers. The spontaneous language sampling task may not have provided opportunities for the children in Experiment 1 to use the full set of tense markers. Some children in Experiment 1 may have acquired the ability to produce additional tense markers.

A structured task that provides children with opportunities to produce all tense markers could potentially be used to identify these additional tense markers. In Experiment 3, elicitation probes found evidence that both participants had acquired the ability to produce tense markers that were not produced in their spontaneous language samples. This two-step process for identifying tense markers that children can produce provided opportunities for both spontaneous and elicited production following the two-step process used by Powell et al.'s (1991) procedures for identifying a child's phonetic inventory. The use of structured elicitation probes to supplement observations of spontaneous production provides a more complete inventory of tense markers children have acquired than the use of spontaneous observations alone.

In Experiment 2, no significant age-by-communication modality interaction effects were found for measures of tense marker stimulability, indicating that tense marker stimulability grows at the same rate across communication modalities. In addition, no significant age-by-modality interaction effects were found for measures of morpheme category stimulability, indicating that morpheme category stimulability grows at the same rate across communication modalities. In both types of analyses, the models with interaction terms did not have better fit than main effects models. For the sample in Experiment 2, the patterns of growth observed were qualitatively similar across communication modalities. This indicates that spoken and graphic symbol-based assessments of morpheme stimulability capture similar patterns of tense marker emergence and growth. The lack of interaction effects is interpreted with caution because a small sample of children at least 42 months old was used for the analysis of AUXILIARY BE category stimulability. This consistency across analyses provides initial evidence for a hypothesis that the growth of morpheme stimulability is not modality-specific or limited to unaided communication strategies. If this is the case, then stimulability testing should be considered as a potential strategy for assessing expressive morphology skills in verbally speaking children and pediatric AAC speakers.

The main effect of communication modality was a significant predictor of tense marker stimulability for six different tense markers. The odds of a child being stimulable for each of these six tense markers were higher for children tested in the spoken modality than they were for children tested in the graphic symbol modality. This is consistent with the prediction that children tested in the spoken modality would have higher performance because they were being tested in a familiar, naturally occurring modality.

Although the direction of modality effects is consistent across tense markers, it is striking that significant main effects of modality on tense marker stimulability were only found for six tense markers. One possible explanation of this pattern would be that differences in stimulability between communication modalities are more robust for some grammatical features than they are for others. Another possible explanation would be that some grammatical features are more difficult to represent in the graphic symbol modality than others. In either of these cases, the distribution of tense markers with significant main effects for communication modality would be expected to pattern with the distribution of grammatical features. This is not consistent with the distribution observed in the data. The tense markers with significant main effects for communication modality were equally distributed across past and present tense forms and four different morpheme categories. This distribution is not consistent with any clear pattern of functional processing features or positional processing features.

Another possible explanation is that some test items were not equivalent across communication modalities. One modality-specific error pattern was consistently observed during administration of the auxiliary *was* and *were* task. In the graphic symbol modality, several children used *sing* in place of the *–ing* suffix during the practice item, as in (34) and failed to complete subsequent probe items that required the use of *sing* as a lexical verb.

34. The fox was dance sing

A more comprehensive review of children's response patterns across modalities may reveal modality-specific error patterns or error patterns that occurred at a higher frequency in the graphic symbol modality.

The main effect of age was a significant predictor of tense marker stimulability for ten different tense markers. The odds of a child being stimulable for each of these ten tense markers increased with age as expected for children developing morphosyntactic skills over time.

The main effect of age was the only significant predictor of tense marker stimulability for seven different tense markers. The high prevalence of this pattern provides initial evidence that the growth of tense marker stimulability is not modality-specific or limited to unaided communication strategies. This also suggests that it may be possible to develop improved stimulability tasks for all tense markers that are sensitive to the emergence of tense marker stimulability and have high convergent validity across communication modalities. It may also be possible to develop similar sets of stimulability tasks for assessing stimulability of other morphosyntactic systems.

A first step towards this goal would be to conduct an error analysis to identify modality-specific error patterns and other patterns of incorrect responses that are not consistent with errors found in the spoken language of typically developing children. The most common tense marking errors reported in the speech of typically developing children are omission errors consistent with the optional infinitives described by Wexler (1994, 1998) and shown in (4). It should be possible for children to produce errors with this structure on all stimulability probe items in both modalities. An example of an incorrect response consistent with the structure of an optional infinitive is shown in (35).

35. Verbal cue: Andy painted the house red. Let's see what color the house was before. The

Response: *house yellow

Most other types of incorrect responses are inconsistent with error patterns observed in typical development. Any other recurrent pattern of incorrect responses in the data could potentially reflect problems with the design or presentation of test items. Once these patterns are identified, it will be possible to test hypotheses about limitations of existing test items and develop improved test items that yield patterns of responses that are more consistent with the patterns observed in typical development. These improved test items may have better sensitivity to the emergence of tense marker stimulability and better convergent validity across communication modalities.

As an example, several children tested in the graphic symbol modality for auxiliary *was* and *were* produced a modality-specific error pattern related to the use of *sing* and *-ing*, as described above. This modality specific error pattern may help explain why a surprisingly small number of children were stimulable for auxiliary *was* or *were* in the graphic symbol modality. In the graphic symbol modality, one 47-month old was stimulable for auxiliary *was* and one 54-month old was stimulable for both auxiliary *was* and *were*. In the spoken modality, 9 children were stimulable for each form. One could test a hypothesis that the phonological similarity between the lexical verb *sing* and the *–ing* suffix substantially increased the complexity of the auxiliary *was* and *were* stimulability task in the graphic symbol modality and led to a modality-specific error pattern. According to this hypothesis, replacing test items to eliminate this phonological similarity would eliminate the modality-specific error pattern and improve convergent validity across communication modalities.

Two items were used to probe stimulability of each form of COPULA BE, AUXILIARY DO, and AUXILIARY BE. In contrast, -*3s* and *–ed* were probed with five items each in order to administer enough test items to obtain -*3s* and *–ed* category stimulability scores. Results from the forms of COPULA BE, AUXILIARY DO, and AUXILIARY BE indicate that two back-to-back probe items in sufficiently different contexts are sufficient for probing stimulability of one tense marker. It may be possible to reduce the length and complexity of the *-3s* and *–ed* stimulability tasks by reducing each task to one practice item and two back-to-back probe items probing different lexical verbs.

Children completing the 5-item *-3s* stimulability task or the 5-item *–ed* stimulability task were required to select target lexical verb stem symbols from a relatively large array of options. This may have added a symbol recognition component that increased the task complexity in the graphic symbol modality. In addition, some of the lexical verbs used to probe the *-3s* form had irregular past tense forms. This mix of regular and irregular verbs may have influenced the *–ed* task in unknown ways. An error analysis of the *-3s* and *-ed* items may find evidence of modality-specific error patterns that should be accounted for in the design of simplified 2-item probe tasks.

5.3 STIMULABILITY OF MORPHEME CATEGORIES

The measures of tense marker stimulability indicated whether or not a child gave at least one correct response across all stimulability probe items corresponding to a given tense marker. In analyses of tense marker stimulability, significant main effects of communication modality were only found for a minority of the tense markers studied.

In contrast, measures of morpheme category stimulability measured the total number of correct responses across all stimulability probe items corresponding to a morpheme category. Significant main effects of communication modality were found for *-3s*, *-ed*, and AUXILIARY DO category stimulability scores. In these categories, children tested in the spoken modality gave more correct responses than children tested in the graphic symbol modality when age was held constant. The non-significant main effects of communication modality for COPULA BE category stimulability scores and for AUXILIARY BE category stimulability scores in children at least 42 months old followed this same pattern. This was consistent with the prediction that children would demonstrate higher performance in the spoken modality.

For the AUXILIARY BE category, there was a 10 month age difference between the youngest child who gave correct responses to stimulability probes in the spoken modality (32 months) and the youngest child who gave correct responses to stimulability probes in the graphic symbol modality (42 months). This age difference is difficult to interpret given the small sample size of the current study. A larger study is needed to determine whether this age difference is replicable and reflects a real developmental difference across communication modalities.

The main effects of communication modality on morpheme category stimulability scores indicated that explicit modeling of target morphemes in the graphic symbol modality was a necessary component of the assessment procedure, as suggested by Binger and Light (2008). Children tested in the graphic symbol modality generated fewer correct responses because they were unlikely to demonstrate correct tense marker use before they were exposed to a grammatically correct, contextualized model of tense marker use in the graphic symbol modality. After this model was provided, children demonstrated correct tense marker use on a subsequent probe item if they were stimulable. Elicitation tasks are not sufficient for assessing emergent knowledge of grammatical morphemes in the graphic symbol modality, especially when children have no prior exposure to relevant language models in the modality they are using for production. There is an inherent and requisite role for explicit modeling of target structures in the graphic symbol modality during assessment of developmental language skills of pediatric

AAC speakers just as there is an inherent role for aided language simulation in AAC intervention based on typical language development (Mirenda, 2008).

Significant main effects of age were found for morpheme category stimulability scores in 4/5 categories. The odds of a child giving a correct response on any given probe item in these categories increased with age as expected for children developing morphosyntactic skills over time. The main effect of age was not a significant predictor of *-3s* category stimulability. This may be related to limitations of the *-3s* stimulability task discussed earlier in this chapter. An item analysis and an error analysis are needed to determine whether the non-significant age effect for *-3s* category stimulability is related to item effects or unusual patterns of errors.

5.4 EXPERIMENT 3

Experiment 3 provided a practical demonstration of a process for assessing unified morphosyntactic systems in pediatric AAC speakers and using assessment results to guide the selection of developmentally appropriate intervention goals. The tense and agreement systems of two pediatric AAC speakers were evaluated in a comprehensive assessment process. Assessment results were used to identify morphosyntactic treatment goals for each participant and guide the development of targeted intervention strategies. Pretest language samples and elicitation probes revealed different patterns of pretest tense marker use in each participant. Participant A used 8 tense markers, distributed across four morpheme categories. Participant B used most tense markers, but did not produce auxiliary *does* or the corresponding subject pronouns *he* and *she*. Stimulability testing was used to successfully identify a tense marker that each participant was stimulable for but did not produce in their pretest assessments (copula *was* for Participant A and auxiliary *does* for Participant B). A short course of intervention focused on increasing production of these target tense markers was provided for each participant. Production of target tense markers increased across sessions in bot treatment and probe sessions. Both participants spontaneously produced their target tense markers in multiple contexts during a post-test language sample.

Treatment sessions provided robust opportunities to practice target tense marker production in diverse contexts, positive reinforcement, and exposure to aided language stimulation modeling correct production of target tense markers on the participants' SGDs. The treatment sessions also provided comparable opportunities to practice production of copula *is*, a contrasting, non-target tense marker that the participants used in their pretest language samples. These experiences with copula *is* were incorporated into treatment sessions to enhance the conceptual saliency of the target tense markers (Fey et al., 2003). Target tense marker productions in multiple contexts and productions of copula *is* in a more diverse range of contexts were found in both participants' posttest language samples. The target tense marker productions are evidence that the participants used a new tense marker in spontaneous conversation after a short course of focused intervention. The increased productivity of copula *is* provides evidence that the participants expanded their spontaneous use of a familiar tense marker that was practiced during treatment activities to create contrasts with a new target tense marker. For Participant A, copula *is* was in the same morpheme category as the target tense marker and shared positional processing features. Increased productivity of copula *is* in Participant A may relate to over-practice of forms in the COPULA BE category. For Participant B, copula *is* was in a different morpheme category than the target tense marker but shared functional processing features. Increased productivity of copula *is* in Participant B's data cannot be explained by over-practice of one morpheme category. If increased productivity of both target and contrasting morphological forms are replicable treatment effects, clinicians may be able to strategically design treatment activities to support increased production of two contrasting forms. Replication studies are needed to verify this possibility.

Evidence of cross-morpheme generalization was found in the posttest language samples of both participants. In one case, treatment effects generalized to posttest production of a new tense marker: Participant A produced the *-3s* form of one verb in his posttest sample and did not produce any *-3s* verbs in his pretest assessment. In other cases of cross-morpheme generalization, a participant used a familiar tense marker in a more diverse range of contexts in his posttest language sample. Such cases were found for Participant A in both past tense markers and third person present tense singular markers. In the past tense, Participant A used *–ed* in a more diverse range of contexts in his posttest sample. Among third person present tense singular markers, Participant A used auxiliary *is* in a more diverse range of contexts in his posttest sample. Participant B's production of familiar present tense markers increased in his posttest language sample. Most notably, use of *-3s* verbs in a more diverse range of contexts suggests generalization across third person present tense singular forms. These findings suggest that the OI theory's strong null hypothesis stating that cross-morpheme generalization is not possible should be rejected. These findings also indicate that the weaker null hypothesis predicting unbounded cross-morpheme generalization across all tense markers should be rejected. This conclusion is stated tentatively based on pretest and posttest assessments in two children.

GML theory predicted that cross-morpheme generalization would occur between tense markers with similar features when intervention supports learning of grammatical features. Some patterns in the data suggest that treatment effects generalized across tense markers with similar features in both participants. For Participant B, both the target tense marker (auxiliary *does*) and the contrasting non-target tense marker used in treatment sessions (copula *is*) encoded features for third person present tense singular forms. Although Participant B used a majority of the tense markers in his pretest language sample, increased productivity of familiar tense markers was only observed for present tense markers in his posttest language sample. Although Participant B's productivity increased for tense markers with different positional processing features, his cross-morpheme generalization was limited to tense markers that shared the [-tense] functional processing feature. The most robust example of increased productivity in Participant B's posttest sample was for *-3s*, which shared functional processing features with the auxiliary *does* and copula *is*.

For Participant A, the target tense marker (copula *was*) was a past tense form and the contrasting non-target tense marker (copula *is*) was a third person present tense singular form. Evidence of cross-morpheme generalization to both past tense and third person present tense singular forms was found in Participant A's posttest language sample. If cross-morpheme generalization occurs between tense markers with shared grammatical features, this would suggest that treatment focused on increased production of copula *was* generalized to increased productivity of other past tense forms (*-ed*) while practice producing copula *is* generalized to production of a new third person present tense singular form (*-3s*) and increased productivity of a familiar third person present tense singular forms (auxiliary *is*).

The non-target tense marker was a familiar tense marker selected to provide contrasts and increase conceptual saliency of the target tense marker (Fey et al., 2003). The cross-morpheme generalization observed in Participant A's language samples suggests that use of a familiar non-target tense marker to create contrasts can generalize to improved productivity of other tense markers that share features with the familiar non-target tense marker. If this is the case, then it may be possible to strategically select both target and non-target tense markers to encourage optimal cross-morpheme generalization.

Morpheme stimulability tests did not unambiguously predict patterns of cross-morpheme generalization, but there is preliminary evidence to suggest that treatment effects are more likely to generalize to stimulable tense markers than non-stimulable tense markers. Evidence of cross-morpheme generalization in Participant A's posttest language sample was found for 3/8 stimulable tense markers and 0/5 non-stimulable tense markers. Evidence of cross-morpheme generalization in Participant B's posttest language sample was found for 1/10 stimulable tense markers and 2/3 non-stimulable tense markers. Both of these non-stimulable tense markers were tense markers that Participant B used in multiple contexts during his pretest language sample. Evidence of cross morpheme generalization also was found in Participant B's elicitation probe for 2 stimulable tense markers (copula and auxiliary *am*). If replication studies find similar patterns, this would provide a stronger indication that morpheme stimulability is a prognostic indicator of cross-morpheme generalization, but that treatment effects are not guaranteed to generalize to stimulable morphemes.

Experiment 3 used a comprehensive pretest assessment of tense marker production and stimulability to identify a tense marker to target in each participant's intervention and a comprehensive posttest assessment to test for cross-morpheme generalization. This comprehensive assessment process systematically identified intervention goals that could be addressed during a short, intensive intervention period. Progress was rapid enough that the intervention phase was concluded before the anticipated 10 session intervention phase was complete. The pretest/posttest design accommodated assessment of generalization in a complete morphosyntactic system. Prior treatment studies focusing on improving production of grammatical morphemes in pediatric AAC speakers have used more robust single-subject designs to demonstrate replication of treatment effects within and between participants (Binger et al., 2011; Kent-Walsh et al., 2015). Experiment 3 focused on investigating a rigorous assessment protocol to guide selection of treatment goals. In contrast, treatment studies have focused on characterizing the effects of rigorous intervention programs. Future studies are needed to investigate the combined effects of rigorous assessment and intervention programs.

5.5 CLINICAL IMPLICATIONS

In Experiment 1, the growth of morpheme category productivity between 30 and 54 months varied across morpheme categories with different positional processing features as predicted by GML theory (Rispoli & Hadley, 2011; Rispoli et al., 2012). This is consistent with the hypothesis that children gradually learn grammatical features of morphosyntactic systems from input using probabilistic learning. A divergence between *-3s* productivity and *–ed* productivity suggests that children may also learn functional processing features. If morphosyntactic systems are gradually learned, then intervention that builds knowledge of grammatical features may encourage development of morphosyntactic systems.

In Experiment 2, the probability of children being stimulable for a given tense marker increased with age for most tense markers. This age-related change in tense marker stimulability was found using two-item or five-item stimulability probes in in either the spoken or graphic symbol modality. Two-item probes in either modality should be sufficient for assessing stimulability of any given tense marker. For some tense markers, children were more likely to be stimulable in the spoken modality than the graphic symbol modality. More research is needed to improve the design of some stimulability tasks.

Stimulability tasks incorporated explicit models of graphic symbol use to introduce symbols and model correct production of target forms in the graphic symbol modality. Aided language stimulation has been recognized as a necessary component of aided AAC intervention based on typical development (Mirenda, 2008). Explicit modeling of targets in the graphic symbol modality may also be a necessary component of morphosyntactic assessment with AAC speakers. These models must be provided in a way that does not reduce the assessment task to an elicited imitation task. The stimulability tasks used in this dissertation overcame the inherent limitations of an elicited imitation task by using multiple exemplars of each target form.

In general, morpheme category stimulability grew at similar rates across morpheme categories. This is consistent with the hypothesis that the tense and agreement system grows maturationally, with highly similar developmental trajectories across categories. Growth of morpheme category stimulability was more consistent with predictions of OI theory while growth of morpheme category productivity was more consistent with predictions of GML theory. Reasons for this discrepancy are not clear. However, the category stimulability scores were obtained in a structured task that is specifically designed to assess all tense markers. Measures of productivity are composite measures obtained from spontaneous language samples that may not account for all tense markers. Structured assessment tasks such as stimulability tasks may be used to supplement spontaneous production tasks and ensure that all forms in a morphosyntactic system are accounted for in a comprehensive assessment.

Experiment 3 demonstrated that it is possible to conduct a comprehensive assessment of a unified morphosyntactic system in pediatric AAC speakers. Measures of productivity and stimulability were used to identify tense markers that each participant used and tense markers that each participant was stimulable for. A similar assessment process can be used in clinical practice to identify developmentally appropriate intervention goals targeting morphosyntactic skills.

Treatment consisted of activities focused on production of a target tense marker that was stimulable and a contrasting non-target tense marker that the participants were already using in their pretest assessment. Increased production of both the target and non-target contrasting tense markers was observed in both participants. Cross-morpheme generalization to other tense markers that shared grammatical features with the target tense marker or non-target contrasting tense marker also was found. These observed patterns of cross-morpheme generalization were most consistent with predictions of GML theory. This suggests that clinicians may be able to streamline treatment programs focused on tense and agreement morpheme development by creating activities that focus on strategically selected pairs of contrasting tense markers and capitalizing on predicted patterns of cross-morpheme generalization.

Two key principles of grammatical intervention with verbally speaking children involve increasing the frequency and saliency of intervention targets (Fey et al., 2003). Frequency of target tense markers was increased though scripted opportunities in storybooks and probe tasks and through manipulation of discourse in treatment sessions. Customized storybooks were designed around participants' individual interests and used in highly interactive treatment activities for several sessions. Saliency of intervention targets was increased by highlighting differences between target tense markers and contrasting, non-target tense markers. Similar strategies have been used in treatment studies with verbally speaking children (e.g., Cleave & Fey, 1997).

Experiment 3 makes a two novel contributions to the literature on morphosyntactic intervention that may be applied with pediatric AAC speakers and verbally speaking children. First, a comprehensive pretest assessment of a morphosyntactic system was used to identify both a target tense marker that was developmentally appropriate and a contrasting, non-target tense marker that the participants already used. Second, a comprehensive posttest assessment of the same system tested for cross-morpheme generalization. Prior studies (e.g., Cleave & Fey, 1997) have shown that incorporating this type of contrastive episode into children's stories can encourage growth of target forms. Experiment 3 extends these findings with evidence that this type of treatment can encourage cross-morpheme generalization to non-target forms that share features with either the target or non-target contrasting form. It may be possible for clinicians to systematically design intervention programs around similar contrasting pairs of target and familiar morphemes and encourage learning of grammatical features that may generalize across morphemes.

Morpheme stimulability tests did not unambiguously predict patterns of cross-morpheme generalization, but there is preliminary evidence to suggest that treatment effects are more likely to generalize to stimulable tense markers than non-stimulable tense markers. Evidence of cross-morpheme generalization in Participant A's posttest language sample was found for 3/8 stimulable tense markers and 0/5 non-stimulable tense markers. Evidence of cross-morpheme generalization in Participant B's posttest language sample was found for 1/10 stimulable tense markers and 2/3 non-stimulable tense markers. Evidence of cross morpheme generalization also was found in Participant B's elicitation probe for 2 stimulable tense markers (copula and auxiliary *am*). If replication studies find similar patterns, this would provide a stronger indication that morpheme stimulability is a prognostic indicator of cross-morpheme generalization, but that treatment effects are not guaranteed to generalize to stimulable morphemes.

Experiment 3 also provides preliminary evidence that morpheme stimulability may be a prognostic indicator of cross-morpheme generalization. Use of new tense markers in a posttest language sample was exclusively observed in stimulable tense markers. Increased productivity of familiar tense markers was observed more often in stimulable tense markers than non-stimulable tense markers. This preliminary evidence should be interpreted with caution because it is drawn from a study using a pretest posttest design with a small number of participants. More robust research studies are necessary to clarify the potential prognostic nature of morpheme stimulability tests.

5.6 CONCLUSIONS

The analysis of morpheme category productivity in Experiment 1 and the analysis of cross-morpheme generalization in Experiment 3 are both consistent with predictions of GML theory. Experiment 1 found that growth of morpheme category productivity between 30 and 54 months patterned with positional processing features, consistent with the hypothesis that children learn the grammatical features of morphosyntactic systems over time through a probabilistic learning process. Experiment 1 also found divergence between morpheme categories differentiated by the functional processing feature of tense, suggesting that children may also learn functional processing features. This evidence is reinforced by the finding that within-participant patterns of cross-morpheme generalization in Experiment 3 were bounded by functional processing features. In Participant B, treatment activities focused exclusively on present tense forms and cross-morpheme generalization was limited to present tense markers. In Participant A, treatment activities focused on contrasting past and present tense forms and cross-morpheme generalization occurred in both past and present tense markers. Evidence of cross-morpheme generalization included production of new tense markers in a posttest language sample and increased productivity of known tense markers in a posttest language sample relative to a pretest language sample.

These findings indicate that grammatical features of morphosyntactic systems are acquired through input-driven learning. Under this input-driven learning system, well-designed, individualized intervention programs can contribute to generalized learning of grammatical features and support morphosyntactic development. Exposure to contrasting morphological forms in a diverse range of contexts (e.g., many different subject-tense marker combinations) can increase perceptual saliency of grammatical features of interest and may encourage generalized learning of these features. Whenever possible, this focused input should be provided in the modality the child uses for expressive language production. For pediatric AAC speakers, this means using aided language stimulation to produce each contrasting form on the child's SGD.

Although other results were most consistent with predictions of GML theory, the models of morpheme category stimulability across categories in Experiment 2 were most consistent with predictions of OI theory. In general, morpheme category stimulability scores grew at similar rates across morpheme categories. Reasons for this discrepancy are not clear. Additional research is needed to reconcile the differences between these models and other results that provide robust evidence for gradual learning of grammatical features.

In a cross-sectional sample of typically developing children, tense marker stimulability increased with age at similar rates across communication modalities for a majority of individual tense markers. Probe items for other tense markers may be improved in future studies to achieve similar patterns of performance.

Parallel assessment of tense marker use, productivity, and stimulability is a viable means of assessing development of the tense and agreement system in verbally speaking children and pediatric AAC speakers. This comprehensive assessment process was used to identify individualized intervention goals for pediatric AAC speakers and guide the development of highly interactive treatment activities based on a child's individual topics of interest. Treatment effects of intervention targeting increased production of a stimulable tense marker generalized to stimulable tense markers more than non-stimulable tense markers. This provides initial evidence to suggest that morpheme stimulability tests can be prognostic.

5.7 NEXT STEPS

In the current dissertation, a series of complementary experiments investigated assessment of tense and agreement morphology skills in both typically developing children and pediatric AAC speakers. This process expanded tense and agreement morphology assessment to a new population of pediatric AAC speakers and verified that a series of novel tense marker stimulability tasks is sensitive to developmental change in typically developing children. Future studies will expand on these novel contributions in two steps. First, the assessment and intervention protocol used in Experiment 3 will be replicated in a larger and more diverse population of pediatric AAC speakers to continue testing hypotheses about patterns of cross-morpheme generalization and establish external validity of the assessment process. Then, similar series of complementary experiments will be conducted to develop strategies for assessing other unified morphosyntactic systems, beginning with the determiner system.

Replication studies with a larger and more diverse population of pediatric AAC speakers are needed to continue testing hypotheses about patterns of cross-morpheme generalization and establish external validity of the assessment process. Experiment 3 tested competing hypotheses about patterns of cross-morpheme generalization in a small sample of two pediatric AAC speakers with cerebral palsy who used the same language application program on their personal SGDs. This small sample produced compelling initial evidence of treatment effects generalizing across tense markers with shared grammatical features and initial evidence that treatment effects are more likely to generalize to stimulable tense markers than non-stimulable tense markers. Replication studies are needed to determine if and when clinicians should expect to find similar patterns of cross-morpheme generalization with pediatric AAC speakers on their case-loads. Researchers using single subject experimental designs to study intervention programs use systematic replication to establish the external validity of intervention programs (Gast, 2010). Evidence that treatment effects can be replicated in different participants provides crucial information about the range of patients who can potentially benefit from a given intervention program. Experiment 3 used a pretest-posttest design to study a comprehensive assessment process and used results from individual participants to test hypotheses about patterns of cross-morpheme generalization. Systematic replication should be used to establish external validity of this assessment process in the same way that systematic replication is used to establish external validity of intervention programs.

In these replication studies, it will be crucial to include participants with a more diverse range of clinical profiles. Experiment 3 focused on assessing pediatric AAC speakers with cerebral palsy because language is often a relative strength in children with cerebral palsy. Rapid improvement of target tense marker production was observed in these children during a short course of treatment. Replication studies can determine whether or not this comprehensive process can lead to similar treatment effects and patterns of cross-morpheme generalization in children who have known difficulty with morphology and syntax. For example, many children with Down syndrome use AAC systems. Verbally speaking children with Down syndrome are known to have relative weaknesses in morphosyntax, with selective weaknesses in expressive morphosyntax relative to both lexical skills and receptive morphosyntax (Chapman, 2006; Chapman & Hesketh, 2000; Chapman, Schwartz, & Kay-Raining Bird, 1991; Chapman, Seung, Schwartz, & Kay-Raining Bird, 1998). If treatment of stimulable morphemes leads to robust treatment and generalization effects in pediatric AAC speakers with Down syndrome, then this assessment process can inform intervention targeting morphosyntactic skills in a population of children who are at risk for life-long difficulties in the area of morphosyntax.

When these replication studies are completed, a similar series of complementary experiments will be needed to investigate assessment of the determiner system, another unified morphosyntactic system. Both OI and GML theories hold that the English tense and agreement morphemes system emerge simultaneously as a unified morphosyntactic system, or a group of functionally related morphemes that share an underlying grammatical system. This morphosyntactic system was initially described by Radford (1990), who argued that adult grammar includes several morphosyntactic systems, which are absent from children's earliest grammars. Radford (1990) argued that children's earliest grammars only include lexical categories, such as nouns, verbs, and adjectives, and that grammatical morphemes appear when morphosyntactic systems emerge developmentally. Morphemes in the tense and agreement system encode crucial grammatical information about verbs and take verb phrases as complements. The determiner system is another crucial morphosyntactic system, which includes morphemes that encode crucial grammatical information about nouns and take noun phrases as complements.

Clinical AAC intervention often focuses on building a lexicon in the graphic symbol modality and establishing a base of words that contribute to the lexical categories described by Radford (1990). Some highly structured AAC intervention programs explicitly emphasize building a lexicon at the single word level during early stages of intervention before systematically introducing multi-word constructions (e.g., Frost & Bondy, 2002). Intervention focused on building a lexicon, communicating at the single word level, or producing early multi-word constructions is frequently implemented without assessment data to verify that intervention is targeting developmentally appropriate skills. Clinicians may invest more contact time building language skills above the single-word level if assessment data indicate that a pediatric AAC speaker is stimulable for more advanced morphosyntactic skills in the graphic symbol modality. Additional tasks for assessing stimulability of morphemes in the determiner system will contribute to a series of tasks for assessing stimulability of crucial morphosyntactic skills necessary for expressive communication above the developmental level commonly referred to as Brown's Stage I.

Experiments with typically developing children can be used to determine whether or not growth of the determiner system reflects gradual learning of grammatical features. Experiments with typically developing children can also be used to develop and refine a series of determiner system stimulability tasks that are sensitive to developmental changes in both spoken and graphic symbol modalities. This will require identifying the elements of the determiner system, identifying positional and functional processing features that differentiate those elements, and developing probes that can be used to assess stimulability in a younger participant population. A comprehensive review of responses to stimulability probes in the current data set, including error analysis can be used to identify factors that influence the validity of morpheme stimulability probes and inform development of an optimal set of tasks for earlier developing forms.

Experiments assessing a variety of pediatric AAC speakers can expand the evidence base for understanding expressive morphology assessment in pediatric AAC speakers. Experiment 3 of this project provided initial evidence that morpheme productivity and stimulability can be assessed in pediatric AAC speakers and that the combined results of these tasks can be used to drive the development of intervention goals. Experiment 3 established this groundwork with pediatric AAC speakers who had already acquired determiners and were in the process of acquiring later developing tense and agreement morphemes. Future experiments can use similar strategies for assessing children who are acquiring determiners and help move pediatric AAC speakers beyond intervention limited to teaching items in early-developing lexical categories.

APPENDIX A

OPERATIONAL DEFINITIONS OF DEPENDENT VARIABLES

	Spontaneous	Stimulability	Potential
Tense Marker Use/Stimulability	Whether or not a tense marker is used correctly at least once in a language sample. Contractions to pronominal subjects not counted. Specifically negative forms *don't, ain't* not counted.	Whether or not a tense marker is used correctly in at least one probe item on a stimulability task.	Whether or not a tense marker can be selectively used on SGD in at least 1 spontaneous novel utterance w/o spell mode. Contractions not counted. For *–ed*, *-3s*: use same lexical verb in 2 sentences, 1 w/un-inflected form, 1 w/inflected For AUX BE forms, progressive main verb
Tense Marker Total (Rg: 0-15)	N. tense markers used in sample at least once	N. stimulable tense markers	N. tense markers that can be selectively used
Productivity/ Stimulability Score	Sum of 5 category productivity scores	Sum of 5 category stimulability scores	Sum of 5 category productivity scores
Criteria for sufficiently diff. use *-ed* *-3s*	Correct inflection of different lexical verbs in a language sample. Overregularizations counted.	Items elicit selective production of inflected forms of different lexical verbs	Correct use of same lexical verb in 2 sentences, 1 w/ uninflected form, 1 w/ inflected form.
Criteria for sufficiently diff. use COPULA BE AUX. BE AUX. DO	Correct use of diff. subject-tense marker combinations in a language sample. Contractions to pronominal subjects & specifically negative forms not counted.	Correct response on items probing stimulability of tense markers in sufficiently different subject-tense marker combinations	Correct use in sentences w/different subject-tense marker combinations. Contractions not counted. Pronominal subjects only counted for *am*.
Category Productivity/ Stimulability Score (Rg: 0-5)	N. sufficiently diff. uses of morphemes in a category in a language sample. Max. 5/category.	Proportion of correct responses on stimulability probe items to all tense markers in each morpheme category.	N. sufficiently diff. contexts for potential use of morphemes in a category on the SGD w/o spell mode. Max. 5/category.

APPENDIX B

GENERAL PROCEDURES FOR PROBE & PRACTICE ITEMS

General Procedures for Practice Items

Introduction: One practice item is presented in the target modality at the beginning of each new task. In the graphic symbol modality, practice items are presented after the child has been familiarized to the current vocabulary page. Each practice item will provide input demonstrating correct use of a tense marker that will be probed in the corresponding task in the target output modality. The child is always given the opportunity to produce a correct response during the practice item with no penalty for incorrect responses.

- Present item & visual stimuli following script provided.
- 10 second expectant delay. Wait longer if the child is still clearly generating a response.
- If **correct response**, give positive reinforcement by praising the participant and repeating the participant's utterance, "Good job! I like how you said _____."
- If **partial (incomplete) response**, recast with a complete correct response by repeating the cue and demonstrating correct response in the target modality. Elicit repetition. Reinforce (i.e. bubbles)
- If **incorrect**, **verbal response**, or **no response**, repeat verbal cue and produce correct response in the target modality. Elicit repetition. Give positive reinforcement.

General Procedures for Probe Items

Introduction: The probe items follow a similar format to the practice items.

- Present item and visual stimuli following script provided.
- 10 second expectant delay. Wait longer if the child is still clearly generating a response.
- If **correct response**, Mark correct, give positive reinforcement by praising the participant and repeating the participant's utterance, "Good job! I like how you said _____."
- If **no response, repeat cue once** & wait again. If no response after repetition, Mark incorrect.
- If **verbal response** *(when target modality is graphic symbols in Exp. 2 only),* add verbal cue "show me how to say that with the device." Wait again.
- If **partial (incomplete) response**, recast with a complete correct response by repeating the cue and demonstrating correct response in the target modality. Mark incorrect. Record the partial response on the score sheet. Allow child to repeat if self-initiated. Reinforce correct repetition.
- If **incorrect response**, mark incorrect. Repeat cue and demonstrate correct response in the target modality. Allow child to repeat if self-initiated. Reinforce correct repetition.

APPENDIX C

Productivity and Stimulability Data Analysis Form

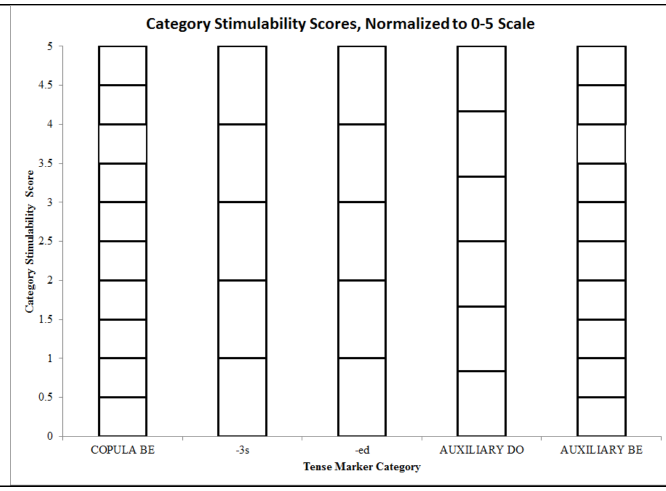
Participant ID _________ Sample Age _________ Lang. Sample Date _________

Stim. Task Age:______ Stim. Task Date _________

Morpheme Category Summary Charts

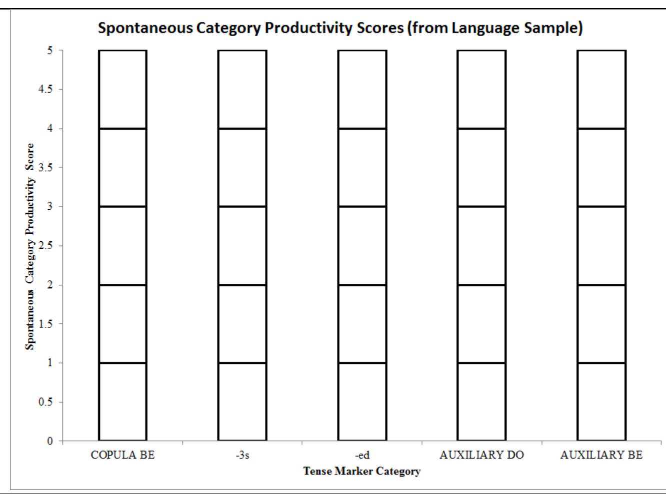
To illustrate category stimulability scores on a normalized scale, shade in the number of boxes corresponding to the number of correct responses on items corresponding to each category in the stimulability tasks.

Categories in spoken modality:	Categories in graphic symbol modality:



To illustrate spontaneous category productivity score, shade in the number of boxes corresponding to the spontaneous category productivity score for each category, as observed in the spontaneous language sample.

Spontaneous tense marker total from language sample	Spontaneous productivity score from language sample



Tense Marker Summary Table

	Cop is	Cop are	Cop am	Cop was	Cop were	-3s	-ed	Aux does	Aux do	Aux did	Aux is	Aux are	Aux am	Aux was	Aux were	Total
Used (lng sample)																
Stimulable																

Mark 1 for a tense marker that is used/stimulable, mark 0 for a tense marker that is not.

Analysis of Tense-Marker Stimulability

Category	**Tense Marker**	**Modality Speech**	**G Sym**	**Items**	**# Correct**	**Stimulable? (# Correct > 0)**
COPULA BE	***is***			1, 2		
COPULA BE	***are***			3, 4		
COPULA BE	***am***			5, 6		
COPULA BE	***was***			7, 8		
COPULA BE	***were***			9. 10		
-3s	***-3s***			11-15		
-ed	***-ed***			16-20		
AUXILIARY DO	***does***			21, 22		
AUXILIARY DO	***do***			23, 24		
AUXILIARY DO	***did***			25, 26		
AUXILIARY BE	***is***			27, 28		
AUXILIARY BE	***are***			29, 30		
AUXILIARY BE	***am***			31, 32		
AUXILIARY BE	***was***			33, 34		
AUXILIARY BE	***were***			35, 36		

Analysis of Category Stimulability

Category	**Modality Speech**	**G Sym**	**Items**	**#Correct/#Possible**
COPULA BE			1-10	/10
-3s			11-15	/5
-ed			16-20	/5
AUXILIARY DO			21-26	/6
AUXILIARY BE			27-36	/10

Stimulability Task Responses

Notes: In spoken modality, non-target contexts are acceptable as long as the subject is not pronominal (e.g. a child says *the eagle is* instead of *the bird is*). If a pronominal subject is used, prompt *start with the*.

In graphic symbol modality, correct contexts are needed. Incorrect contexts would be correct responses on other items.

	Response		**Verb Stem & Tense Marker or Subj. & Tense Marker**	
	Target	**Observed (if different)**	**Correct**	**Incorrect**
P1	bird **is** [in the basket]			
1	bird **is** [on the chair]		☐	☐
2	moose **is** [next to the plant]		☐	☐
3	zebras **are** [in the basket]		☐	☐
4	bears **are** [on the chairs]		☐	☐
5	I **am** [under the box]		☐	☐
6	I **am** [next to the box]		☐	☐
P2	[The] boat **was** white			
7	[The] wagon **was** red		☐	☐
8	[The] house **was** yellow		☐	☐
9	[The] cars **were** blue		☐	☐
10	[The] trucks **were** white		☐	☐
P3	[The boy] run**s**			
11	[The man] paint**s**		☐	☐
12	[The girl] read**s**		☐	☐
13	[The girl] climb**s**		☐	☐
14	[The boy] sing**s**		☐	☐
15	[The woman] dance**s**		☐	☐
P4	[She] climb**ed**			
16	[He] paint**ed**		☐	☐
17	[He] jump**ed**		☐	☐
18	[It] snow**ed**		☐	☐
19	[He] splash**ed**		☐	☐
20	[She] rak**ed**		☐	☐
P5	**Do** you want pizza			
21	**Does** the bird want pizza		☐	☐
22	**Does** the moose want pizza		☐	☐
23	**Do** the zebras want chicken		☐	☐
24	**Do** the bears want chicken		☐	☐
25	**Did** the bird eat chicken		☐	☐
26	**Did** the zebras eat pizza		☐	☐
P6	moose **is** [reading a book]			
27	moose **is** [sleeping]		☐	☐
28	bird **is** [flying in the sky]		☐	☐

	Response		**Verb Stem & Tense Marker or Subj. & Tense Marker**	
	Target	**Observed (if different)**	**Correct**	**Incorrect**
29	zebras **are** [reading a book]		☐	☐
30	bears **are** [drinking water]		☐	☐
31	I **am** [sitting]		☐	☐
32	I **am** [sitting]		☐	☐
P7	The fox **was** dancing			
33	The moose **was** drawing		☐	☐
34	The bird **was** sleeping		☐	☐
35	The bears **were** dancing		☐	☐
36	The zebras **were** singing		☐	☐

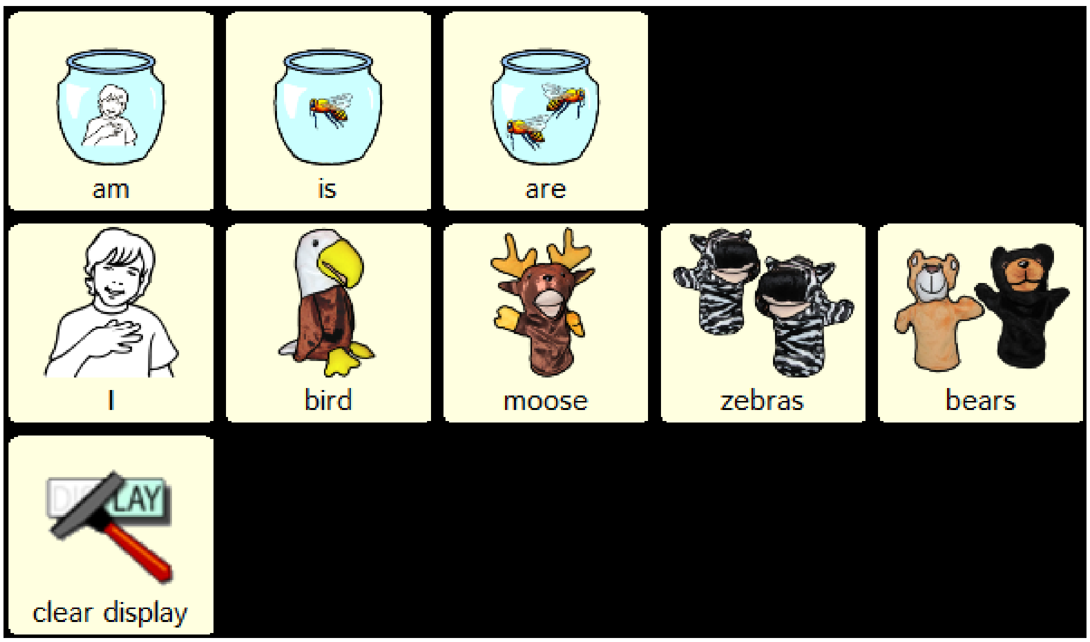
*P = Practice item

APPENDIX D

VOCAB PAGES FOR STIMULABILITY TASKS IN GRAPHIC SYMBOL MODALITY

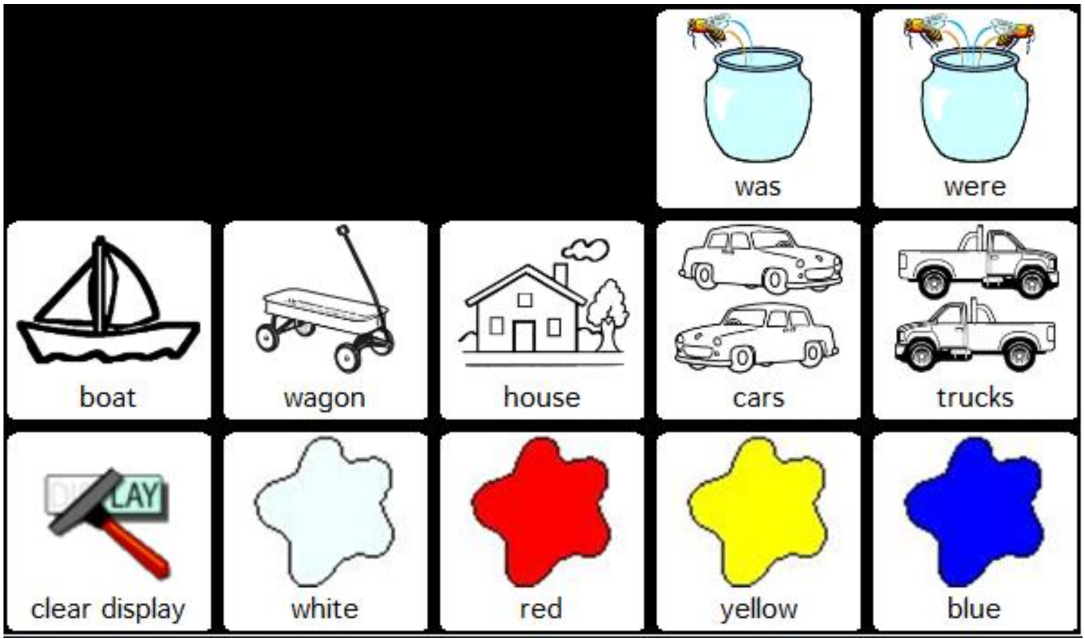
Vocabulary Page #1: Present tense BE (text glosses removed for experiments)

- Use for copula *am*, *is*, *are* task
- Use for auxiliary *am*, *is*, *are* task



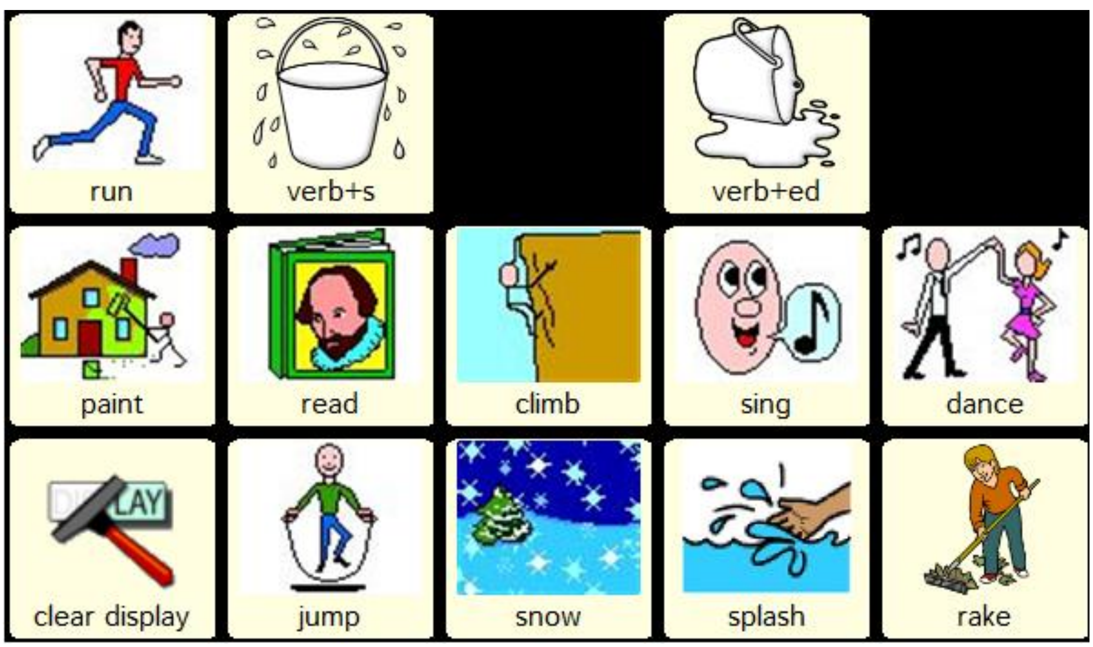
Vocabulary Page #2: Past tense COPULA BE (text glosses removed for experiments)

- Use for copula *was*, *were* task



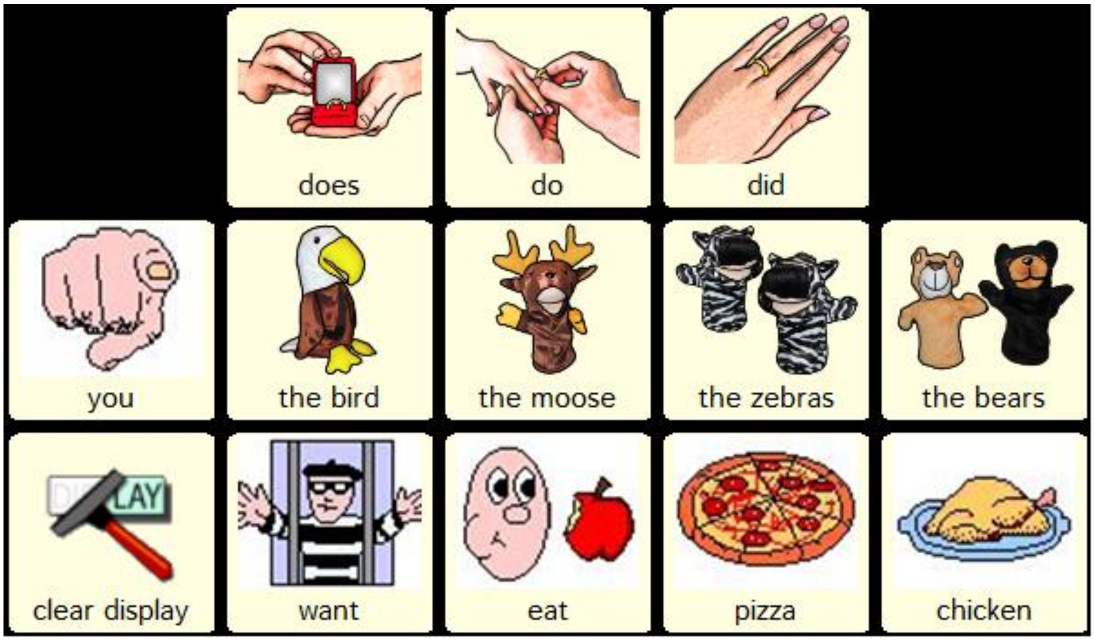
Vocabulary Page #3: *-3s* and *–ed* (text glosses removed for experiments)

- Use for -*3s* task
- Use for –*ed* task



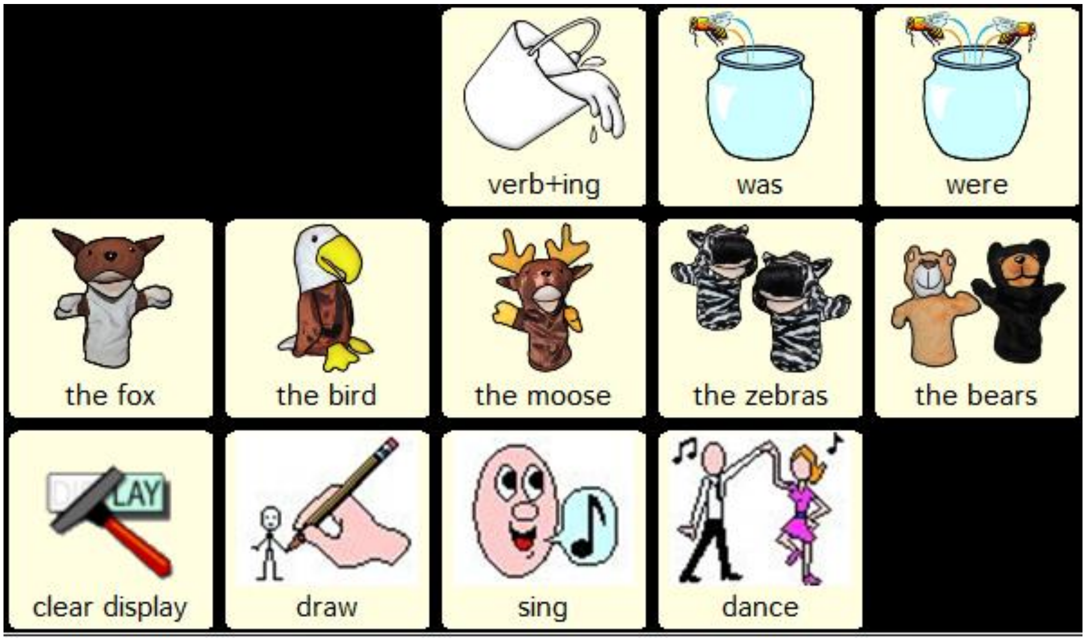
Vocabulary Page #4: AUXILIARY DO (text glosses removed for experiments)

- Use for auxiliary *do*, *does*, *did* task (AUXILIARY DO task)



Auxiliary Page #5: Past tense AUXILIARY BE (text glosses removed for experiments)

- Use for auxiliary *was*, *were* task



APPENDIX E

SPONTANEOUS NOVEL UTTERANCES WITH TENSE MARKERS GENERATED BY PEDIATRIC AAC SPEAKERS IN EXPERIMENT 3 LANGUAGE SAMPLES

Table E.1. Spontaneous novel utterances containing tense markers generated in Participant A's pretest language sample

Tense Marker	Utterance[a]	Tense Marker Context
copula *is*	is [Name] there	[Name] is
copula *is*	what is your last name	your last name is
copula *are*	we are going to hug you	we are
-ed	arrived	arrived
-ed	arrived	arrived
auxiliary *did*	I didn't understand	did I
auxiliary *did*	you didn't leave your number	did you
auxiliary *are*	then are you going	you are
auxiliary *are*	are we going go outside	we are
auxiliary *was*	I was asking you on the phone	I was

Note. [a]Only utterances containing a correct production of a tense marker in context are counted. Productions of tense markers in personal pronoun + tense marker contractions (*it's*, *he's*, *I'm*, etc.) are not counted.

Table E.2. Spontaneous novel utterances containing tense markers generated in Participant A's posttest language sample

Tense Marker	Utterance[a]	Tense Marker Context
copula *is*	who is it	it is
copula *is*	its behind you	it is
copula *is*	[Name] is here	[Name] is
copula *is*	she is home	she is
copula *is*	where is is your hotel	your hotel is
copula *was*	it was huge	it was
copula *was*	it was cloudy	it was
copula *was*	it was good	it was
copula *was*	it was excited	it was
copula *was*	car was dirty	car was
copula *was*	my wheelchair was damaged	my wheelchair was
copula *was*	she was home	she was
copula *was*	that was a good idea	that was
copula *was*	the car was dirty	the car was
copula *was*	the cat was in the flolor	the cat was
copula *was*	the dog was in the car	the dog was
copula *was*	the taxi was red	the taxi was
-3s	he sounds funny	sounds
-ed	tom called	called
-ed	we missed the airplane	missed
-ed	talked to you yesterday	talked
-ed	I wanted to see her	wanted
auxiliary *is*	he is calling	he is
auxiliary *is*	he is calling	he is
auxiliary *is*	he is paying the people	he is
auxiliary *is*	the aiprport is calling us all	the aiprort (airport) is
auxiliary *is*	the airp. Is calling [PhonNum]	the airp. (airport) is

Table E.2. Continued

Tense Marker	Utterance[a]	Tense Marker Context
auxiliary *is*	phone is ranging	phone is
auxiliary *is*	the airplane is coming	the airplane is
auxiliary *are*	we are calling to talk	we are
auxiliary *are*	we are going on it	we are

Note. [a]Only utterances containing a correct production of a tense marker in context are counted. Productions of tense markers in personal pronoun + tense marker contractions (*it's*, *he's*, *I'm*, etc.) are not counted.

Table E.3. Spontaneous novel utterances containing tense markers generated in Participant B's pre-test language sample

Tense Marker	Utterance[a]	Tense Marker Context
copula *is*	is it on	it is
copula *is*	is it on	it is
copula *is*	is it open	it is
copula *is*	where is it	it is
copula *is*	because it's fun	it is
copula *is*	it's seventeen dollars	it is
copula *is*	it's your turn mom	it is
copula *is*	when is New_Year's_Day	New_Year's_Day is
copula *is*	please ask when is our turn please	our turn is
copula *is*	boys first since there is one	there is
copula *are*	if you are OK with it Mr_Tom	you are
copula *are*	where are you living when you are in [CITY]	you are
copula *are*	the hand cookies are so good	the hand cookies are
copula *are*	when we are done can I stop it please	we are
copula *was*	when I was like three	I was
copula *was*	because it was funny	it was
-*ed*	I haven't played it forever so I forgot	played
auxiliary *do*	how do you win	do you
auxiliary *do*	do you need a receipt	do you
auxiliary *do*	tomorrow how come do I have to come	do I
auxiliary *did*	yesterday what did you go	did you
auxiliary *is*	who is helping it	who is
auxiliary *are*	where are we sleeping	we are
auxiliary *are*	tomorrow what time are we working	we are
auxiliary *are*	when are we flying to this year	we are
auxiliary *are*	are we leaving to get there in the morning	we are
auxiliary *are*	how will we know what day are we cooking	we are

Table E.3. Continued

Tense Marker	Utterance[a]	Tense Marker Context
auxiliary *are*	we are getting money for the [ORGANIZATION] in Utah	we are
auxiliary *are*	we are doing a thing for money for the [ORGANIZATION]	we are
auxiliary *are*	how come are you writing if we are taking a movie	we are
auxiliary *are*	when are you flying home	you are
auxiliary *are*	when are you grading it	you are
auxiliary *are*	how are you seeing the movie	you are
auxiliary *are*	when are you coming to my house	you are
auxiliary *are*	where are you living when you are in [CITY]	you are
auxiliary *are*	we'll see when are you coming to today	you are
auxiliary *are*	how come are you writing if we are taking a movie	you are

Note. [a]Only utterances containing a correct production of a tense marker in context are counted. Productions of tense markers in personal pronoun + tense marker contractions (*it's*, *he's*, *I'm*, etc.) are not counted.

Table E.4. Spontaneous novel utterances containing tense markers generated in Participant B's post-test language sample

Tense Marker	Utterance[a]	Tense Marker Context
copula *is*	Is you person a boy	your person is
copula *is*	is your person a boy	your person is
copula *is*	Is your person a boy	your person is
copula *is*	Is your person a a boy	your person is
copula *is*	is their a boy orr a girl	there is
copula *is*	this is your key	this is
copula *is*	What is on your hand	what is
copula *is*	tomorrow what time is your check out	your check out is
copula *is*	is your person 's hair white	your person's hair is
copula *are*	you are in 26	you are
copula *are*	What number are you	you are
copula *was*	how was the room	the room was
copula *were*	yesterday were you full	you were
-3s	who ever gets the most wins	gets
-3s	What are you writing if he yes gets it right or wrong	gets
-3s	can we see how she coes please	closes (spelling error)
-3s	who ever gets the most wins	wins
-ed	Lets play how you check ed in	checked
auxiliary *does*	Does he have glasses	does he
auxiliary *does*	what does he look like	does he
auxiliary *does*	Does he has brown eye	does he
auxiliary *does*	Does he have a black	does he
auxiliary *does*	What color hair does he have	does he
auxiliary *does*	Does your person have glasses	does your person
auxiliary *does*	does your person have a hat	does your person
auxiliary *does*	what hair does your person have	does your person

Table E.4. Continued

Tense Marker	Utterance[a]	Tense Marker Context
auxiliary *does*	Does your person start with a b	does your person
auxiliary *do*	Do you have geoe	do you
auxiliary *do*	Do you have sosn	do you
auxiliary *do*	Do you have a kitchen	do you
auxiliary *do*	How do you open it	do you
auxiliary *do*	Do you have a little refrigerator	do you
auxiliary *do*	Do you have your thing so that you'll remember	do you
auxiliary *do*	tomorrow what time do you think that Mr Tom can come	do you
auxiliary *do*	do you want me to walk you to to your room	do you
auxiliary *do*	how do I turn it on	do I
auxiliary *do*	how many an more things do we have	do we
auxiliary *is*	If mom is not coming you'll call her	mom is
auxiliary *are*	What are you writing	you are
auxiliary *are*	tomorrow what time are you coming	you are
auxiliary *are*	What are you writing if he yes gets it right or wrong	you are
auxiliary *are*	we try and get them when they are mov ing that	they are
auxiliary *are*	I thought that tomorrow we are working	we are

Note. [a]Only utterances containing a correct production of a tense marker in context are counted. Productions of tense markers in personal pronoun + tense marker contractions (*it's*, *he's*, *I'm*, etc.) are not counted.

APPENDIX F

ASSESSMENT OF SGD POTENTIAL TO SUPPORT TENSE MARKER USE

Table F.1. Sentences demonstrating potential to support selective tense marker use in Experiment 3 participants' language application programs

Tense Marker	Sentence W/TM or Pronominal Subj.	Sentnce W/O TM or W/Nominal Subj.
copula *is*	He is happy.	The man is happy.
copula *am*	I am hungry.	I eat cereal for breakfast.
copula *are*	We are in the car.	The boys are in the car.
copula *was*	It was at the park.	The dog was at the park.
copula *were*	They were with their family.	The children were with their family.
-*3s*	The fox jumps over the dog.	The foxes jump over the dog.
-ed	The foxes jumped over the dog.	The foxes jump over the dog.
auxiliary does	Does he want to go?	Does your friend want to go?
auxiliary do	Do you like pizza?	Do the children like pizza?
auxiliary *did*	Did they take it with them?	Did the girls take it with them?
auxiliary *is*	She is working on her homework.	The girl is working on her homework.
auxiliary *am*	I am going for a walk.	I went for a walk.
auxiliary *are*	We are making breakfast.	My parents are making breakfast.
auxiliary *was*	He was singing.	The man was singing.
auxiliary *were*	They were dancing.	The people were dancing.

Table F.2. Sentences demonstrating potential category productivity in Experiment 3 participants' language application programs

Category	Sentences With TM or Pron. Subj.[a]	Sentences W/O TM or W/Nom. Subj.
COPULA BE	He is happy.	The man is happy.
	I am hungry.	I eat cereal for breakfast.
	We are in the car.	The boys are in the car.
	It was at the park.	The dog was at the park.
	They were with their family.	The children were with their family.
-3S	The fox jumps over the dog.	The foxes jump over the dog.
	The boy walks home after school.	The children walk home after school.
	She likes him.	They like him.
	The baby loves his mom.	The babies love their mom.
	The woman dances.	They dance.
-ED	The foxes jumped over the dog.	The foxes jump over the dog.
	I wanted cereal for breakfast.	I want cereal for breakfast.
	The cars raced around the circle.	The cars race around the circle.
	The man cleaned the house.	The man and woman clean the house.
	The children helped their mom.	The children help their mom.
AUXILIARY DO	Does he want to go?	Does your friend want to go?
	Does she need any more?	Does the lady need any more?
	Do you like pizza?	Do the children like pizza?
	Did they take it with them?	Did the girls take it with them?
	Did it make you sneeze?	Did the cat make you sneeze?
AUXILIARY BE	She is working on her homework.	The girl is working on her homework.
	I am going for a walk.	I went for a walk.
	We are making breakfast.	My parents are making breakfast.
	He was singing.	The man was singing.
	They were dancing.	The people were dancing.

Note. [a]Up to 5 sentence pairs per category.

BIBLIOGRAPHY

Ambridge, B., Rowland, C. F., Theakston, A. L., & Tomasello, M. (2006). Comparing different accounts of inversion errors in children's non-subject wh-questions: 'What experimental data can tell us?'*. *Journal of Child Language, 33*, 519-557.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59*, 390-412.

Bain, B. A. (1994). A framework for dynamic assessment in phonology: Stimulability revisited. *Clinics in Communication Disorders, 4*, 12-22.

Baker, B. (1982). Minspeak: A semantic compaction system that makes self-expression easier for communicatively disabled individuals. *Byte, 7*, 186-202.

Baker, B. (1986). Using images to generate speech. *Byte, 11*, 160-168.

Barr, D. J. (2008). Analyzing 'visual world' eyetracking data using multilevel logistic regression. *Journal of Memory and Language, 59*, 457-474.

Bedore, L. M., & Leonard, L. B. (1998). Specific language impairment and grammatical morphology: A discriminant function analysis. *Journal of Speech, Language, and Hearing Research, 41*, 1185-1192.

Berninger, V. W., & Gans, B. M. (1986). Language profiles in nonspeaking individuals of normal intelligence with severe cerebral palsy. *Augmentative and Alternative Communication, 2*, 45-50.

Binger, C., Kent-Walsh, J., Berens, J., Del Campo, S., & Rivera, D. (2008). Teaching Latino parents to support the multi-symbol message productions of their children who require AAC. *Augmentative and Alternative Communication, 24*, 323-338.

Binger, C., & Light, J. (2007). The effect of aided AAC modeling on the expression of multi-symbol messages by preschoolers who use AAC. *Augmentative and Alternative Communication, 23*, 30-43.

Binger, C., & Light, J. (2008). The morphology and syntax of individuals who use AAC: Research review and implications for effective practice. *Augmentative and Alternative Communication, 24*, 123-138.

Binger, C., Maguire-Marshall, M., & Kent-Walsh, J. (2011). Using aided AAC models, recasts, and contrastive targets to teach grammatical morphemes to children who use AAC. *Journal of Speech, Language, and Hearing Research, 54*, 160-176.

Blockberger, S., & Johnston, J. R. (2003). Grammatical morphology acquisition by children with complex communication needs. *Augmentative and Alternative Communication*, *19*, 207-221.

Blom, E., & Wijnen, F. (1999). How Dutch children's root infinitives become modal *Proceedings BCULD 24* (pp. 128-139).

Bloom, L., & Lahey, M. (1978). *Language Development and Language Disorders*. New York: John Wiley & Sons.

Bock, J. K. (1982). Toward a cognitive psychology of syntax: Information processing contributions to sentence formulation. *Psychological Review, 89*, 1-47.

Bock, J. K., & Levelt, W. (1994). Language production: Grammatical encoding. In M. Gersenbacher (Ed.), *Handbook of Psycholinguistics* (pp. 945-984). San Diego, CA: Academic Press.

Bondy, A. S., & Frost, L. A. (1994). The picture exchange communication system. *Focus on Autistic Behavior, 9*, 1-19.

Borer, H., & Wexler, K. (1987). The maturation of syntax. In T. Roeper & E. Williams (Eds.), *Parameter setting* (pp. 123-172). Dordrecht: D. Reidel.

Borer, H., & Wexler, K. (1992). Bi-unique relations and the maturation of grammatical principles. *Natural Language & linguistic Theory, 10*, 147-189.

Brown, R. (1973). *A First Language: The Early Stages*. Cambridge, MA: Harvard University Press.

Bruno, J., & Trembath, D. (2006). Use of aided language stimulation to improve syntactic performance during a weeklong intervention program *Augmentative and Alternative Communication, 22*, 300-313.

Bybee, J. L. (2006). From usage to grammar: The mind's response to repetition. *Language, 82*, 711-733.

Carrow-Woolfolk, E. (1985). *Test for Auditory Comprehension of Language* (Revised ed.). Allen, TX: DLM Teaching Resources.

Carrow-Woolfolk, E. (1999). *Test for auditory comprehension of language* (3rd ed.). Austin, TX: Pro-Ed.

Chapman, R. S. (2006). Language learning in Down syndrome: The speech and language profile compared to adolescents with cognitive impairment of unknown origin. *Down Syndrome Research and Practice*, *10*, 61-66.

Chapman, R. S., & Hesketh, L. J. (2000). Behavioral phenotype of individuals with Down syndrome. *Mental Retardation and Developmental Disabilities Research Reviews, 6*, 84-95.

Chapman, R. S., Schwartz, S. E., & Kay-Raining Bird, E. (1991). Language skills of children and adolescents with Down syndrome: I. Comprehension. *Journal of Speech, Language, and Hearing Research, 34*, 1106-1120.

Chapman, R. S., Seung, H. K., Schwartz, S. E., & Kay-Raining Bird, E. (1998). Language skills of children and adolescents with Down syndrome: II. Production deficits. *Journal of Speech, Language, and Hearing Research, 41*, 861-873.

Chomsky, N. (1965). *Aspects of the theory of syntax.* Cambridge, MA: MIT Press.

Chomsky, N. (1995). *The minimalist program*. Cambridge, MA: MIT Press.

Cleave, P. L., & Fey, M. E. (1997). Two approaches to the facilitation of grammar in children with specific language impairment: Rationale and description. *American Journal of Speech Language Pathology, 6*, 22-32.

Cohen, J. (1960). A coefficient of agreement for normal scales. *Educational and Psychological Measurement*, *20*, 37-46.

Crain, S., & Nakayama, M. (1987). Structure dependence in grammar formation. *Language, 63*, 522-543.

Crain, S., & Thornton, R. (2000). *Investigations in universal grammar: A guide to experiments on the acquisition of syntax and semantics*. Cambridge, MA: MIT Press.

Cumley, G. D., & Swanson, S. (1999). Augmentative and alternative communication options for children with developmental apraxia of speech: Three case studies. *Augmentative and Alternative Communication, 15*, 110-125.

Dawson, J., Stout, C., Eyer, J. A., Tattersall, P., Fonkalsrud, J., & Croley, K. (2005). *Structured Photographic Expressive Language Test Preschool Second Edition (SPELT-P 2).* DeKalb, IL: Janelle Publications.

de Villiers, J. G., & de Villiers, P. A. (1973). A cross-sectional study of the acquisition of grammatical morphemes in child speech. *Journal of Psycholinguistic Research, 2*, 267-278.

Dr. Seuss. (1960). *Green Eggs and Ham*. New York: Random House.

Drager, K., Light, J., Carlson, R., D'Silva, K., Larsson, B., Pitkin, L., & Stopper, G. (2004). Learning of dynamic display AAC technologies by typically developing 3-year-olds: Effects of different layouts and menu approaches. *Journal of Speech, Language, and Hearing Research, 47*, 1133-1148.

Drager, K., Light, J., Speltz, J. C., Fallon, K. A., & Jeffries, L. Z. (2003). The performance of typically developing 2 1/2-year-olds on dynamic display AAC technologies with different system layouts and language organizations. *Journal of Speech, Language, and Hearing Research, 46*, 298-312.

Fenson, L., Marchman, V. A., Thal, D. J., Dale, P. S., Reznick, J. S., & Bates, E. (2007). *MacArthur-Bates Communicative Development Inventories: User guide and technical manual* (2nd ed.). Baltimore, MD: Brookes.

Fey, M. E., Long, S. H., & Finestack, L. H. (2003). Ten principles of grammar facilitation for children with specific language impairment. *American Journal of Speech Language Pathology*, *12*, 3-15.

Foster, R., Giddan, J. J., & Stark, J. (1983). *Assessment of children's language comprehension, 1983 revision*. Chicago, IL: Riverside Publishing.

Frost, L. A., & Bondy, A. S. (2002). *The picture exchange communication system training manual* (2nd ed.). Newark, DE: Pyramid Educational Products.

Garrity, A. W., & Oetting, J. B. (2010). Auxiliary BE production by African American English-speaking children with and without specific language impairment. *Journal of Speech, Language, and Hearing Research, 53*, 1307-1320.

Gast, D. L. (ED.) (2010). *Single subject research methodology in behavioral sciences*. New York: Routledge.

Gierut, J. A., Morrisette, M. L., Hughes, M. T., & Rowland, S. (1996). Phonological treatment efficacy and developmental norms. *Language, Speech, and Hearing Services in Schools, 27*, 215-230.

Gladfelter, A., & Leonard, L. B. (2013). Alternative tense and agreement morpheme measures for assessing grammatical deficits during the preschool period. *Journal of Speech and Hearing Research, 56*, 542-552.

Goldman, R., & Fristoe, M. (2015). *Goldman-Fristoe Test of Articulation 3 (GFTA-3)*. San Antonio, TX: Pearson.

Goosens', C. (1989). Aided communication intervention before assessment: A case study of a child with cerebral palsy. *Augmentative and Alternative Communication, 5*, 14-26.

Hadley, P., & Walsh, K. M. (2014). Toy talk: Simple strategies to create richer grammatical input. *Language, Speech, and Hearing Services in Schools, 45*, 159-172.

Hadley, P., & Rice, M. L. (1996). Emergent uses of BE and DO: Evidence from children with specific language impairment. *Language Acquisition, 5*, 209-243.

Hadley, P., Rispoli, M., Fitzgerald, C., & Bahnsen, A. (2011). Predictors of morphosyntactic growth: Contributions of parent input and child sex. *Journal of Speech, Language, and Hearing Research, 54*, 549-566.

Hadley, P., Rispoli, M., Holt, J. K., Fitzgerald, C., & Bahnsen, A. (2014). Growth of finiteness in the third year of life: Replication and predictive validity. *Journal of Speech, Language, and Hearing Research, 57*, 887-900.

Hadley, P., & Short, H. (2005). The onset of tense marking in children at risk for speific language impairment. *Journal of Speech, Language, and Hearing Research, 48*, 1344-1362.

Halle, M., & Marantz, A. (1993). Distributed morphology and the pieces of inflection. In K. Hale & S. J. Keyser (Eds.), *The view from building 20: Essays in linguistics in honor of Sylvian Bromberger* (pp. 111-176). Cambridge, MA: MIT Press.

Harris, A., & Wexler, K. (1996). The optional-infinitive stage in child English: Evidence from negation. In H. Clahsen (Ed.), *Generative perspective on language acquisition* (pp. 1-42). Philadelphia, PA: John Benjamins.

Hart, B., & Risley, T. R. (1995). *Meaningful differences in the everyday experience of young American children*. Baltimore, MD: Paul H. Brookes Publishing Co.

Hill, K. (2004). Augmentative and alternative communication and language: Evidence-based practice and language activity monitoring. *Topics in Language Disorders, 24*, 18-30.

Hill, K. (2010). Advances in augmentative and alternative communication as quality-of-life technology. *Physical Medicine and Rehabilitation Clinics of North America, 21*, 43-58.

Hill, K., Kovacs, T., & Shin, S. (2014). Reliability of brain computer interface language sample transcription procedures. *Journal of Rehabilitation Research & Development, 51*, 579-590.

Hill, K., & Romich, B. (1999). AAC language monitoring and analysis for clinical intervention and research outcomes *Proceedings of the 14th Annual Technology and Persons with Disabilities Conference*. California State University, Northridge (CSUN). Los Angeles, CA.

Hill, K., & Romich, B. (2001). A language activity monitor for supporting AAC evidence-based clinical practice. *Assistive Technology, 13*, 12-22.

Hughes, D. L., & Till, J. A. (1982). A comparison of two procedures to elicit verbal auxiliary and copula in normal kindergarten children. *Journal of Speech and Hearing Disorders, 47*, 310-320.

Iacono, T., Mirenda, P., & Beukelman, D. (1993). Comparison of unimodal and multimodal AAC techniques for children with intellectual disabilities. *Augmentative and Alternative Communication, 9*, 83-93.

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language*, *59*, 434-446.

Kamhi, A. G. (2014). Improving clinical practices for children with language and learning disorders. *Language, Speech, and Hearing Services in Schools, 45*, 92-103.

Kent-Walsh, J., Binger, C., & Buchanan, C. (2015). Teaching children who use augmentative and alternative communication to ask inverted yes/no questions using aided language modeling. *American Journal of Speech Language Pathology, 24*, 222-236.

King, M. R., Binger, C., & Kent-Walsh, J. (2015). Using dynamic assessment to evaluate the expressive syntax of children who use augmentative and alternative communication. *Augmentative and Alternative Communication, 31*, 1-14.

Klee, T., & Fitzgerald, M. D. (1985). The relation between grammatical development and mean length of utterance in morphemes. *Journal of Child Language, 12*, 251-269.

Klein, C. (2017). *Research and AAC speakers.* Paper presented at the Clinical AAC Research Conference, Washington, DC.

Kovacs, T., & Hill, K. (2013). *Acquisition of auxiliary syntax in twins using AAC*. Paper presented at the Clinical AAC Research Conference, Bloomington, IN.

Kovacs, T., & Hill, K. (2015). A tutorial on reliability testing in AAC language sample transcription and analysis. *Augmentative and Alternative Communication, 31*, 148-158.

Kovacs, T., & Hill, K. (2017). Language samples from children who use speech generating devices: Making sense of small samples and utterance length. *American Journal of Speech Language Pathology, 6*, 939-950.

Kozulin, A. (2011). Learning potential and cognitive modifiability. *Assessment in Education: Principles, Policy & Practice, 18*, 169-181.

Kuczaj, S. (1979). Influence of contactibility on acquisition of BE: Substantial, meager, or unknown? *Journal of Psycholinguistic Research, 8*, 1-11.

LaPointe, S. G., & Dell, G. (1989). A synthesis of some recent work in sentence production. In G. Carlson & M. Tannehaus (Eds.), *Linguistic structure in language processing* (pp. 107-156). Dordrecht, The Netherlands: Kluwer.

Legate, J. A., & Yang, C. (2007). Morphosyntactic learning and the development of tense. *Language Acquisition, 14*, 315-344.

Leonard, L. B., Camarata, S., Brown, B., & Camarata, M. N. (2004). Tense and agreement in the speech of children with specific language impairment: Patterns of generalization through intervention. *Journal of Speech, Language, and Hearing Research, 47*, 1363-1379.

Leonard, L. B., Camarata, S., Pawtowska, M., Brown, B., & Camarata, M. N. (2006). Tense and agreement morphemes in the speech of children with specific language impairment during intervention: Phase 2. *Journal of Speech, Language, and Hearing Research, 49*, 749-770.

Leonard, L. B., Deevy, P., Miller, C. A., Charest, M., Kurtz, R., & Rauf, L. (2003). The use of grammatical morphemes reflecting aspect and modality by children with specific language impairment. *Journal of Child Language, 30*, 769-795.

Lidz, C. S., & Peña, E. A. (1996). Dynamic assessment: The model, its relevance as a nonbiased approach, and its application to Latino American preschool children. *Language, Speech, and Hearing Services in Schools, 27*, 367-372.

Light, J., Drager, K., McCarthy, J., Mellott, S., Millar, D., Parrish, C., . . . Welliver, M. (2004). Performance of typically developing four- and five-year-old children with AAC systems using different language organization techniques. *Augmentative and Alternative Communication, 20*, 63-88.

Loban, W. (1976). *Language development: Kindergarten through grade twelve*. Urbana, IL: National Council of Teachers of English.

Long, S. H., & Olswang, L. B. (1996). Readiness and patterns of growth in children with SELI. *American Journal of Speech Language Pathology, 5*, 79-85.

MacWhinney, B. (2000). *The CHILDES project: tools for analyzing talk* (3rd ed.). Mahwah, NJ: Erlbaum.

Maratsos, M., & Kuczaj II, S. (1978). Against the transformationalist account: A simpler analysis of auxiliary overmarkings. *Journal of Child Language, 5*, 337-345.

Millisen, R. (1954). The disorder of articulation: A systematic clinical and experimental approach. *Journal of Speech and Hearing Disorders, Monograph Supplement 4*.

Mirenda, P. (2008). A back door approach to autism and AAC. *Augmentative and Alternative Communication, 24*, 220-234.

Moyle, M., Karasinski, C., Ellis Weismer, S., & Gorman, B. (2011). Grammatical morphology in school-age children with and without SLI: a discriminant function analysis. *Language, Speech, and Hearing Services in Schools, 42*, 550-560.

Myers, R. H., Montgomery, D. C., Vining, C. G., & Robinson, T. J. (2010). *Generalized linear models with applications in engineering and the sciences* (2nd ed.). Hoboken, NJ: John Wiley & Sons, Inc.

Nakamura, K., Newell, A. F., Alm, N., & Waller, A. (1998). How do members of different language communities compose sentences with a picture-based system?: A cross-cultural study of picture-based sentences constructed by English and Japanese speakers. *Augmentative and Alternative Communication, 14*, 71-19.

Nigam, R., Schlosser, R. W., & Lloyd, L. L. (2006). Concomitant Use of the Matrix Strategy and the Mand-Model Procedure in Teaching Graphic Symbol Combinations. *Augmentative and Alternative Communication, 22*, 160-177.

Norris, R. L. (2004). *Acquisition of the T and C system in clausal complements.* (MS Master's Thesis), Massachusetts Institute of Technology.

Olswang, L. B., Bain, B. A., Rosendahl, P. D., Oblak, S. B., & Smith, A. E. (1986). Language learning: Moving from a context-dependent to -independent state. *Child Language Teaching and Therapy, 2*, 180-209.

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.

Ortloff, M. (2010). *Language development in physically and motor impaired children without speech: A case study of a child using a speech-generating device.* (MS Master's thesis), University of Cologne, Cologne, Germany. Retrieved from http://www.linguist.org/AAC.html

Paul, R. (2001). *Language Disorders from Infancy through Adolescence Assessment & Intervention* (2nd ed.). St. Louis, MO: Mosby.

Pinker, S. (1999). *Words and rules*. New York: Basic Books.

Pinker, S., & Prince, A. (1994). Regular and irregular morphology and the psychological status of rules and grammar. In S. D. Lima, R. L. Corrigan, & G. K. Iverson (Eds.), *The reality of linguistic rules* (pp. 321-351). Amsterdam: John Benjamins Publishing Company.

Pollock, J.-Y. (1989). Verb movement, universal grammar, and the structure of the IP. *Linguistic Inquiry, 20*, 365-424.

Powell, T. W., Elbert, M., & Dinnsen, D. A. (1991). Stimulability as a factor in the phonological generalization of misarticulating preschool children. *Journal of Speech and Hearing Research, 34*, 1318-1328.

Powell, T. W., & Miccio, A. W. (1996). Stimulability: A useful clinical tool. *Journal of Communication Disorders, 29*, 237-253.

Radford, A. (1990). *Syntactic Theory and the Acquisition of English Syntax*. Cambridge, MA: Basil Blackwell, Inc.

Redmond, S. M., & Johnston, S. S. (2001). Evaluating the morphological competence of children with severe speech and physical impairments. *Journal of Speech, Language, and Hearing Research, 44*, 1362-1375.

Rice, M. L. (2003). A unified model of specific and general language delay: grammatical tense as a clinical marker of unexpected variation. In Y. Levy & J. Schaeffer (Eds.), *Language competence across populations: Toward a definition of specific language impairment* (pp. 63-95). Mahwah, NJ: Lawrence Erlbaum Associates.

Rice, M. L., & Wexler, K. (1996). Toward tense as a clinical marker of specific language impairment in English-speaking children. *Journal of Speech and Hearing Research, 39*, 1239-1257.

Rice, M. L., & Wexler, K. (2001). *Rice/Wexler test of early grammatical impairment*. San Antonio, TX: PsychCorp.

Rice, M. L., Wexler, K., & Cleave, P. L. (1995). Specific language impairment as a period of extended optional infinitive. *Journal of Speech and Hearing Research, 38*, 850-863.

Rice, M. L., Wexler, K., & Hershberger, S. (1998). Tense over time: the longitudinal course of tense acquisition in children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 41*, 1412-1431.

Rispoli, M. (2016). Cross-morpheme facilitation: The systematic emergence of agreement in 2-year-olds. *Language Acquisition, 23*, 293-306.

Rispoli, M., & Hadley, P. (2011). Toward a theory of gradual morphosyntactic learning. In I. Arnon & E. V. Clark (Eds.), *Experience, variation ad generalization: Learning a first language* (pp. 15-33). Amsterdam: John Benjamins Publishing Company.

Rispoli, M., Hadley, P., & Holt, J. K. (2009). The growth of tense productivity. *Journal of Speech, Language, and Hearing Research, 52*, 930-944.

Rispoli, M., Hadley, P., & Holt, J. K. (2012). Sequence and system in the acquisition of tense and agreement. *Journal of Speech, Language, and Hearing Research, 55*, 1007-1021.

Romski, M. A., & Sevcik, R. A. (1996). *Breaking the speech barrier: Language development through augmented means*. Baltimore, MD: Paul H Brookes Publishing.

Rowland, C. F., & Theakston, A. L. (2009). The acquisition of auxiliary syntax: A longitudinal elicitation study. Part 2: The modals and auxiliary do. *Journal of Speech, Language, and Hearing Research, 52*, 1471-1492.

Rvachew, S. (2005). Stimulability and treatment success. *Topics in Language Disorders, 25*, 217-219.

Rvachew, S., & Nowak, M. (2001). The effect of target-selection strategy on phonological learning. *Journal of Speech, Language, and Hearing Research, 44*, 610-623.

Schütze, C. T., & Wexler, K. (1996). Subject case licensing and English root infinitives. In A. Stringfellow, D. Cahan-Amitay, E. Hughes, & A. Zukowski (Eds.), *Proceedings of the 20th Boston University Conference of Language Development* (pp. 670-681). Boston, MA: MIT Press.

Smith, M. (1996). The medium of the message: A study of speaking children using communication boards. In S. von Tetzchner & M. Jensen (Eds.), *Augmentative and alternative communication: European perspectives* (pp. 119-137). London: Whurr.

Smith, M., & Grove, N. (1999). The bimodal situation of children learning language using manual and graphic signs. In F. Loncke, J. Clibbens, & L. L. Lloyd (Eds.), *New directions in AAC: Research and practice* (pp. 9-30). London: Whurr.

Smith, M., & Grove, N. (2003). Asymmetry in input and output for individuals who use AAC. In J. Light, D. Beukelman, & J. Reichle (Eds.), *Communicative competence for individuals who use AAC* (pp. 163-198). Baltimore, MD: Paul H Brookes.

Snell, M. E. (2002). Using dynamic assessment with learners who communicate nonsymbolically. *Augmentative and Alternative Communication, 18*, 163-176.

Sutton, A., Gallagher, T., Morford, J., & Shahnaz, N. (2000). Relative clause sentence production using augmentative and alternative communication systems. *Applied Psycholinguistics, 21*, 473-486.

Sutton, A., & Morford, J. (1998). Constituent order in picture pointing sequences produced by speaking children using AAC. *Applied Psycholinguistics, 19*, 525-536.

Sutton, A., Soto, G., & Blockberger, S. (2002). Grammatical issues in graphic symbol communication. *Augmentative and Alternative Communication, 18*, 192-204.

Sutton, A., Trudeau, N., Morford, J., Rios, M., & Poirier, M. A. (2010). Preschool-age children have difficulty constructing and interpreting simple utterances composed of graphic symbols. *Journal of Child Language, 37*, 1-26.

Szumilas, M. (2010). Explaining odds ratios. *Journal of the Canadian Academy of Child and Adolescent Psychiatry, 19*, 227-229.

Tönsing, K. M., (2012). *Using a matrix strategy to teach graphic symbol combinations to children with limited speech during shared storybook reading*. (Ph.D. Dissertation). University of Pretoria, Pretoria, South Africa.

Tönsing, K. M., Dada, S., & Alant, E. (2014). Teaching graphic symbol combinations to children with limited speech during shared story reading. *Augmentative and Alternative Communication, 30*, 279-297.

Trudeau, N., Sutton, A., Dagenais, E., de Broeck, S., & Morford, J. (2007). Construction of graphic symbol utterances by children, teenagers, and adults: The effect of structure and task demands. *Journal of Speech, Language, and Hearing Research, 50*, 1314-1329.

Ullman, M. T. (2001). A neurocognitive perspective on language: The declarative/procedural model. *Nature Reviews Neuroscience, 2*, 717-727.

von Tetzchner, S., & Jensen, M. (Eds.). (1996). *Augmentative and alternative communication: European perspectives*. London: Whurr.

Vygotsky, L. S. (1934/2012). *Thought and Language* (Rexised and expanded ed.). Cambridge, MA: MIT Press.

Westby, C. E., Stevens Dominguez, M., & Oetter, P. (1996). A performance/competence model of observational assessment. *Language, Speech, and Hearing Services in Schools, 27*, 144-156.

Wexler, K. (1994). Optional infinitives, head movement, and the economy of derivations. In D. Lightfoot & N. Hornstein (Eds.), *Verb Movement* (pp. 305-350). Cambridge, UK: Cambridge University Press.

Wexler, K. (1998). Very early parameter setting and the unique checking constraint: A new explanation of the optional infinitive stage. *Lingua, 106*, 23-79.

Wexler, K. (2003). Lenneberg's dream: Learning, normal language development, and specific language impairment. In Y. Levy & J. Schaeffer (Eds.), *Language competence across populations: Toward a definition of specific language impairment* (pp. 11-62). Mahwah, NJ: Lawrence Erlbaum.

Wexler, K. (2011). Grammatical computation in the optional infinitive stage. In J. de Villiers & T. Roeper (Eds.), *Handbook of Generative Approaches to Language Acquisition* (pp. 53-118). New York: Springer Science.

Wexler, K., Schaeffer, J., & Bol, G. (2004). Verbal syntax and morphology in typically developing Dutch children and children with SLI: How developmental data can play an important role in morphological theory. *Syntax, 7*, 148-198.

Wiig, E. H., Secord, W., & Semel, E. (2000). *Clinical Evaluation of Language Fundamentals-Preschool UK*. London: The Psychological Corporation.

Wilkinson, K. M., Romski, M. A., & Sevcik, R. A. (1994). Emergence of visual-graphic symbol combinations by youth with moderate or severe mental retardation. *Journal of Speech and Hearing Research, 37*, 883-895.

Wilson, M. S., & Fox, B. J. (2015a). *Language Links to Literacy Experessive Syntax Assessment*. Verona, WI: Attainment Company, Inc.

Wilson, M. S., & Fox, B. J. (2015b). *LanguageLinks to literacy implementation guide*. Verona, WI: Attainment Company, Inc.