

NEUROCOGNITIVE PLASTICITY IN VERB BIAS LEARNING
IN CHILDREN AND ADULTS

BY

ZHENGHAN QI

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Neuroscience
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2012

Urbana, Illinois

Doctoral Committee:

Associate Professor Susan Garnsey, Chair, Director of Research
Professor Cynthia Fisher
Associate Professor Kara Federmeier
Assistant Professor Christopher Grindrod

ABSTRACT

Verb-specific preference for syntactic structure (verb bias) is considered as a critical parsing constraint that guides online sentence comprehension. Both adults and preschoolers show great sensitivity to verb bias in their temporary parsing commitment as sentences unfold in time. How do people learn verb bias in the first place? In natural language, frequency-sensitive verb argument structure is closely intertwined with the event information delivered by the verb and its argument, which raises complexity in teasing apart the information from linguistic co-occurrence frequency and the information from the event semantics.

In this dissertation I began by examining the independent roles of each information source during the process of updating familiar verb bias. The rest of the study focused on the verb bias learning without event cues from verb semantics. Two parallel approaches were applied to explore the details of the learning mechanisms. One set of studies used eye tracking to monitor the time course of online usage of newly learned verb bias during sentence ambiguity resolution across different age ranges. The other set of studies examined the neural stages of verb bias learning as well as the individual differences of verb bias retrieval during online sentence reading with event-related brain potential (ERP) techniques.

I demonstrated with very brief training paradigm in both listening and reading modality that children and adults were capable of quickly adapting to new information about verb-specific structural preference from the dynamic language input. The results provided evidence for a central role of linguistic distributional information in verb bias learning. Newly learned verb bias plays a similar role as the existing verb bias knowledge in affecting language users' parsing commitment and online ambiguity resolution. In addition ERP results revealed separate neural

stages that transits from semantic prediction to syntactic rule-based processing as learners continuously collected distributional information of verb-specific structural preference. Individuals who were highly sensitive to familiar verb bias also showed greater use of newly learned verb bias during conflict detection, further indicating the same mechanism underlying natural verb bias acquisition.

ACKNOWLEDGEMENTS

Looking back into my years in graduate school and how I stumbled my way through all the growing pains and joy as a researcher, I feel greatly blessed for the support and help I had received from professors, colleagues, family and friends!

First and foremost I would like to thank my advisor Dr. Susan Garnsey. I switched from animal neurobiology to human cognitive neuroscience at the end of my second year. The transition across two completely different fields was made possible only because of Susan's patient guidance and encouragement during the most difficult stage of my graduate study. Since then, she has always provided helpful advice and support as the kindest mentor, at the same time she allowed me great freedom to explore my research interest and consistently encouraged me during the exploration.

I would like to express my deep gratitude to Dr. Cynthia Fisher. I have been very fortunate to work with her since the summer in 2009. Her insights and comments immensely motivated the research direction of this dissertation. Cindy taught me how to approach scientific questions and how to enjoy the process with her own example of academic endeavor.

Dr. Kara Federmeier has also been great help in my graduate study. Over the years, I often turn to Kara and her lab for their advice about EEG experiment setup and data interpretation. I greatly appreciate many thoughtful questions, discussions and comments Kara has generously provided for a lot of my graduate projects. I would also like to thank Dr. Christopher Grindrod for his dedication as my committee member and his helpful input at various stages of my dissertation research.

The projects described in this dissertation were not possible without the assistance from many colleagues and friends. I am fortunate to have incredible help from members in the Language and Brain Lab and Language Acquisition Lab. I would like to thank all of them for their contribution in creating stimuli, running experiments and providing advice. I would like to thank especially to Lynn Brault, Cynthia Lukyanenko, Katherine Messenger, Kathleen Miller, Alison Trude and Sylvia Yuan, who provided great help in designing and setting up the eye-tracking experiments. I would like to thank Scott Jackson for his great contribution in developing scripts and toolbox for stimuli presentation and data analysis in the EEG experiments. I would like to thank Scott Fraundorf, Chia-Lin Lee, Hsu-wen Huang, Eun-Kyung Lee, Joseph Toscano, and Edward Wlotko for providing generous help in trouble-shooting technical issues as well as devoting their intellectual energy in many discussions.

I would also like to thank the language community in University of Illinois. I have been greatly benefited from various seminars, joint-lab meetings and collaborations. In particular, I would like to thank Dr. Sarah Brown-Schmidt, Dr. Duane Watson and Dr. Gary Dell for their helpful advice and intellectual input in my projects. I would like to thank Dr. Paul Gold, Dr. Donna Korol, Dr. Samuel Beshers, Dr. Gene Robinson, Dr. Arthur Kramer and Dr. Diane Beck in Neuroscience Program for their incredible help and advice in shaping my career path.

Finally, I owe a great deal to my family. My parents Meixia Huang and Huyang Qi have supported and encouraged my pursuit of science even that meant their only child being far away across the Pacific. I particularly would like to thank my mother for being my role model in my whole life with her integrity and curiosity as a scholar in her own research field. I also greatly appreciate the love and care from Xiaodan Zhuang, who has shared all the joyful and frustrating moments in my graduate school and deeply understands my enthusiasm.

TABLE OF CONTENTS

List of Tables	ix
List of Figures	x
Chapter 1 Introduction	1
1.1 Verb Bias Guides Sentence Comprehension	1
1.2 Statistical Learning Approaches for Verb Bias	5
1.3 Neurophysiological Markers of Statistical Learning	8
1.4 Overview of Dissertation	12
Chapter 2 Experience With Particular Verbs Affects Online Sentence Processing: An Eye-Movement Study	14
2.1 Introduction	15
2.2 Experiment 1: Malleability of familiar verb bias in 5-year-olds	22
2.2.1 Method	22
2.2.2 Results and Discussion	29
2.2.3 Summary	38
2.3 Experiment 2: Malleability of familiar verb bias in adults	39
2.3.1 Methods	39
2.3.2 Results and Discussion	42
2.3.3 Summary	47
2.4 Experiment 3: Verb bias learning without linguistic distributional information	50
2.4.1 Methods	50
2.4.2 Results and Discussion	53

2.4.3 Summary	58
2.5 Experiment 4: Verb bias learning without verb semantics	61
2.5.1 Methods.....	62
2.5.2 Results and Discussion	66
2.5.3 Summary	71
2.6 General Discussion	73
Chapter 3 Electrophysiological Process Underlying Novel Verb Bias Learning	80
3.1 Introduction.....	81
3.2 Experiment 5: Encoding novel verb bias	88
3.2.1 Methods.....	88
3.2.2 Results.....	97
3.2.3 Discussion	105
3.3 Experiment 6: Neural stages of novel verb bias learning	108
3.3.1 Methods.....	109
3.3.2 Results.....	111
3.3.3 Discussion	126
3.4 Experiment 7: Newly-learned verb bias in conflict detection	129
3.4.1 Methods.....	129
3.4.2 Results.....	134
3.4.3 Discussion	144
3.5 General Discussion	146
Chapter 4 Conclusion And Discussion	152
4.1 Summary Of Results.....	152

4.2 Contribution of Event-distributional Information.....	154
4.3 Ambiguous Context Boosts Effective Learning	157
4.4 Future Direction	158
4.5 Conclusions.....	160
Bibliography	161

LIST OF TABLES

Table 1: Experiment 2 Adults' offline measures.	43
Table 2: Experiment 5 Examples of the training stimuli.	91
Table 3: Experiment 5 & 6 Lexical properties of critical nouns.....	91
Table 4: Experiment 6 Picture-matching task result.....	113
Table 5: Experiment 7 Examples of the test trials..	133
Table 6: Experiment 7 Lexical properties of critical nouns.....	133
Table 7: Experiment 7 Behavioral results for target sentences by participants' responses.	137

LIST OF FIGURES

Figure 1: Experiment 1 Study phase and test phase design.	26
Figure 2: Experiment 1 Eye-movement results..	35
Figure 3: Experiment 1 Eye-movement results by test orders.	37
Figure 4: Experiment 2 Coarse fixation results.	44
Figure 5: Experiment 2 Eye-movement results.	46
Figure 6: Experiment 2 Eye-movement results by test order.	47
Figure 7: Experiment 3 Action results.	54
Figure 8: Experiment 3 Eye-movement results.	57
Figure 9: Experiment 3 Eye-movement results by vocabulary.....	58
Figure 10: Experiment 4 Eye-movement results.	70
Figure 11: Experiment 4 Eye-movement results by age.....	71
Figure 12: Experiment 5 & 6 Example of an ambiguous training trial.	95
Figure 13: Schematic map of electrodes layout.....	96
Figure 14: Experiment 5 Ambiguity effect.....	99
Figure 15 : Experiment 5 Attachment effect.....	101
Figure 16: Experiment 5 Attachment effect shifts across blocks.	103
Figure 17: Experiment 5 Predicative effect.	105
Figure 18: Experiment 6 Example of picture-matching task.	111
Figure 19: Experiment 6 Individual difference of familial handedness behavioral results.	114
Figure 20: Experiment 6 Ambiguity effect.....	116
Figure 21: Experiment 6 Attachment effect.....	117

Figure 22: Experiment 5 & 6 Attachment effect shifts across training blocks.	119
Figure 23: Experiment 6 Predicative effect.	120
Figure 24: Experiment 6 Attachment effect during the second half of training.	121
Figure 25: Experiment 6 Individual difference of familial handedness ERP results I	124
Figure 26: Experiment 6 Individual difference of familial handedness ERP results II	125
Figure 27: Experiment 7 Behavioral results for DO/SC sentences.....	135
Figure 28: Experiment 7 Behavioral results for target sentences..	136
Figure 29: Experiment 7 ERP results for DO/SC sentences.....	139
Figure 30: Experiment 7 ERP results for target sentences..	140
Figure 31: Experiment 7 Relationship between ERPs during training and ERPs during testing.	142
Figure 32: Experiment 7 Relationship between sensitivity to biases of familiar verbs and ERPs in test trials	144

CHAPTER 1 INTRODUCTION

1.1 Verb Bias Guides Sentence Comprehension

Verbs are considered a critical source of expectancy generation in sentence comprehension. The sorts of constituents that co-occur with individual verbs are known as argument structures. Verbs with the same argument structures are expected to participate in the same way in syntactic frames, and therefore provide systematic constraints on the upcoming sentence constructions (MacDonald, Pearlmuter & Seidenberg, 1994; Trueswell, 1996). A wealth of empirical findings suggests that the most likely structure for a specific verb, also called verb bias (Jurafsky, 1996), rapidly leads people's online parsing decisions during sentence comprehension.

Comprehenders actively take the frequency-based verb bias into account to predict the following words in the sentence and the intended meaning of an utterance. Adult studies of temporary ambiguity during reading have revealed that structural and semantic analyses at the disambiguating point are determined by verb bias. For example, the verbs *warn* and *worry* can both occur with sentence complements (SC) or with direct objects (DO). But these verbs differ in their co-occurrence frequency with these two types of structures: *warn* is more frequently used with a direct object, and *worry* with a sentence complement. Therefore, listeners tend to analyze a noun phrase following *warn* as a direct object. However, as the sentence (1a) unfolds, the new information in the sentence turns out to lead to inconsistency with the verb bias evidence. Readers experience temporary comprehension difficulty, showing typical "garden-path" effects. Such effects are reduced or eliminated in two ways. First, if the verb more frequently occurs with a sentence complement as *worried* in sentence (1b), readers are less likely to experience

difficulty because the sentence continuation confirms their expectation based on the verb's bias. Alternatively, if the optional complementizer *that* is included in the sentence as in (1c), the post-verbal noun phrase *the spectators* is no longer structurally ambiguous. The violation of expectation, as in (1a), has been found to cause readers to garden-path at the disambiguating word (*would*), reflected in increased reading times and different eye fixation patterns (Garnsey, Pearlmutter, Myers, & Lotocky, 1997; Trueswell, Tanenhaus, & Kello, 1993).

- (1) a. The referees warned the spectators would probably get too rowdy.
- b. The bus driver worried the passengers were starting to get annoyed.
- c. The referees warned that the spectators would probably get too rowdy.

Using online electroencephalogram (EEG) recording, Osterhout et al. reported that in the cases of DO-SC ambiguity described above, a larger P600 was elicited by the disambiguating auxiliary words (e.g. *would*, *were*) in the sentences with DO-biased verbs than that elicited by sentences with SC-biased verbs. The presence of an overt complementizer reduced the amplitude of P600 in the DO-biased conditions. The data suggested that the subcategorization bias of the verb affected the degree to which participants were garden-pathed (Osterhout, Holcomb, & Swinney, 1994).

Children's parsing preferences are also guided by verb bias (e.g., Snedeker & Trueswell, 2004). Snedeker and Trueswell (2004) tracked 5-year-olds' eyes as they responded to instructions to act on toys. Critical instructions were globally ambiguous, as in (2): each contained a prepositional phrase that could be attached to the verb and given an instrument interpretation (e.g., using the feather to tickle), or attached to the noun and given a modifier interpretation (e.g., the frog that has the feather). Snedeker and Trueswell compared instrument-bias verbs, which often occurred with instruments in a norming study (e.g., *tickle*), modifier-bias

verbs, which rarely did (e.g., *choose*), and equi-biased verbs, which fell between these extremes (e.g., *feel*). Children showed strong effects of verb bias: they more often used instruments when hearing instrument-bias verbs, and eye-movement analyses revealed that the effects of verb bias emerged quickly as children heard the noun phrase following with (e.g., *feather*).

(2) Tickle the frog with the feather.

These findings document the rapid use of verb bias in online sentence processing by children and adults. When they encounter a verb, native speakers retrieve frequency-sensitive information that they use to anticipate what kinds of phrases may follow that verb, and how those phrases should be interpreted.

What is the source of frequency-sensitive verb bias? On a syntactic view, verb bias is defined and measured with the probabilistic co-occurrence of the individual verb and the argument structure. Given the semantic richness provided by the verb argument structure in addition to the syntactic relationship, linguistic and psycholinguistic studies have explored the close relationship between the verb semantics and verb bias to account for the formation of verb bias (Grimshaw, 1979; Hare, McRae, & Elman, 2004; Levin, 1993). For example, the structural probabilities of a set of verbs were found to be correlated with the semantic subcategorization probabilities of their corresponding nouns (e.g. *proposed* – *proposal*), supporting the view that verb's argument structural choice and verb semantics are associated (Argaman & Pearlmuter, 2002). Hare et al., 2003 demonstrated that comprehenders' online expectation about the syntactic continuation after a particular verb was dependent on the sense of that verb (e.g. concrete action vs. mental attitude), manipulated by the discourse information. For instance, the DO sense of the verb *find* usually represents a concrete action "locating" (e.g. *He found the book on the shelf*), while the SC sense of the verb *find* describes a mental event (e.g. *He found the plane has taken*

off) (Hare, McRae, & Elman, 2003). Corpus analysis showed that verb bias has to be attached to senses of verbs rather than a superordinate lexical representation (Hare et al., 2004). These results suggested probabilistic verb bias could be partially determined by the semantic profile of the verb.

In addition, the semantic relationship between a verb and its argument structure incorporates the link between the described event information and the syntactic construction. How event knowledge interacts with linguistic knowledge in influencing people's online comprehension has been profoundly studied. Empirical evidence found priming effect between a particular verb and its thematic roles, such as agent and instrument, in both directions (Ferretti, McRae, & Hatherell, 2001; McRae, Ferretti, & Amyote, 1997; McRae, Hare, Elman, & Ferretti, 2005). For example, on one hand, during online sentence comprehension, people anticipated an edible object (e.g. *cake*) in a visual-world paradigm more quickly after hearing the verb *eat*, as opposed to after hearing the verb *move*, reflecting the activation of the combinatorial semantic/thematic knowledge that is verb-specific (Altmann & Kamide, 1999; Kamide, Altmann, & Haywood, 2003). On the other hand, Boland (2005) argued that the anticipatory looks are limited to verb argument. For instance, listeners would look at potential recipient argument (e.g. *to the owner*) of a dative verb (e.g. *mention*) in a visual-world paradigm before it has been mentioned in the sentence. However, they rarely showed anticipatory looks to the potential benefactive adjunct (e.g. *for the owner*) after hearing an action verb (e.g. *fix*). These results add into the close relationship between verb semantics and the structures that follow the verbs. Taking together, the situational schema built upon the semantic relationship between verbs and their argument structure may serve as another important component of verb bias.

The current study attempts to address the formation of verb bias from the perspective of language learners, by examining how children and adults learn new verb bias from their language experiences, in particular how linguistic distributional and event distributional history of a particular verb each contributes into verb bias learning process.

1.2 Statistical Learning Approaches for Verb Bias

The existence of statistical regularities at multiple linguistic levels has been shown to be of use in learning language. A growing body of empirical evidence have shown that young infants, children, and adults are capable of exploiting statistical information such as raw frequency, frequency of co-occurrence, or transitional probability to learn speech categories (Anderson, Morgan, & White, 2003; Maye, Weiss, & Aslin, 2008; Maye, Werker, & Gerken, 2002), to detect phonotactic structures (Chambers, Onishi, & Fisher, 2003), to locate word boundaries in continuous speech flow (Saffran, 2001; Saffran, Aslin, & Newport, 1996), to learn word sequences (Gomez & Gerken, 1999) to form syntactic categories (Altmann, 2002; Gerken, Wilson, & Lewis, 2005), and to build abstract representations of sequential patterns (Gerken et al., 2005; Gomez & Gerken, 1999). For example, preverbal infants who are exposed to a continuous pseudospeech stream for 2 min can learn that some syllables are more likely to co-occur than others, providing a bootstrapping mechanism by which “words” can be identified (Saffran et al., 1996). In the visual domain, adult observers are sensitive to contingencies between shapes in temporal sequences (Fiser & Aslin, 2002; Turk-Browne, Jungé, & Scholl, 2005) and spatial configurations (Fiser & Aslin, 2001). Gómez and Gerken (1999) showed that after brief exposure to a simple artificial grammar, 12-month-old children could distinguish new

grammatical from nongrammatical sequences, suggesting that language learners are capable of detecting distributional regularities in the input statistics to acquire language structures.

Even for the acquired statistical pattern, adults continuously adapt themselves to the dynamic probabilistic cues in the language input and update their parsing strategy during online language comprehension (Bock & Griffin, 2000; Chang, Dell, Bock, & Griffin, 2000; Chang, Dell, & Bock, 2006). For example, several weeks of increased experience with less-frequent object-relative structures alleviated the difficulty in processing object-relative structures (Wells, Christiansen, Race, Acheson, & MacDonald, 2009). A recent study reported rapid updating of parsing strategy based on the changes of frequency-sensitive cue validity (Fine & Jaeger, 2011). Participants read sentences containing direct-object and sentential-complement biased verbs. Across experimental trials, participants learned to adjust their expectation based on the probabilistic structural experience with the familiar verbs. In a low-reliability group, verb bias was constantly violated in the sentence and no longer provided valid cues for structural prediction. Therefore participants learned to rely more on a different disambiguating cue (e.g. complementizer *that*) for their parsing decisions.

The robustness of successful learning of statistics in various aspects of language raises the question what kinds of learning experiences might lead children and adults to learn to expect a particular structure following a certain verb, *e.g.* an instrument prepositional phrase following *tickle* but not *choose*?

Multiple information sources (not mutually exclusive) may contribute to verb-bias learning. One potential source is event-distributional information, via observation of events in the world. Upon encountering a verb, listeners might retrieve conceptual knowledge of the referent event, and use it to determine what kinds of event participants are plausible (e.g.,

Ferretti, McRae & Hatherell, 2001). On the other hand, many have argued that word learning and sentence interpretation depend on analyses of the linguistic contexts in which words appear (e.g., Gleitman, 1990; Trueswell & Gleitman, 2004). Such considerations suggest another possibility, that verb bias learning might be based on linguistic distributional information about how speakers use each verb in sentences. Thus, we might treat *tickle* as an instrument-bias verb because we often hear this verb in sentences with a verb-attached prepositional phrase that describes an instrument.

In natural language exposure, these information sources are nearly always confounded. Attempts have been made to disentangle these sources and to explore their contributions in verb learning. Some of the strongest evidence for the independent role of linguistic distributional information in the verb lexicon comes from experiments in which 2-year-olds learn combinatorial facts about novel verbs before being shown any events at all (Arunachalam & Waxman, 2010; Yuan & Fisher, 2009). Yuan and Fisher (2009) showed 2-year-olds dialogues in which two women used a made-up verb in transitive (e.g., “Jane blicked the baby!”) or intransitive sentences (“Jane blicked!”). In a later test, children’s interpretations of the novel verb were influenced by the dialogue experience: those who had heard transitive dialogues looked longer at a two-participant event (as opposed to a one-participant event) than did those who had heard intransitive dialogues. These studies examined the learning of absolute subcategorization facts about novel verbs rather than probabilistic verb biases; nonetheless, these data show that toddlers can encode linguistic distributional facts about verbs, without situational knowledge.

Artificial grammar learning is another approach to uncouple the sources for statistical learning of verb bias. For example, Wonnacott et al. (2008) found that adults who were learning to produce and understand sentences in a miniature language encoded verb-specific constraints

on word order that affected online sentence comprehension. During the study phase, adults listened to and repeated the sentences containing novel verbs either in the verb-agent-patient or verb-patient-agent-article structure. Across subjects, the two structures referred to the same events. The newly learned verb-specific combinatorial facts about the subsequent structures affected both language production and online sentence comprehension. The authors argued that the verb bias learning effect was independent from the semantic cues of the verbs, because no covert event categories differentiated the verb distribution in the experiment (Wonnacott, Newport & Tenenhaus, 2008). Amato and MacDonald (2010) reported online reading time data supporting more complex probabilistic learning in an artificial language learning task. Adults learn to anticipate a particular direct object (e.g. *clate*) based on its co-occurrence with the combination of the subject noun (e.g. *pim*) and verb (e.g. *dak*), while neither the noun nor the verb alone had predictive values. Although it was clear that linguistic distributional learning played an important role in verb-specific structural learning, the learning process in both experiments took place in a referential context where nonsense verbs and their constituents were provided with semantic meanings.

To summarize, these results yield strong evidence for the experience-dependent plasticity of the language system that is sensitive to linguistic distributional information, and the accessibility of the newly learned combinatorial facts during online sentence comprehension.

1.3 Neurophysiological Markers of Statistical Learning

Studies of statistical learning have explored its power and flexibility across multiple modalities, suggesting it is important for language acquisition, as the process involves the extraction of regularities and patterns distributed across a set of exemplars in time and/or space.

However the neural processes recruited during linguistic statistical learning has received little attention until recently. The following paragraphs will review a number of recent studies providing evidence from Electroencephalography (EEG) and functional magnetic resonance imaging (fMRI).

Two major event-related potential (ERP) components, N400 and P600 (or SPS, syntactic positive shift), are considered as indices for the degree of semantic and syntactic processing (Hagoort, Brown, & Groothusen, 1993; Kutas & Hillyard, 1984; Osterhout & Holcomb, 1992) and therefore are studied widely in the language learning studies. N400 was found to be sensitive to the degree of training exposure from multiple training modalities, mostly using a word /sequence learning paradigm. McLaughlin and colleagues studied learning-related N400 changes in French L2 word learning and found a gradual learning process of word meaning in the brain without overt lexicality judgment. The word-pseudoword N400 effect increased as a function of hours of instruction during French learning (McLaughlin, Osterhout, & Kim, 2004). Another cross-sectional word-learning study showed reading words from a second language elicited reduced amplitude of N400 as opposed to reader's native language. More importantly, the L1-L2 N400 effect became smaller in more proficient language learners (Midgley, Holcomb, & Grainger, 2009). A recent study provided evidence for a similar word learning effect changing across the training stages within only fourteen minutes of passive exposure to new spoken pseudo-words (Shtyrov, Nikulin, & Pulvermüller, 2010). Besides evidence from second language learning literature, N400 also responds to the degree of learning in a study using nonlinguistic auditory sequences. Participants with high behavioral accuracy in the post-training recall task showed a larger N400 effect to the onset of the familiar sequence compared to those with low behavioral accuracy (Abla, Katahira, & Okanoya, 2008).

Artificial grammar learning studies using statistical training paradigm found the P600 effect responding to ungrammatical artificial structure undistinguishable from the effect found in natural language processing (Christiansen, Conway, & Onnis, In Press; Friederici, Steinhauer, & Pfeifer, 2002; Hsu, 2009), suggesting a similar neural processes shared between statistical learning in artificial sequences and grammar learning in natural language. These studies provided support for the crucial role of statistical learning in syntactic acquisition. A series of second language studies recently in Osterhout group tracked different learning stages to explore when and how learners started to categorize the information from language input and initiated their grammatical knowledge, such as gender, number and person agreement. The findings showed an N400 effect at an earlier stage and a P600 effect at a later stage of learning, when subjects were tested on a grammaticality judgment task. The two discrete neural stages were explained to reflect the transition from rote-memorized lexical processing to rule-based grammatical processing (McLaughlin et al., 2010; Osterhout et al., 2008, 2006).

Using functional MRI, several studies examined the localization of neural substrate for statistical learning. Activity in left inferior and middle frontal gyri was taken to index the detection for word boundary, which showed sensitivity to the frequency of the trisyllabic combinations in a continuous speech stream (McNealy, Mazziotta, & Dapretto, 2006). In the visual domain, responses to statistically structured sequences of shapes were observed in the striatum and medial temporal lobe, raising interesting possibility that multiple memory systems contribute to statistical learning in parallel (Turk-Browne et al., 2005). A recent case study of a split-brain patient revealed the right-hemisphere dominance in visual statistical learning (Roser, Fiser, Aslin, & Gazzaniga, 2010). In contrast to normal control participants, the patient could not discriminate fixed-pair shapes from randomly-combined shapes, except when the training and

testing items were both presented to her left visual field / right hemisphere. These results suggest the important role of right hemisphere in both learning and retrieval of statistical patterns. Very few studies have directly explored how these neural substrates for statistical learning are recruited in language acquisition, except for the evidence about the influence of proficiency in the second language acquisition literature (Kotz, 2009). The left inferior frontal gyrus in the Broca's area has been linked with syntactic learning. It appears that late learners showed an increase recruitment of inferior frontal gyrus compared to early learners (Rüschemeyer, Fiebach, Kempe, & Friederici, 2005; Rüschemeyer, Zysset, & Friederici, 2006; Tatsuno & Sakai, 2005). Despite the implications of the neural characteristics that are sensitive to the extent of language exposure, the fine-grained process underlying distributional learning in language is yet unknown.

In sum, the existing neural evidence for language learning lent support for the neural sensitivity to the statistical distribution in language input. However, learning efficacy was usually measured hours, days and even weeks after language exposure. Therefore little is known about how and what type of information is processed in real time during learning. As a reliable frequency-sensitive parsing cue for sentence comprehension, verb bias is presumably derived from the distributional information in events and language. Studies on verb bias learning over the course of training will reveal the real-time neural responses to the dynamic statistical information in the language input.

1.4 Overview of Dissertation

The present study seeks to explore the statistical learning mechanisms of verb bias acquisition. The type of verb bias chosen for training across multiple experiments is *with* prepositional phrase (*with*-PP) attachment, because the same surface structure shared between the structural alternatives made both global ambiguity and temporary ambiguity possible. *With*-PP attachment ambiguity offers great flexibility for studying verb bias effect during online ambiguity resolution as well as conflict detection in garden-path sentences.

I examined learners' sensitivity to linguistic distributional information in 5-year-olds and adults. The goals of the study are to address (1) whether children and adults are able to capture the verb-specific structural preference from a brief listening or reading experience; (2) how linguistic and event distributional information plays a role in verb bias learning; (3) whether language users readily rely on the newly learned verb bias to make syntactic commitment and to resolve sentence ambiguity as they would have done with the existing verb bias knowledge.

The rest of the dissertation is organized as follows. Chapter 2 presents a set of eye-movement experiments, testing the usage of newly learned verb bias during online sentence ambiguity resolution. The chapter begins by examining the malleability of familiar verb bias in both 5-year-olds (Experiment 1) and adults (Experiment 2). Then Experiment 3 and 4 further teased apart the event and linguistic distributional training information to examine the necessity and sufficiency of the latter for verb bias learning in 5-year-olds. Chapter 3 presented three event-related potential (ERP) experiments, providing real-time neural evidence for the stages of verb bias learning as well as verb bias usage after acquisition. Experiment 5 and 6 focused on identifying the ERP responses at different stages of verb bias learning as adults were collecting distributional information from sentence reading. Experiment 7 tested how newly learned verb

bias was used for conflict detection in garden-path sentences on the next day after training. In addition, Experiment 6 and 7 reported individual differences in verb bias learning efficacy and proposed possible explanation. The last chapter summarized and discussed the key findings from the eye-movement and ERP studies.

CHAPTER 2 EXPERIENCE WITH PARTICULAR VERBS AFFECTS ONLINE SENTENCE PROCESSING: AN EYE-MOVEMENT STUDY

Verbs' probabilistic subcategorization biases affect sentence processing in adults and children (Snedeker & Trueswell, 2004). For example, in "Tickle the pig with the flower" the underlined prepositional-phrase (PP) could stipulate an instrument for tickling, or a modifier specifying which pig to tickle. Children's interpretations of such sentences depend on whether the verb takes instrument-PPs often ("tickle"), or rarely ("choose"). These effects emerge quickly as children listen, showing that verb bias guides incremental sentence comprehension. Artificial-grammar-learning experiments suggest that adults can learn verb biases from the statistics of language experience, independent of meaning differences among verbs (Wonnacott et al., 2008). Moreover, 2-year-olds learned a novel verb's transitivity through listening experience, before learning the verb's semantic content (Yuan & Fisher, 2009). The present study tests the possibility whether children and adults' keep track of the statistics in a very brief language experience and learn new information of specific verbs, as reflected in online measures of sentence processing. Across four experiments, we assessed the independent role of linguistic and event-distributional information in verb bias learning. The findings demonstrated that hearing familiar verbs combined with clear instrument- or modifier-PPs influenced 5-year-olds' and adults' later incremental interpretation of sentences containing the same verbs. Importantly, our results also support the central role of linguistic distributional information in verb bias learning. For 5-year-olds, the combinatorial fact of a particular verb and its following structures is critical for effective retrieval of verb-specific information during online sentence parsing and such effect is independent from verb semantic knowledge.

2.1 Introduction

A wealth of experimental findings suggest that probabilistic knowledge of the sentence structures co-occurring with particular verbs -- *verb bias* -- plays an important role during real-time sentence parsing. Across a number of different experimental settings and sentence types, adults show robust sensitivity to verb bias differences (Boland, Tanenhaus, & Garnsey, 1990; Garnsey, Pearlmutter, Myers, & Lotocky, 1997; Trueswell & Kim, 1998; Trueswell, Tanenhaus, & Kello, 1993). The influence of verb bias knowledge on online interpretation emerges early in language development. In a seminal study, Snedeker and Trueswell (2004) tracked adults' and children's eyes as they followed instructions to act on toys. The critical instructions were globally ambiguous, as in (1): Each contained a prepositional phrase that could be attached to the verb phrase and given an instrument interpretation (e.g., use the fan to tickle the pig), or attached to the noun phrase and given a modifier interpretation (e.g., tickle the pig that has the fan). Snedeker and Trueswell compared sentences containing instrument-bias verbs, which often occurred with instruments in a norming study (e.g., *tickle*), modifier-bias verbs, which rarely did (e.g., *choose*), and equi-biased verbs, which fell between these extremes (e.g., *feel*). Both 5-year-olds and adults showed strong effects of verb bias in off-line and online measures of sentence interpretation. They more often used instruments to carry out the requested actions when hearing instrument-bias verbs, and eye-movement analyses revealed that the effects of verb bias emerged quickly as listeners heard the noun phrase following the preposition (e.g., *fan*).

(1) *Instrument bias*: Tickle the pig with the fan.

Equi-bias: Choose the cow with the barrette.

Modifier bias: Feel the frog with the feather.

Later studies confirmed children's sensitivity to verb bias in offline and online measures of sentence interpretation, and also revealed intriguing evidence that preschoolers rely more strongly on verb bias than on other constraints that strongly guide adult parsing. This includes constraints based on the referential context in which the sentence is presented. Adults more readily arrive at a modifier interpretation of an ambiguous prepositional phrase when a modifier is needed in context to specify the intended referent. That is, upon hearing "the pig with the fan," adults give less consideration to the instrument interpretation if there are two pigs present (one with a fan) rather than only one (Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995). Five-year-olds, in contrast, show little or no sensitivity to this referential context manipulation (Kidd & Bavin, 2007; Snedeker & Trueswell, 2004; Trueswell, Sekerina, Hill, & Logrip, 1999; Weighall, 2008). Based on such evidence, Trueswell and Gleitman (2004) have argued that the language comprehension system becomes sensitive to various information-sources in a developmental sequence that is partly predictable from the validity of each information source. On this view, verb bias effects emerge early in development, and strongly, because verbs are very reliable cues to upcoming sentence structure.

How do people learn verb bias? Multiple information sources (not mutually exclusive) may contribute to verb-bias learning.

One potential source is knowledge about the world events named by each verb. Upon encountering a verb, adult listeners retrieve conceptual knowledge about the referent event, and use it to determine what kinds of event participants are plausible (e.g., Ferretti, McRae & Hatherell, 2001; see also Altmann & Kamide, 1999). Even preschoolers do something similar, looking towards a juice glass rather than a car when they hear *drink* (Fernald, Zangl, Portillo, & Marchman, 2008). Such inferences can be mediated by multiple parts of sentences, not only by

the verb itself. For example, adult listeners quickly integrated a subject noun phrase and verb to predict likely direct object nouns, looking more at a motorcycle (rather than a carroussel) when hearing "The man will ride..." than when hearing "The (little) girl will ride..." (Kamide et al., 2003). Such findings suggest that real-world knowledge of events rapidly guides linguistic predictions in context. Conceptual knowledge of plausible event participants, in turn, might in large part be derived from observation of world events; we will call this *event distributional information*. Thus, we might think of instruments when we hear *tickle* because this word refers to a class of events that, in our experience, have often involved an instrument.

On the other hand, many have argued that word learning and sentence interpretation depend on analyses of the linguistic contexts in which words appear (e.g., Gleitman, 1990; Trueswell & Gleitman, 2004). Verbs in particular pose problems for learning based on observation of world events alone (Gillette, Gleitman, Gleitman, & Lederer, 1999), and even toddlers use syntactic evidence to guide the assignment of meanings to new verbs (e.g., Naigles, 1990; Yuan, Fisher, & Snedeker, in press). Such considerations suggest another possibility, that verb bias learning might be based on *linguistic distributional information* about how speakers use each verb in sentences. Thus, we might treat *tickle* as an instrument-bias verb because we often hear this verb in sentences with a verb-attached prepositional phrase that describes an instrument.

In natural language exposure, these information sources are nearly always confounded. Speakers use instrument prepositional phrases with *tickle* because they are talking about events in which tickling involved an instrument. Recently, however, several experiments have attempted to disentangle these sources and to explore their separate contributions to verb learning.

Some of the strongest evidence for the independent encoding of linguistic distributional information in the verb lexicon comes from experiments in which 2-year-olds learned combinatorial facts about novel verbs before being shown any events at all (Arunachalam & Waxman, 2010; Scott & Fisher, 2009; Yuan & Fisher, 2009). Yuan and Fisher (2009) showed 2-year-olds dialogues in which two women used a made-up verb in transitive (e.g., “Jane blicked the baby!”) or intransitive sentences (“Jane blicked!”). In a later test, children’s interpretations of the novel verb were influenced by the dialogue experience: those who had heard transitive dialogues looked longer at a two-participant event (as opposed to a one-participant event) than did those who had heard intransitive dialogues. These dialogue effects persisted after a day’s delay, but disappeared if children heard no novel verb at test (Yuan & Fisher, 2009), or heard a different novel verb (Scott & Fisher, 2009). These studies examined the learning of absolute subcategorization facts about novel verbs rather than probabilistic verb biases (i.e. whether the new verb is transitive or intransitive); nonetheless, these data show that toddlers can link linguistic distributional facts with particular verbs, without knowing what kind of event the verb refers to.

Artificial grammar learning provides another way to uncouple the data sources for verb-bias learning. For example, Wonnacott, Newport and Tanenhaus (2008) found that adults who were learning to produce and understand sentences in a miniature language encoded verb-specific constraints on word order that affected online sentence comprehension. The experimental language permitted two word orders for transitive verbs (agent-first or patient-first), differentiated by a function-word-like marker; particular verbs appeared preferentially with one or the other of these orders. Crucially, although all exposure sentences were accompanied by the events they described, Wonnacott et al. argued that verb-bias learning in their experiment

resulted from linguistic distributional and not from event distributional learning, because verbs were assigned randomly to particular word-order biases; thus no situational criteria could be used to predict verbs' syntactic behavior.

Relatedly, Amato and MacDonald (2010) reported online reading time data suggesting more complex probabilistic learning in an artificial language learning task. As noted earlier, adults integrate knowledge of familiar subject noun-phrases and verbs to predict likely direct objects (Kamide et al., 2003). Amato and MacDonald taught adult participants a similar pattern in brief exposure to an artificial language. Adults learned to anticipate a particular direct object noun based on its co-occurrence with a particular subject noun and verb, where neither the subject nor the verb alone predicted the direct object noun. Note that these results do not permit us to disentangle the effects of event versus linguistic distributional learning. Exposure sentences were always accompanied by depictions of the referent events; therefore participants could have learned that a particular creature often did a particular action to a particular object, or that particular nouns and verbs co-occurred consistently, or both.

It is difficult to extend this artificial language learning paradigm to children. However, in one recent study 6-year-olds learned a semi-artificial language in which familiar nouns were combined with made-up determiner-like particles (Wonnacott, 2011). In a post-training production task, children showed that they had learned not only which particles were common in the task, but which particle went with each noun. These findings suggest that children, like adults, can learn semantically arbitrary distributional facts about words. These data comport with a wealth of evidence that infants and toddlers learn distributional facts about nonsense syllables, even when no word has yet been assigned any meaning (e.g., Gómez & Gerken, 1999; Marcus, Vijayan, Bandi Rao, & Vishton, 1999; Saffran, Aslin, & Newport, 1996).

One recent language production experiment suggests that the biases of familiar verbs can be modified in adults. Coyle and Kaschak (2008) asked participants to complete sentence stems containing dative verbs. In the biasing phase, the sentence stems contained object nouns that induced participants to produce double-object datives with one verb (e.g., *The teacher sent the student ...*) and prepositional datives with another verb (e.g., *The man handed the book ...*). At test, the stems to be completed ended at the verb (e.g., *The boy handed ...*), leaving it up to the participant which structure to use. Coyle and Kaschak found effects of the biasing phase on later productions: Participants produced more double-object sentences for the verb that they had been induced to produce in that structure in the biasing phase. Thus, adults encoded new information about the structural biases of known verbs when producing sentences, and this information biased later sentence production. Coyle and Kaschak's (2008) finding also suggests that the modification of familiar verbs bias in this case derived from linguistic distributional information, because dative verbs such as *hand* describe essentially the same events whether they are used in double-object or prepositional dative sentences (Coyle & Kaschak, 2008).

Building on these findings, our experiments explore how experience with particular verbs influences children and adults' online sentence comprehension. All four experiments comprised an initial study phase and a subsequent test phase. In the study phase, participants watched dialogue videos in which two women used a set of verbs to talk about unseen events. Each verb appeared in multiple sentences, many of which contained an instrument phrase for the verb or a modifier phrase for the direct object of the verb. In Experiment 1, these instrument or modifier phrases were syntactically ambiguous *with* prepositional phrases; the surrounding sentences and noun choices in these critical sentences were designed to bias listeners toward either an instrument or a modifier interpretation (see Example 2). In the subsequent test phase, listeners

responded to sentence instructions that included ambiguous *with* prepositional phrases. Participants' actions and eye-movements as the sentences unfolded were monitored to determine whether their interpretations of the ambiguous *with* phrases were affected by whether they had heard the verb in each test sentence in an instrument- or in a modifier-training dialogue during the initial study phase. If this brief language experience can modify children's verb biases, then children should retrieve the new verb-bias information encoded during the training dialogue when they encounter the same verb in the test phase. As a result, they should be more likely to consider a modifier interpretation of an ambiguous *with* phrase when encountering modifier-trained verbs, and an instrument interpretation when encountering instrument-trained verbs. In Experiment 2, we tested adults to further explore the plasticity of the language processing system's use of verb bias. If adults are also able to update their verb biases from a brief learning experience, this experiment will provide stronger evidence from the continuous influence of linguistic distributional information on verb bias acquisition.

(2a) A: What did Tim use to point at the tiger?

B: He pointed at the tiger with the red pencil.

(2b) A: Which tiger did Tim point at?

B: He pointed at the tiger with the large paws.

The purpose of the current study is to address the separate roles of linguistic and event-distributional information in verb-bias learning. In Experiments 1 and 2, we tested the malleability of familiar verb bias in both 5-year-olds and adults. But note in (2) that the dialogue exposure confounds linguistically-conveyed event information and linguistic distributional information. Here we used two tactics to disentangle the effects of linguistic exposure from the event information conveyed by sentences. First, Experiment 3 explored whether linguistic-

distributional information is necessary for verb-bias learning, by replacing the syntactically ambiguous *with* phrases with unambiguous phrases that differed in their words and their syntactic structures from the test sentences. In the training dialogues for Experiment 3, instruments were expressed with a phrase such as *using the red pencil*, and modifiers were expressed with a relative clause such as *that has the large paws*. This allowed us to ask whether linguistically-conveyed event information that was not couched in the same words and structures could modify children's verb-bias knowledge. Finally, in Experiment 4 we replaced the familiar verbs in the dialogues with made-up verbs. Thus children heard sentences such as "I veebed the tiger with the sharp teeth" rather than the sensible sentences in (2). This manipulation allowed us to drastically reduce the semantic information conveyed by the verb and to focus more on the contribution from linguistic-distributional information.

2.2 Experiment 1: Malleability of familiar verb bias in 5-year-olds

In Experiment 1 we provided children with additional experience with a set of familiar verbs in the laboratory, to determine whether simply listening to sentences could modify the biases of known verbs. If children's familiar verb bias is malleable, then new linguistic experience involving these verbs might affect children's online sentence processing during the test phase.

2.2.1 Method

2.2.1.1 Participants

Sixty-four 5-year-olds (mean 5; 0 months, range 4; 3 - 5; 10 months; 31 boys, 33 girls) participated. Sixteen additional children were eliminated from analysis, because they declined to

participate (8), contributed too few usable trials (6; see Coding, below); or showed evidence of misunderstanding the study phase dialogues (2; see Procedures, below).

2.2.1.2 Materials

The materials included two types of dialogue videos to be presented in the study phase, Instrument-training and Modifier-training dialogues, and prerecorded instructions accompanied by appropriate sets of toy animals and objects for the test phase. The verbs in the dialogues and the critical instructions were the 8 equi-bias target verbs [*point at*, *turn over*, *throw*, *scratch*, *pinch*, *feel*, *blow on*, and *drag*] from Snedeker and Trueswell (2004).

For the study phase, we created for each target verb two short videos showing two women talking about unseen events; one of the dialogue videos for each verb was an Instrument-training dialogue, and the other was a Modifier-training dialogue. Each dialogue contained four critical sentences containing structurally ambiguous *with* phrases, with discourse context and noun-phrase choices strongly promoting either modifier or instrument interpretation of the *with* phrases. For example, (3) shows the Instrument- and Modifier-training dialogues for the verb *point*. In the Instrument-training dialogue (3a), the question “What did Tim use...” and the prepositional phrase objects “red pencil” and “magic sword” strongly implied an instrument interpretation of the *with* phrases. In the Modifier-training dialogue (3b), the question “Which tiger...” and the prepositional phrase objects “sharp teeth” and “large paws” implied a modifier interpretation. We also created 2 filler dialogues that did not contain a target verb or *with* phrases. The dialogue participants produced the sentences in an animated, child-directed style, and produced gestures that emphasized the identity or properties of the prepositional phrase objects. For example, while referring to a magic sword, the speaker gestured as if holding a

sword; while referring to tiny tweezers, the speaker held up thumb and forefinger as if measuring something very small. This was done both to make the dialogues more engaging, and to provide additional cues guiding the children toward the intended interpretation of the sentences.

(3a) Instrument-training dialogue

A: Do you remember that story about the tiger? What did Tim use to point at the tiger?

B: He pointed at the tiger with the red pencil.

A: Right, he pointed at the tiger with the red pencil. Hmm. If I were him, I would point at the tiger with the magic sword.

B: Wow! You want to point at the tiger with the magic sword! How exciting!

(3b) Modifier-training dialogue

A: Do you remember that story about the tiger? Which tiger did Tim point at?

B: He pointed at the tiger with the large paws.

A: Right, he pointed at the tiger with the large paws. Hmm. If I were him, I would point at the tiger with the sharp teeth.

B: Wow! You want to point at the tiger with the sharp teeth! How exciting!

For the test phase, we recorded 8 critical sentences and 24 filler sentences; all were taken from the materials of Snedeker and Trueswell (2004). Each critical sentence contained a target verb followed by an animal name in direct object position, and an ambiguous *with* phrase, as in example (4). None of the nouns used in the test sentences had appeared in the training dialogues.

(4) Point at the pig with the flower.

As shown in Figure 1, each test sentence was accompanied by a set of toys: a Target Animal with a small replica of the target instrument (e.g., a pig holding a flower), a Distracter

Animal with a small replica of the distracter instrument (e.g., an elephant holding a crayon), a Target Instrument (e.g., a large flower) and a Distracter Instrument (e.g., a large crayon). These toys made available both modifier and instrument interpretations of the *with* phrase. That is, the display contained a large flower with which to point at the pig, and a pig that had a flower. These trials thus represented the one-referent condition of Snedeker and Trueswell's (2004) experiment. In their two-referent condition, the distracter and target animals were of the same kind (e.g. two pigs holding different mini-instruments). The one-referent condition was chosen for the current experiment so that we could use the same materials in Experiment 2 for adults: Snedeker and Trueswell found that adults overwhelmingly interpreted ambiguous prepositional phrases as modifiers in the two-referent context with equi-bias verbs. However, as discussed earlier, children have repeatedly failed to show such effects of referential context in previous experiments.

2.2.1.3 Procedure

In the study phase (Figure 1A), children watched 10 dialogue videos presented on a laptop computer. They saw eight critical dialogues, one for each of the eight equi-bias verbs, and two filler dialogues. Instrument and modifier training dialogues were arranged in a quasi-random order. The first and the last dialogues in the study phase were fillers. After each dialogue, the experimenter asked the child to repeat the last sentence of the dialogue ("What did she just say?") while pointing at the actor who produced the last sentence. The experimenter then repeated the last sentence whether or not the child had repeated it correctly, and the child got a sticker as a reward if the answer was on-task. This was done to ensure that children attended to the dialogues, and to assess their comprehension of the dialogue sentences.

Children's attempts to repeat the dialogue sentences were transcribed for later coding. These repetitions varied greatly in accuracy, with some children typically repeating only a few words of the sentence ("with the sword"). However, in some cases children's repetitions of the dialogue sentences revealed clear evidence of misunderstanding. For example, instead of answering *she wanted to point at the tiger with the sharp teeth / magic sword*, children might say "point at the teeth" or "the tiger has a sword". Two children who produced two or more such responses (out of 4 dialogues per condition) were replaced in the design.

1A. Study phase



1B. Test phase

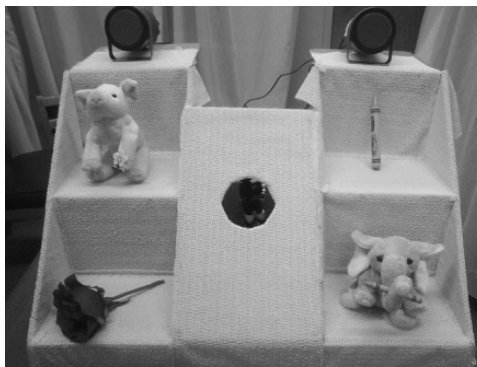


Figure 1: Experiment 1 Study phase and test phase design. A. Dialogue video showing two women using a target verb in a modifier- or instrument-bias context; B. Toy layout for the sentence *Point at the pig with the flower*. The target animal is the pig, which is holding a flower. The target instrument is the big flower.

During the test phase (Figure 1B), children sat in front of an inclined podium with a shelf in each quadrant on which toys could be placed, and a central opening for a camera that recorded the children's eye-movements. Another camera positioned behind the children recorded their actions. The recorded instructions were played from a laptop computer through external speakers positioned at the top of the podium display. Children were told that they would play a game involving following instructions. At the start of each trial, the experimenter placed the toys in each quadrant of the podium, introducing each by name as she did so. The mini-instruments attached to each animal were introduced as separate objects (e.g., "Here's a crayon, a pig, a

flower, another flower, an elephant, and another crayon”). After naming the objects in this manner a second time, the experimenter played the pre-recorded instructions for the trial. Each trial began with an instruction to "look at the camera", followed by two other instructions. In critical trials the first instruction was the critical sentence and the second was a filler sentence. Each child received 4 critical and 5 filler trials. This procedure is modeled closely on the one described by Snedeker and Trueswell (2004).

Each child also completed the Peabody Picture Vocabulary Test—Third Edition (PPVT-III) at the end of the study, as a measure of receptive vocabulary skills (Dunn & Dunn, 1997). The number of correct items was included as the measure of vocabulary in the statistical analysis.

2.2.1.4. Design

The 8 equi-bias verbs were divided into two lists, with verbs matched across lists on their probability of occurring with instrument phrases in the norming data reported by Snedeker and Trueswell (2004). The lists were combined such that each child heard instrument-training dialogues for 4 verbs and modifier-training dialogues for the other 4, and each verb was presented in a modifier-training dialogue to half of the children, and an instrument-training dialogue to the other half. In the test phase, children received only 4 critical test trials – each child was tested either on the verbs they had heard in instrument-training dialogues, or on the verbs they had heard in modifier-training dialogues. Participants were randomly assigned to the instrument or the modifier-testing group. The two groups did not differ from each other in age or vocabulary.

2.2.1.5 Coding

Trained coders categorized children's actions as Instrument actions (the child performed the action on the target animal using the target instrument), Modifier actions (the child performed the action on the target animal using her hand), or Mini-instrument actions (the child used the mini-instrument attached to the target animal to carry out the action). Ten trials coded as Mini-instrument actions were removed from the eye-movement analysis as in Snedeker & Trueswell, 2004). Two additional trials were excluded from analysis because the child performed the action on the distracter animal in error (1) or because the child's body blocked the coder's view of the action, which prevented accurate action coding (1).

Coders first marked the onset of each critical sentence and then children's eye movements were coded frame by frame from video, played back without sound. Coders recorded where the child was looking from the onset of the critical sentence until 3.5 seconds later, by which time most children had carried out an action (average critical sentence length was 2.1 s). We coded looks to the four quadrants of the podium, away, and to the central camera. Frames were coded as missing if the child's eyes were hidden. Four trials were removed from analysis, because children made fixations to a quadrant of the podium in fewer than 1/3 of frames between the onset of the sentence and the average onset of the action (87 frames). As a result, of 256 possible trials, we eliminated a total of 16 trials (6%) from the eye-movement analyses. Six children were replaced in the design because they contributed too few usable trials. Coding reliability was assessed for 16 children and yielded agreement on 96% of coded video frames.

2.2.1.6 Analysis

Data in this experiment and the subsequent experiments in this chapter were analyzed with mixed-effect models using the lme4 package (Bates & Sarkar, 2011) in R (R Development Core Team, 2011). In action data analysis, whether children did instrument action or modifier action in each trial was treated as a binomial dependent variable. In eye-movement analyses, we calculated the proportion of visual fixations at each point that were directed at particular objects (e.g., Target Animal, Target Instrument), and these proportions were transformed using an empirical-logit function (Barr, 2008; Jaeger, 2008). Participant and item were entered as random effects on the intercept (Baayen, Davidson, & Bates, 2008; Barr, 2008; Jaeger, 2008); models with random slopes were attempted but not included in the final models because they did not improve the model fit. The significance of predictors in the fitted model was tested with likelihood ratio tests based on comparisons of models with and without each fixed effect (Agresti, 2007; Barr, Levy, Scheepers, & Tily, submitted; Jaeger, 2008).

2.2.2 Results and Discussion

2.2.2.1 Behavioral measures

Across all trials in the experiment, children carried out instrument actions on 54% of trials; this value is similar to the mean proportion of instrument actions for equi-bias verbs in Snedeker and Trueswell's (2004) experiment. This suggests that children interpreted the verbs as did their peers in that earlier experiment, as equi-bias verbs with respect to the prepositional phrase attachment ambiguity. However, children tested on instrument-trained verbs also tended to enact more instrument actions ($M = 0.60$, $SE = 0.06$) than did those tested on modifier-trained

verbs ($M = 0.47$, $SE = 0.06$). This difference was in the right direction to indicate an influence of the training dialogues, but was not significant ($\chi^2(1) = 2.49$, $p = 0.11$). A significant vocabulary effect suggests children with higher vocabulary tended to conduct more modifier actions ($\chi^2(1) = 16.33$, $p < 0.001$), but vocabulary did not interact with training ($\chi^2(1) = 0.61$, $p > 0.44$). Analyses with sex and age added as predictors were attempted, but these factors did not significantly improve the models' prediction of children's actions.

2.2.2.2 Online eye-movement measures

Although all of our test verbs are really equi-bias verbs given the full range of children's experience, we asked whether children's eye movements as they listened to the test sentences would reveal the influence of recent experience in the training dialogues. If so, children should more often consider instrument interpretations of the ambiguous prepositional phrases for verbs experienced in instrument than in modifier dialogues. We first analyzed the proportion of trials in which children looked at the target instrument, following the coarse-grained analysis of eye movements in Snedeker and Trueswell (2004), training yielded no significant effect on this critical measure (Instrument *vs.* Modifier: 0.54 *vs.* 0.52, $\chi^2(1) < 0.1$, $p > 0.7$). However, we reasoned that fine-grained eye-movement pattern might reflect the influence of training both early in the sentence, before the ambiguous prepositional phrase, and in response to the ambiguous prepositional phrase. First, the verbs themselves might prompt children to expect instruments (or not); If so, then as children identified the verb and then the direct object noun in the test sentence, they might look more at possible instruments if the verb had appeared in instrument dialogues. In contrast, fixations to animals might dominate in these early regions of the sentence if the verb had appeared in modifier dialogues. Second, as children heard the object

of the ambiguous prepositional phrase, they should be more likely to interpret it as naming an instrument if the verb had appeared in instrument dialogues, or as a modifier if the verb had appeared in modifier dialogues. If so, then children who hear instrument-trained verbs should look longer at the target instrument and less at the target animal, than children who hear modifier-trained verbs. Thus we analyzed visual fixations in time windows anchored on the verb (e.g., *point*), the direct object noun (Noun-1, e.g. *pig* in *Point at the pig with the flower*), and the prepositional phrase object noun (Noun-2, *flower* in the same example). These three time views of the data are shown in Figure 2A-C. The zero-point of each plot's time scale is aligned with the onset of the relevant word.

Figure 2A shows the proportion of fixations to Animals (either the target or distracter animal) and Instruments (either target or distracter) in the Verb region. In this region we examined composite measures of looks to animals and instruments rather than looks to the target animal and target instrument because at this point in the sentence children have not yet heard the names of the target animal or target instrument. As the Figure 2A shows, in the Verb region, looks toward possible Instruments increased more after verb onset in instrument-trained than in modifier-trained trials. We analyzed the proportion of fixations to animals and instruments in a 700-ms window starting 200 msec after verb onset and ending 167 msec after Noun-1 onset. Analysis windows for eye movement in the current experiment and all the subsequent experiments were offset by 200 msec to allow time to program an eye movement (Hallett, 1986). The analysis revealed a significant effect of training dialogue on looks to instruments ($\chi^2(1) = 4.64, p = 0.03$), indicating that children started to expect a potential instrument soon after hearing an instrument-trained but not a modifier-trained verb. Children's looks to animals were not affected by training during the Verb region ($\chi^2(1) = 0.64, p = 0.42$).

Figure 2B shows the proportion of fixations to animals and instruments as the target animal name unfolded during the Noun-1 region (e.g. *pig*). Looks toward animals became dominant more quickly in sentences with modifier-trained rather than instrument-trained verbs. In sentences with instrument-trained verbs, children distributed their attention more equally between animals and instruments. We analyzed these fixation patterns in a 500-ms time window starting 200 msec after Noun-1 onset until 167 ms after the earliest onset of Noun-2 (the prepositional object). Children looked more to animals if they heard modifier- rather than instrument-trained verbs ($\chi^2(1) = 3.46, p = 0.06$), and more to instruments if they heard instrument- rather than modifier-trained verbs ($\chi^2(1) = 13.61, p < 0.001$). Though not shown in Figure 2B, fixations to the target animal were also affected by training in a similar pattern as fixation to either animal ($\chi^2(1) = 3.11, p = 0.08$).

Figure 2C shows the proportion of fixations to the target animal (TA) and target instrument (TI) in the Noun-2 window, anchored on the onset of the prepositional object (*with the flower*). Snedeker and Trueswell (2004) analyzed fixations in a 1-s period beginning at the onset of this word, divided into an early and a late analysis window. The average length of our recorded sentences was approximately 1.2 times longer than those used by Snedeker and Trueswell; thus we expanded the analysis window from a 1-s period to a 1.2-s period. The Noun-2 analysis window began 200 msec after Noun-2 onset, and was divided into two 600-msec analysis windows, the Early and Late Noun-2 windows shown in Figure 2C. In the Early Noun-2 window, we found a reliable training effect on looks to the target animal ($\chi^2(1) = 4.25, p = 0.04$). Children tended to look more at the target animal in trials with a modifier- rather than an instrument-trained verb. The training dialogues did not affect children's fixations to the target instrument in either the Early or Late Noun-2 window ($\chi^2(1) < 0.44, p > 0.51$).

The main effect of dialogue training on children's visual fixations to animals and instruments (again using the composite measures) was verified by a mixed-effect analysis with two levels of the between-subjects training factor (tested on instrument vs. modifier-trained verbs) and four levels of a within-subjects sentence position factor (Verb, Noun-1, early Noun-2, late Noun-2) (Quené & van den Bergh, 2004). Participants' fixations to animals and instruments changed across the sentence, yielding a main effect of sentence position (χ^2 's (3) > 52.0, p 's < 0.0001). There was also a marginal effect of training on fixations to animals (χ^2 (1) = 2.83, p = 0.09) and a significant effect of training on fixations to instruments (χ^2 (1) = 4.78, p = 0.03). Sentence position does not interact with training effects on looks to either animals or instruments (χ^2 's (3) < 3.23, p 's > 0.36).

Why did the training effect appear so early in the sentence, even before the ambiguous *with* phrase, and decay rapidly in the Noun-2 late window? Did children anticipate instruments for instrument-trained verbs as soon as they heard the verb? Or might this early effect be carried by later trials within the experiment, if children learned to anticipate a *with* phrase as the experiment progressed?

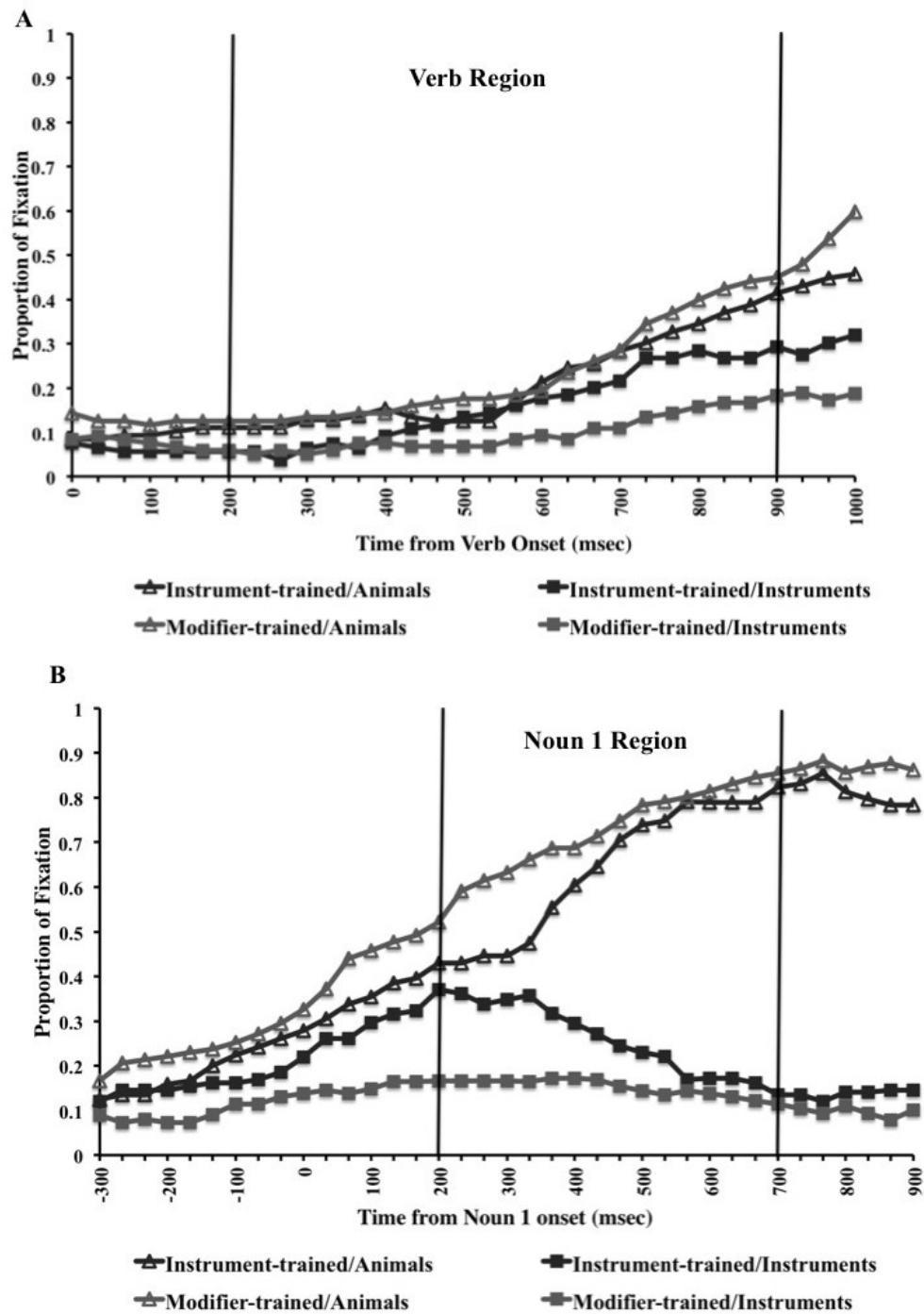


Figure 2 (cont. on next page)

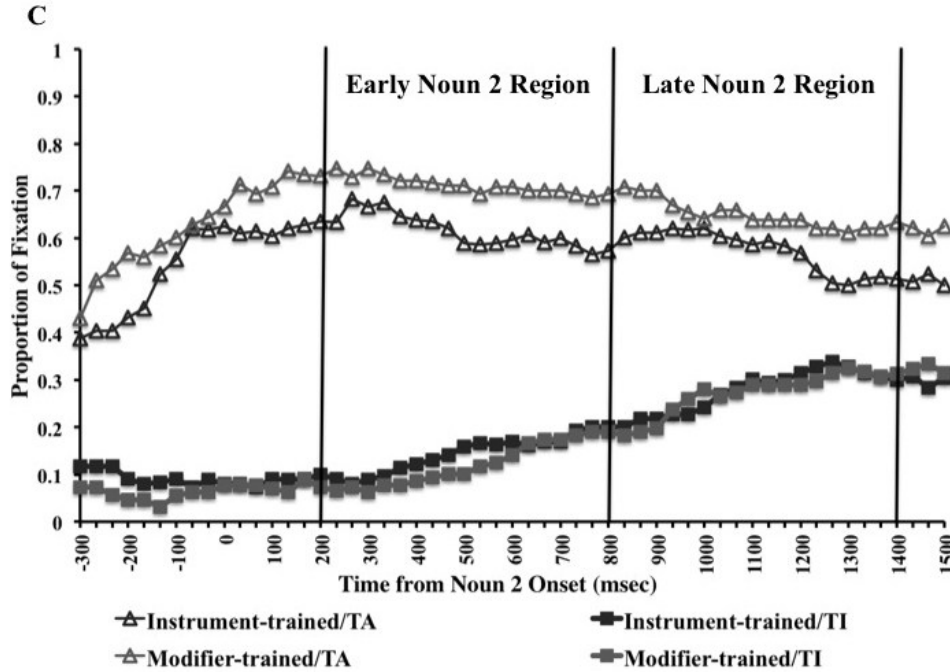


Figure 2: Experiment 1 eye-movement results. 5-year-olds' proportion of fixation to objects, time-locked at the onset of the verb (A), direct object noun (B) and the prepositional phrase object noun (C). Panel A and B plots the composite looks to animals and instruments respectively. Panel C plots the fixation to target animal and target instrument. The dark grey lines represent test trials containing instrument-trained verbs. The light grey lines represent test trials containing modifier-trained verbs. The vertical lines in each plot define the width of the analysis window. TA: target animal; TI: target instrument.

Separate analyses of the first and second halves of the test session were conducted to examine these questions. Figure 3 shows the proportion of fixations to animals (Figure 3A) and instruments (Figure 3B) across all the analysis windows, for instrument-trained and modifier-trained verbs. Note in the figure that in both halves of the experiment, children looked more at instruments and less at animals in trials with instrument- as opposed to modifier-trained verbs. Earlier in the experiment, this training effect was most prominent at the end of the sentence, in the Noun-2 region (e.g. *flower*). Mixed effects analyses like those reported above conducted on data from the first half of the experiment revealed a significant effect of training only in the early Noun-2 region, both in an analysis of fixations to animals ($\chi^2(1) = 5.97, p = 0.01$) and to

instruments ($\chi^2(1) = 4.65, p = 0.03$). Though not shown separately on the graph, the training effect on fixations to animals was driven mainly by looks to target animals in the Noun-2 region. Children looked more at the target animal in trials with modifier-trained verbs than in trials with instrument-trained verbs, in both the early Noun-2 ($\chi^2(1) = 12.14, p < 0.001$) and the late Noun-2 regions ($\chi^2(1) = 5.65, p = 0.02$). Children's looks to the target instrument followed a reversed pattern, though yielding a non-significant training effect ($\chi^2(1) = 2.20, p = 0.14$). In contrast, no reliable training effect was found in either the Verb region or the Noun-1 region during the first half of the experiment ($\chi^2(1) < 2.52, p > 0.11$). Note in Figure 3 that the second half of the experiment saw a shift of the training effect to earlier regions in the sentence. In the Verb region, children's looks to the animals did not differ between the instrument and modifier-trained verbs ($\chi^2(1) = 0.43, p > 0.51$). However, children looked more at the instruments after hearing an instrument-trained verb than after hearing a modifier-trained verb ($\chi^2(1) = 5.72, p = 0.02$). Analyses of fixations in the Noun-1 region in the second half of the experiment revealed a significant training effect on both fixations to animals ($\chi^2(1) = 3.74, p = 0.05$) and to instruments ($\chi^2(1) = 6.54, p = 0.01$). These effects disappeared in the Noun-2 region in the second half of the experiment ($\chi^2(1) < 0.41, p > 0.52$). The shift of the training effect from later to earlier sentence positions across the experiment was further supported by the fact that the test order (first half vs. second half) significantly interacts with the training effect on target animal fixations in the early Noun-2 region ($\chi^2(1) = 8.08, p = 0.004$) and marginally interacts with the training effect on target fixations in the late Noun-2 region ($\chi^2(1) = 3.28, p = 0.07$).

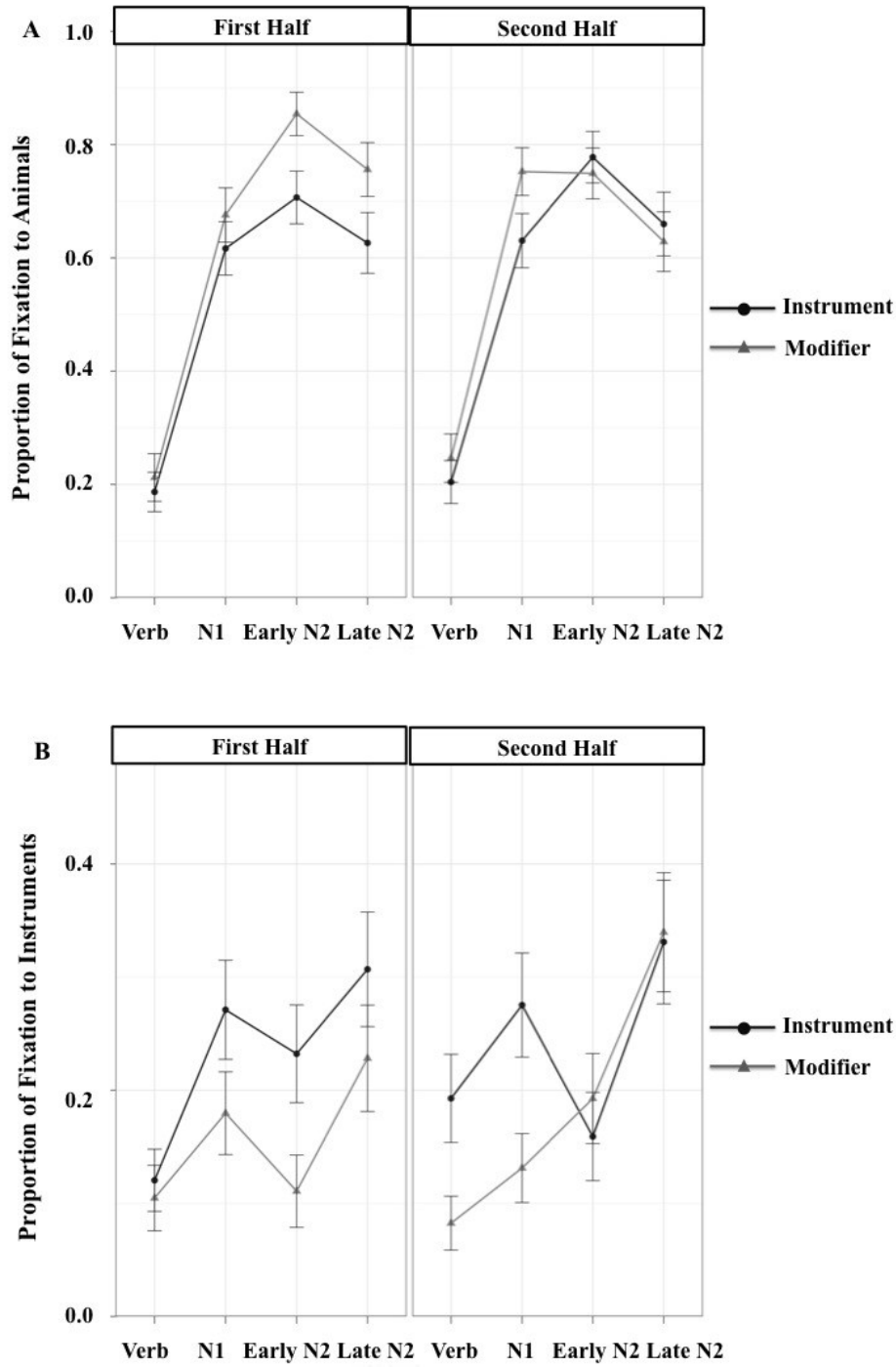


Figure 3: Experiment 1 eye-movement results by test orders. Mean proportion of fixation to Animals (A) and Instruments (B) across four analysis windows during the first half and the second half of the experiment. The training effect shifted earlier across the experiment trials. N1: direct object noun (Noun-1); N2: prepositional object noun (Noun-2). Error bars represent the standard errors.

This pattern suggests that the newly learned verb bias plays a role not only in children's ambiguity resolution, but also in guiding children's anticipation of the upcoming words in the sentence. In the first half of the trials, reliable effects of training emerged as children heard the object of prepositional phrase. In the second half of the trials, after hearing some test trials, children started to anticipate an ambiguous *with* phrase and thus retrieved the trained verb bias soon after they heard the verb. However, despite the earlier retrieval of the learned verb bias, the training effect declined during the later PP-noun regions, which might reflect children's awareness of the possible alternative interpretation for the ambiguous sentences. Children's awareness of the ambiguity also went consistently with the overall equi-bias pattern in their action data.

2.2.3 Summary

The findings show that brief exposure to sentences was sufficient to modify the structural biases of familiar verbs in 5-year-olds. All children heard both instrument and modifier dialogues in the study phase. Their consideration of instrument versus modifier interpretations of *with* phrases in the test phase depended on which verbs appeared in the test sentences – those they had heard in instrument or in modifier dialogues. Thus, during the study phase, children attached to each verb information about its occurrence with instrument vs. modifier phrases in the dialogues. At test, children retrieved this combinatorial information when they encountered the same verbs again. As a result, the linguistic information in the training dialogues influenced children's online processing of sentences with ambiguous *with* phrases. These findings demonstrate that language system of 5-year-olds can dynamically adapt to the updated verb-specific combinatorial

information from language input and make rapid use of the information during online language processing.

2.3 Experiment 2: Malleability of familiar verb bias in adults

In Experiment 2 we tested adults in a similar task to further explore the plasticity of the language processing system's use of verb bias. Evidence has suggested adults can rapidly update their structural expectations during comprehension depending on their recent experiences with particular syntactic structures, also known as syntactic priming effect on comprehension (Chang et al., 2006; Thothathiri & Snedeker, 2008; Tooley, Traxler, & Swaab, 2009). In addition, comprehenders appeared to be sensitive to the probabilistic cues of certain syntactic structures over the course of an experiment (within hours or days) reflected by their shifted parsing strategy (Fine & Jaeger, 2011; Wells et al., 2009). Here we ask whether, the rapid update of information can be associated with specific verbs. Even after 20 years of experience hearing familiar verbs, can a short listening experience with particular verbs still dynamically modulate people's verb bias? If so, this experiment would provide strong evidence for the continuous influence of linguistic distributional info on verb bias acquisition.

2.3.1 Methods

2.3.1.1 Participants

Thirty-two adults, all native speakers of English (18-22 years old, 14 females and 18 males) participated in this study. They were recruited from the University of Illinois student community, and received partial course credit or a small payment in exchange for their

participation. Eight additional subjects participated but were not included in the analysis because they contributed too few coded trials due to excessive central looking¹.

2.3.1.2 Materials and Procedures

The procedure and stimuli were almost identical to those of Experiment 1, with the following exceptions. First, in the study phase, in addition to the 4 instrument-training dialogues and 4 modifier-training dialogues, each participant also heard 8 filler dialogues intermixed with the critical training dialogues in a random order. Second, during the test phase, the adults received all 8 target trials, blocked by their training condition; recall that the children received only 4 target trials, all from the same training condition. Within each block, the target trials were pseudo-randomly mixed with 12 filler trials. Therefore, like the 5-year-olds in Experiment 1, the adults in Experiment 2 heard both types of training dialogues. For the adults, however, the verb-bias training condition in the test phase was manipulated within subjects rather than between subjects.

Finally, at the end of the experiment participants completed a survey asking about their impressions of the task, including their awareness of the ambiguity of the stimulus sentences, their behavioral strategy during the test phase, and a sentence completion task designed to probe their memory for the dialogue sentences. All subjects noticed the target sentences could be resolved using either the instrument or the modifier interpretation. 28 out 32 participants claimed they were not attempting to recall the training dialogues in the test phase. In the sentence completion task, participants received sentence stems from the dialogue videos containing the

¹ These participants appeared to interpret the initial fixation instruction of each trial ("Look at the camera") as an instruction to look only at the camera during the sentences.

eight critical verbs (e.g. *Point at the tiger with _____*), and were asked to complete the stems with whatever phrase first came to mind. Because these sentences included nouns from the dialogue videos, we reasoned that this task might prompt them to retrieve the prepositional objects from the dialogues (*Point at the tiger with the magic sword* versus *with the large paws*, depending on training condition for this verb).

2.3.1.3 Coding

Participants' actions and eye-movement data were coded as in Experiment 1. Adults rarely enacted mini-instrument actions (5 out of 256 trials), but they tended to fixate more frequently at the central camera. 31 trials with fewer than 1/3 of the frames coded as fixations at any quadrant of the podium were removed from analysis; thus 36 trials (14%) were eliminated from the eye movement analyses. Coding reliability was assessed for 8 subjects and yielded 94% agreement.

Responses in the sentence completion task were categorized into three types: instrument completions (with noun choices the coders judged could only serve as instruments for the action, not as attributes of the direct object noun), modifier completions (those judged as clearly describing an attribute of the direct object), and ambiguous completions (those judged to permit either reading; these included completions incorporating words from the test phase sentences, such as *with the flower*). Five participants did not complete the sentence completion task.

2.3.2 Results and Discussion

2.3.2.1 Behavioral measures

The adult participants performed instrument actions on 23% of trials; adults made fewer instrument responses (thus more modifier responses) than did the children in Experiment 1. There was no reliable difference in the proportion of instrument actions between the instrument and modifier training conditions (0.24 vs. 0.21, $\chi^2(1) = 0.42, p > 0.50$)².

In the post-experiment sentence completion task, as shown in Table 1, the dialogue training very strongly influenced the types of endings participants used to complete the sentences (paired t-test: $t(27) = 8.20, p < .001$). These data clearly show that the participants remembered the dialogue sentences. Upon hearing the verb and direct object noun (e.g., *point at the tiger with*) from a dialogue sentence, participants strongly tended to produce a completion that resembled the original training sentences. Note, however, that this large effect of training on sentence recall was not mirrored by a large effect of training on the final interpretations of the ambiguous test sentences. This difference suggests that participants did not explicitly retrieve the dialogue sentences and use them to resolve the ambiguity in the test phase.

² Within female participants, the size of the training effect was similar as was observed in Experiment 1 (0.34 vs. 0.20, $\chi^2(1) = 2.97, p = 0.08$). The interaction between gender and training was marginal ($\chi^2(1) = 2.71, p = 0.10$).

Table 1: Mean proportions of instrument actions performed in the test phase, and instrument completions produced in the post-experiment sentence completion task, Experiment 2. Standard errors are in parenthesis.

Training Type	Proportion of Instrument Actions	Proportion of Instrument Completions
Instrument	0.24 (0.05)	0.70 (0.06)
Modifier	0.21 (0.04)	0.10 (0.04)

2.3.2.2 Online eye-movement measures

Despite the overall preference for modifier actions, coarse-grained analysis of eye-movements showed evidence for implicit retrieval of newly updated verb bias information. For each trial, we determined whether the participant looked at the Target Instrument from the onset of the sentence till the start of the action. Figure 4 shows the proportion of trials with an instrument fixation by training condition. Participants tended to look at the target instrument more when the sentence contained an instrument-trained verb, than when it contained a modifier-trained verb ($\chi^2(1) = 3.13, p = 0.08$)³.

³ Again, this effect is only reliable in female participants ($\chi^2(1) = 4.57, p = 0.03$). The non-significant training effect in male participants was in the same direction.

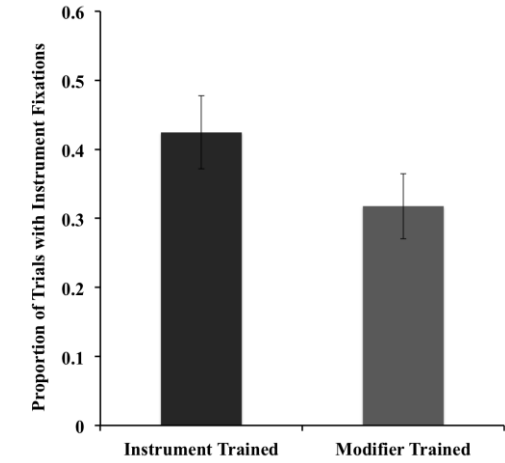


Figure 4: Experiment 2 coarse fixation results. Mean proportion of trials in which participants fixated on the target instrument from the onset of the sentence to the beginning of the action.

We analyzed the fine-grained eye-movement patterns in the same time windows defined in Experiment 1. Figure 5 plots eye fixation to objects during Verb region, Noun-1 (e.g. *Pig*) region, and Noun-2 (e.g. *flower*) region. In both the Verb (Figure 5A) and Noun-1 (Figure 5B) regions, adults' fixation pattern was not differentiable between the two training conditions ($\chi^2(1) < 1.80, p > 0.18$). However, as in Experiment 1, we again found that the training effect in the Verb region grew more prominent in the second half than the first half of the experiment. Figure 6 plots the mean proportion of fixations to instruments in the first half and the second half of the experiment. Analysis revealed a significant interaction between the testing order (first half vs. second half) and the training factor ($\chi^2(1) = 4.23, p = 0.04$). Training affected adults' anticipatory fixation to instruments only in the second half of the experiment ($\chi^2(1) = 6.02, p = 0.01$).

In the Noun-2 region (Figure 5C), fixations to target instruments were significantly affected by training. During the early Noun-2 region, adults looked more at the target instrument when hearing the prepositional object nouns after an instrument-trained verb than when hearing it after a modifier-trained verb ($\chi^2(1) = 4.82, p = 0.03$). Unlike children, the training effect

during the early Noun-2 region did not decline in the second half of the experiment, which may be due to the higher proportion of filler items. In the late Noun-2 region, training did not affect adults' looks to either the target animal or the target instrument ($\chi^2(1) < 0.21, ps > 0.65$).

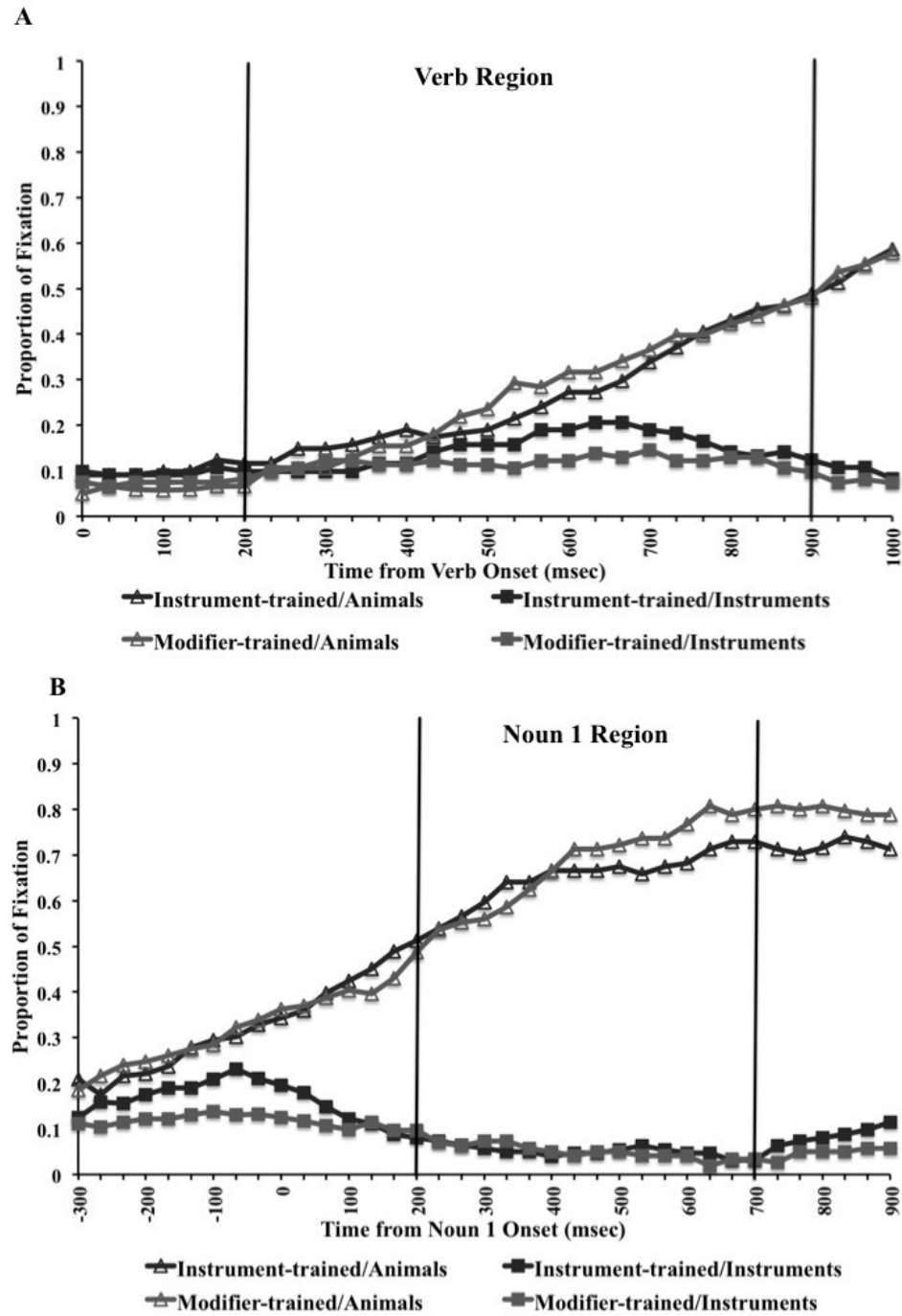


Figure 5 (cont. on next page)

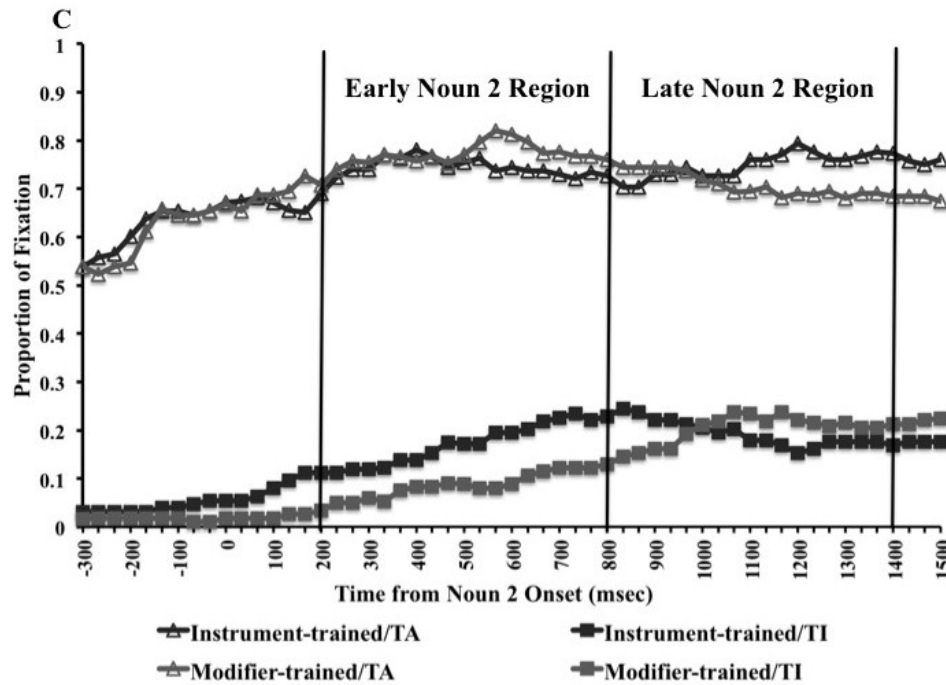


Figure 5: Eye-movement patterns in Experiment 2. Adults' proportion of fixations to objects in the display, time-locked at the onset of the verb (A), direct object noun (B) and prepositional phrase object noun (C). The dark grey lines represent test trials containing instrument-trained verbs. The light grey lines represent test trials containing modifier-trained verbs. The vertical lines in each plot define the width of the analysis window. TA: target animal; TI: target instrument.

In sum, Experiment 2 replicated with adults the main findings of Experiment 1, indicating even with 20 years of linguistic experience, adults' language-processing system remains highly sensitive to distributional information in the language input. Adults attached to each individual verb information about its co-occurrence with modifier- or -instrument *with* prepositional phrases, and retrieved that information when encountering the same verb in an ambiguous test sentence. The early training effect at the Verb region emerged in the second half of the experiment, suggesting that the verb alone could activate the retrieval of the co-occurrence frequency of the individual verb and the *with* structure, once adults became more experienced with the experimental sentences. During the second half of the experiment, the retrieved

information of verb bias guided participants' anticipation of the upcoming words and structures in the sentence. Note that the trained bias in the second half was different from that in the first half for each participant. Therefore the shifting-early phenomenon only reflected possible increase of anticipation of an ambiguous *with* phrase, but not simply perseveration.

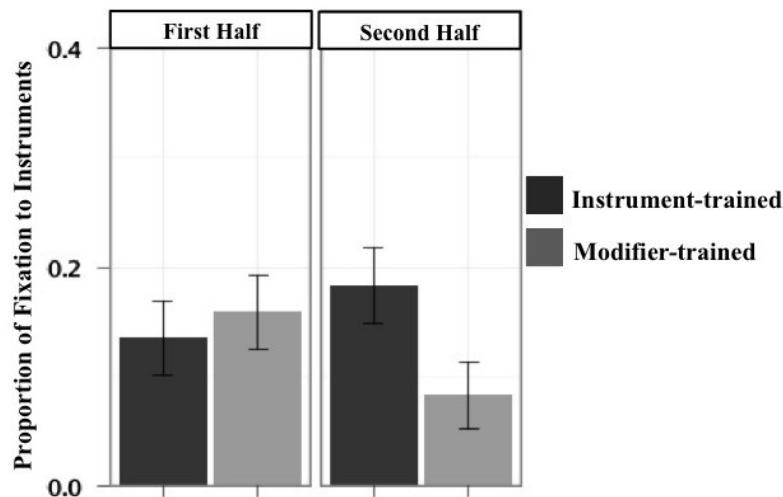


Figure 6: Experiment 2 eye-movement results by test order. Mean proportion of fixations to Instruments during the Verb region in the first and second half of the experiment. The training effect became greater in the second half of the experiment. Error bars represent the standard errors.

2.3.3 Summary

The results of Experiment 2 provide new evidence that the human language-processing system can continuously acquire new evidence about the biases of familiar verbs based on language exposure, and rapidly integrate this updated verb bias information during online sentence comprehension. Both children and adults showed at best a weak tendency for their final actions to be influenced by the dialogue training. Interestingly, in the adult data of Experiment 2, female participants seemed to be more sensitive than males to our training manipulation. This effect resembles the sex differences in sensitivity to linguistic information (Baxter et al., 2003;

Kansaku, Yamaura, & Kitazawa, 2000; Shaywitz et al., 1995) that are sometimes found (though these are controversial and hard to characterize); in our case, of course, the sex differences may simply reflect the tendency of college-aged women to have a bit more patience with our child-friendly dialogues and test sentences than did college-aged men. The substantial training effect on sentence completion task makes it clear that adults are able to memorize the statistical pattern by rote, but retrieval of such information during natural language processing is automatic and implicit.

Taken together, in Experiment 1 and Experiment 2, children and adults attached to each verb new information about its frequency of occurrence with an instrument or modifier *with* phrase. They later retrieved that information when they encountered the same verb again, and used it to guide both online prediction and ambiguity resolution. The findings suggest that new linguistic experience involving known verbs can reshape listeners' knowledge about their structural biases.

However, the training effects found in these two experiments might reflect contributions from both linguistic-distributional and event-distributional knowledge, due to the semantic richness of the familiar verbs. Although the events described by the dialogue sentences were not visually depicted, the linguistic descriptions conveyed event information that might have induced subtle changes in verb sense. For example, upon hearing *point at the tiger with the magic sword / red pencil* in a context inviting an instrument interpretation, children and adults presumably generated a mental representation of the event described. Thus the dialogues simultaneously gave children information about (1) the linguistic distributional information about how likely a verb-attached *with* prepositional phrase co-occurred with this verb, and (2) the event distributional information about how likely an instrument could be used in a pointing event. The

latter was carried by the semantics of the verb, the semantic overlap between the verb and its possible instruments (e.g. *sword* is *pointy*) as well as the semantic category that can be generalized from the similarity between the PP-object nouns (both *pencil* and *sword* are plausible instruments for *pointing*). Therefore, at the time listeners learned about the linguistic distributional information, they might also have encoded the event-distributional information, which later contributed into their online ambiguity resolution.

The difficulty of disentangling event and linguistic distributional information follows from the nature of linguistic communication. On one hand, when we identify a verb, we activate a complex of linguistic and event knowledge. This includes information about the possible syntactic complements of the verb, the participants involved in the event, and the mapping between the two (e.g., Carlson & Tanenhaus, 1988). In adults, presentation of a verb causes readers or listeners to activate detailed information about the participant-roles that are filled by that verb's arguments (e.g., Boland, 2005; Ferretti et al., 2001). On the other hand, the linguistic and event contexts of verb use in people's linguistic experiences are closely related. Toddlers are able to use the structural contexts of verb as a cue for acquiring the meaning of events delivered by the verb (Fisher, Gertner, Scott, & Yuan, 2010; Scott & Fisher, 2009; Yuan & Fisher, 2009). For example, 2-year-olds were able to learn the combinatorial information for a nonsense verb and a transitive structure and later retrieve such information in mapping the unknown verb to a 2-participant event. Therefore, the training effect observed in Experiment 1 and 2 could result from the strong interaction between the linguistic and event information sources. In Experiment 3, we tried to further dissociate the two sources by examining the necessity of the matching sentence structures between study and test phases for verb bias learning.

2.4 Experiment 3: Verb bias learning without linguistic distributional information

In the current experiment, we removed the co-occurrence information of the verb and the *with* PP-phrases from training, but kept the same event information conveyed by the dialogues. For example: instead of *point at the tiger with the magic sword* in an instrument training dialogue, children heard *point at the tiger using the magic sword*. This manipulation permits a further uncoupling of linguistic- and event-distributional learning. If the training effects found here are eliminated by this syntactic difference between the dialogue and test sentences, then this would provide stronger evidence of a central role for linguistic-distributional learning in the acquisition of verb bias. If verb bias knowledge results primarily from knowledge of verb meanings and the events verbs refer to, then we would expect the results of Experiment 3 to be just like Experiment 1. If both the linguistic- and event-distributional information contributes in verb bias learning, then we would expect a reduced training effect but in a similar pattern as observed in Experiment 1.

2.4.1 Methods

2.4.1.1 Participants

Sixty-four 5-year-olds (mean 5; 1 months, range 4; 6 - 5; 10; 32 boys, 32 girls) participated. Eight additional children were eliminated from analysis, because they declined to participate (3), contributed too few codable trials (2); or showed evidence of misunderstanding the study phase dialogues (3). Participants in the current experiment were similar in age and vocabulary as in those in Experiment 1.

2.4.1.2 Materials and Procedure

During the study phase, the *with* phrases in the training dialogues from Experiment 1 were replaced by *using* or *that has* to convey the same event semantics. For example, (5a) and (5b) show the two training dialogues for the verb *point*. As in the previous two experiments, after each dialogue, the experimenter asked the child to repeat the last sentence of the dialogue (“What did she just say”, while pointing at the last speaker).

(5a) Instrument-training dialogue

A: “Do you remember that story about the tiger? What did Tim use to point at the tiger?”

B: “He pointed at the tiger using the red pencil.”

A: “Right, he pointed at the tiger using the red pencil. Hmm. If I were him, I would point at the tiger using the magic sword.”

B: “Wow! You wanna point at the tiger using the magic sword! How exciting!”

(5b) Modifier-training dialogue

A: “Do you remember that story about the tiger? Which tiger did Tim point at?”

B: “He pointed at the tiger that has the large paws.”

A: “Right, he pointed at the tiger that has the large paws. Hmm. If I were him, I would point at the tiger that has the sharp teeth.”

B: “Wow! You wanna point at the tiger that has the sharp teeth! How exciting!”

Children’s attempts to repeat the final dialogue sentences were recorded and transcribed. Interestingly, children often provided responses such as “*point at the tiger with the sword*” or “*the tiger with the sharp teeth*”, in which they replaced the unambiguous *using* or *that has* phrases of the dialogues containing *with* prepositional phrases like those presented in the dialogues of Experiments 1 and 2. Children were about equally likely to produce *with* phrases

after instrument (*using*) and modifier (*that has*) training dialogues (2.11 vs. 2.28 out of 4, respectively). Children's tendency to volunteer *with* as a suitable paraphrase for the modified dialogue sentences of Experiment 3 confirms that these training dialogues indeed conveyed the same meanings as did the dialogues in Experiment 1, as we intended. The first and the last dialogues in the study phase were fillers. We measured children's answer accuracy by counting how many times children completely repeated the trained verb together with either a correct instrument or a correct modifier phrase. In average, children are able to provide complete answers in 4.33 out of 8 dialogues. Comparing the current experiment with Experiment 1 (4.86 out of 8), we verified the training dialogues in two experiments were similar in their processing difficulty ($t(125) = 1.27, p = 0.21$).

The materials and procedures of the test phase were identical to Experiment 1. Thus, though children had viewed training dialogues for all 8 equi-bias verbs, in the test phase they received test trials for only 4 verbs, either those that they had experienced in instrument dialogues, or those they had experienced in modifier dialogues. Participants were randomly assigned to the instrument and modifier-testing group; the two groups of children did not differ from each other in age, vocabulary or the performance in answering the questions during the study session.

Each child also completed the Peabody Picture Vocabulary Test—Third Edition (PPVT-III) at the end of the study. Children's vocabulary was strongly correlated with the number of times they produced *with* during training ($R=0.39, p < 0.001$).

2.4.1.3 Coding

Participants' actions and eye movements were coded as in Experiment 1. Nine trials coded as mini-instrument actions were removed from the eye-movement analyses. Six additional trials were excluded from analysis because no action was performed (1), the action was performed on the distractor animal (2) or on the mini-instrument as if it was the direct object (2), or the child blocked the coder's view of the quadrant of podium she was acting on (1). Seven trials with fewer than 1/3 of frames coded as fixations to a quadrant of the podium were removed from analysis. In total, 22 trials (8.6%) were eliminated from the eye movement analyses. Coding reliability was assessed for 16 children and yielded agreement on 97% of coded video frames.

2.4.2 Results and Discussion

2.4.2.1 Behavioral measures

The dialogue training did not affect children's tendency to enact instrument actions (Instrument training vs. Modifier training: 0.52 vs. 0.47, $\chi^2(1) = 0.05$, $p > 0.80$). However, individual differences in vocabulary appeared to interact with children's response to training. Analysis with vocabulary entered as a continuous predictor revealed a reliable interaction between children's PPVT score and training ($\chi^2(1) = 4.47$, $p = 0.03$). Figure 7 shows the proportion of instrument actions in each training condition for children whose raw PPVT scores fell above (high vocabulary) versus below (low vocabulary) the median score for the group. Children with higher vocabularies enacted more instrument actions for instrument-trained verbs than for modifier-trained verbs. An analysis with sex, age and the number of times children

produced *with* during training was attempted but adding these factors did not improve the model's predictions of children's actions.

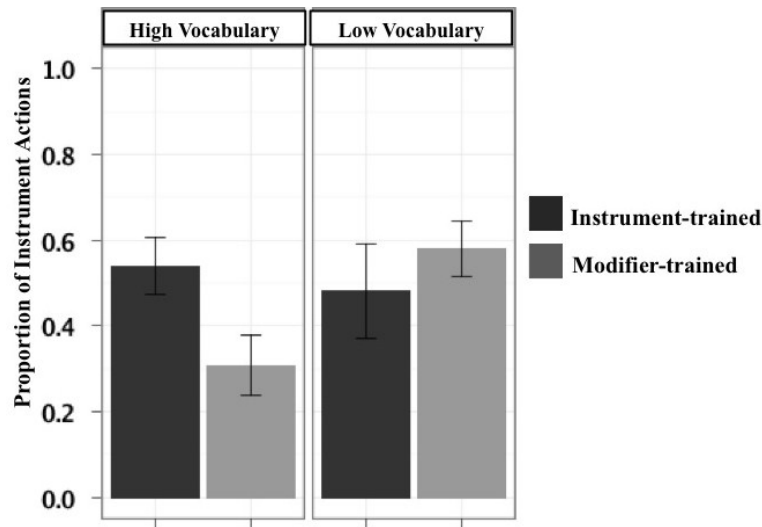


Figure 7: Experiment 3 action results. 5-year-olds' proportion of instrument actions by median-split vocabulary groups. Children with higher vocabulary showed larger sensitivity to training.

2.4.2.2. Online eye-movement measures

Children's eye-movement patterns were analyzed in the same time windows defined in Experiment 1. Figure 8 plots the fixation proportions to objects during Verb, Noun-1 (e.g. *pig*) region and Noun-2 (e.g. *flower*) region. Training effect did not reliably influence children's fixation in either Verb region or the Noun-1 region ($\chi^2(1) < 2.3$, $ps > 0.13$). As a matter of fact, the visually largest difference between two training conditions shown in the Noun-1 region (Figure 8B) appeared to be in the opposite direction, that children looked more at the animals in the trials containing instrument-trained verbs than in those containing modifier-trained verbs.

During the early Noun-2 region (Figure 8C), children looked marginally more at the target instrument upon hearing Noun-2 in an instrument-trained trial, as opposed to when hearing Noun-2 in a modifier-trained trial ($\chi^2(1) = 3.10$, $p = 0.08$). Vocabulary did not interact with

training ($\chi^2(1) = 0.59, p > 0.44$). No training effect was revealed in the late Noun-2 region on children's looks to either the target animal or the target instrument ($\chi^2's(1) < 0.88, p's > 0.35$). However, as Figure 9 shows, high-vocabulary children showed sensitivity to training if examined in a wider measurement window. Combining both the early and late Noun-2 regions, high-vocabulary children looked reliably more to the target instrument in the instrument-trained trials than in the modifier-trained trials ($\chi^2(1) = 4.18, p = 0.04$). No such effect was observed in the low-vocabulary group ($\chi^2(1) = 0.51, p = 0.48$).

Despite the fact that the marginal training effect at the early Noun-2 region resembles the training effect in Experiment 1, analyses with training and test order (first and second half) as factors revealed no difference of the training effect between two halves of the experiment on children's looks to target animal and target instrument in the early Noun-2 region ($\chi^2's(1) < 0.45, p's > 0.50$). The training effect on children's looks to animals and instruments also did not emerge early in either the Verb or Noun-1 region in the second half of the experiment ($\chi^2's(1) < 2.44, p's > 0.11$).

Finally, we conducted a mixed-effect analysis with training condition as a between-subjects factor (instrument vs. modifier), and the 4 sentence positions (Verb, Noun-1, Early Noun-2 and Late Noun-2) as a within-subjects factor. Participants' fixations to animals and instruments changed across the sentence, yielding a main effect of sentence position ($\chi^2's(3) > 62.7, p's < 0.0001$). Training did not affect children's overall fixation pattern ($\chi^2's(1) < 0.71, p's > 0.4$).

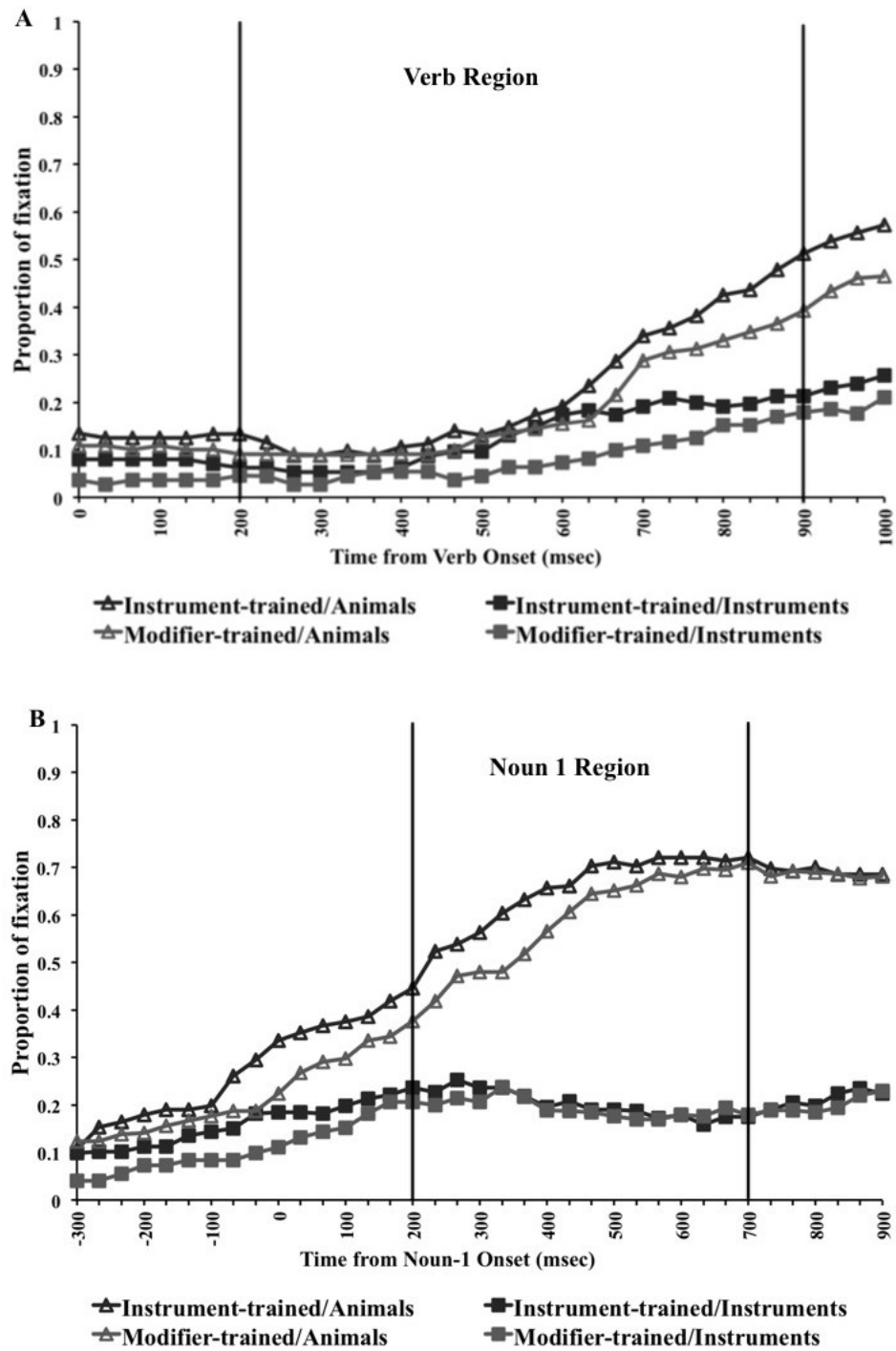


Figure 8 (cont. on next page)

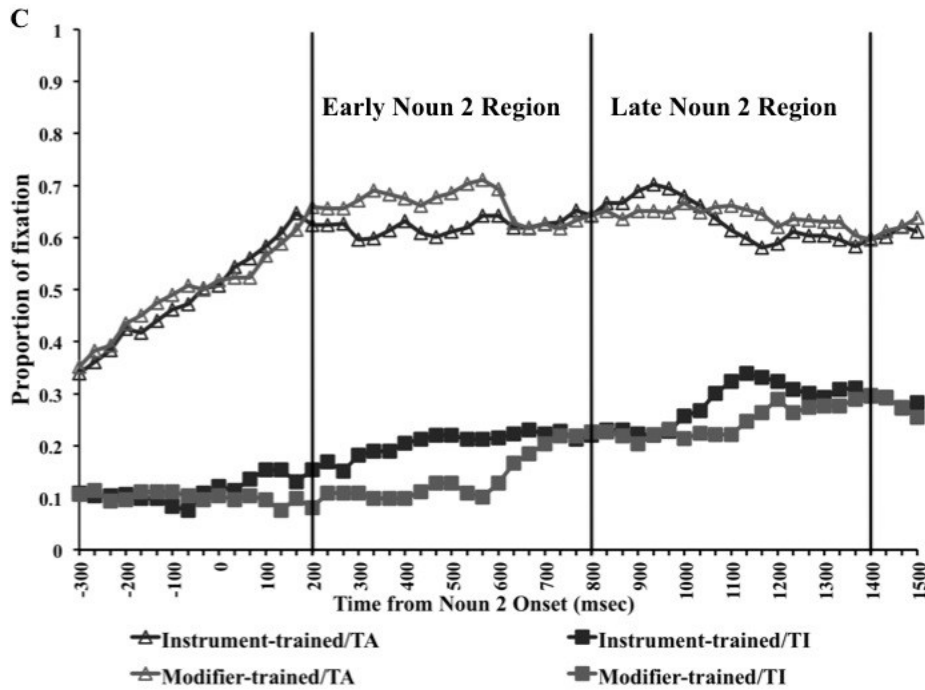


Figure 8: Experiment 3 eye-movement results. 5-year-olds' proportion of fixation to objects, time-locked at the onset of the verb (A), direct object noun (B) and the prepositional phrase object noun (C). The dark grey lines represent test trials containing instrument-trained verbs. The light grey lines represent test trials containing modifier-trained verbs. The vertical lines in each plot define the width of the analysis window. TA: target animal; TI: target instrument.

The overall training effect on children's online eye-movement pattern in the current experiment resembles that of Experiment 1, indicating that the event-distributional information plays a role in updating children's bias of familiar verbs. However, the effect was much weaker and only emerged marginally during Early Noun-2 region. In addition only high-vocabulary children showed significant training effect in a wider Noun-2 analysis window, suggesting retrieving newly learned verb bias was more difficult without the structural co-occurrence information from training. However, given the fact that children used *with* ambiguous phrases to paraphrase training sentences in more than 50% of the dialogues, it is likely that at least some children were able to transfer parts of the event distributional information into linguistic

distributional information, and then retrieve the latter upon hearing the same verb. The results of vocabulary effect went along with such speculation. Only high-vocabulary children were able to show sensitivity to training manipulation in both their actions and eye-movement, suggesting learning verb bias from linguistic-conveyed event-distributional information is a demanding task, yielding in the maturation of one's language skills.

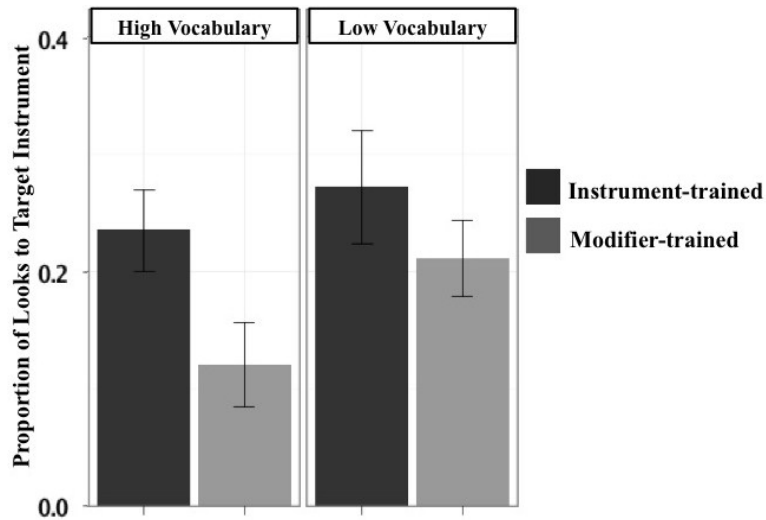


Figure 9: Experiment 3 eye-movement results by vocabulary. Mean proportion of fixation to target instrument during the 1-s Noun-2 region. High-vocabulary children looked more to the target instrument in instrument-trained trials and less in modifier-trained trials. Error bars represent standard errors.

2.4.3 Summary

The current experiment investigated the necessity of the linguistic-distributional information in verb-bias learning by presenting the same linguistically described events in sentences that involve neither the same syntactic structure nor the same key words as the later test sentences. That is, during the dialogues children heard sentences in which instruments and attributes were described without *with* (e.g., *pointing at a tiger using a magic sword*, vs *pointing at a tiger that has large paws*). These results suggest that the statistical information embedded in

the events (conveyed by language) plays a role in updating children's biases of familiar verbs, but in a much weaker fashion. Children considered the instrument interpretation marginally more for the instrument-trained verbs than for the modifier-trained verbs with large between-subject variability. Only high-vocabulary children showed clear signs of sensitivity to the training manipulation. The overall weaker training effects in the current experiment might be due to two potential factors. First, the discrepancy of the sentence structures between the study phase (i.e. *using* and *that has*) and the test phase (*with* prepositional phrases) may lead to an extra cost in retrieving event-distributional information without the support of structural similarity. If so, these results will be consistent with the view that linguistic distributional information plays a crucial role in applying the newly learned verb bias during online ambiguity resolution. As mentioned before, children with higher vocabulary tended to produce *with* ambiguous phrases more frequently when paraphrasing the training sentences, indicating they treated the sentences containing *with* structures as an equivalent semantic representation of the described events. Although the production data might reflect various other aspects of children's social and cognitive development, it is likely that high-vocabulary children encoded the verb-specific *with* structural preference simultaneously together with the event-distributional information.

Furthermore, even before reaching *with* in the test sentence, the training effect in the current experiment appeared to be weaker than that in Experiment 1. Children showed no evidence of anticipating an instrument soon after hearing an instrument-trained verb, suggesting the association between a verb and its preference for an instrument (or instrument structure) was not strong enough to activate children's anticipation of an instrument. Therefore, the second possible explanation for the weaker training effects might be due to the reduced learning efficacy in unambiguous sentences. Because *using* and *that has* serve as strong disambiguating cues in

the training sentences in addition to the discourse context, it was not necessary to use the verb to predict the outcome of the temporary across dialogue sentences under this circumstances. In addition, according to the expectation-driven account of sentence comprehension, the lack of anticipatory looks in the current experiment suggests compared to *using* and *that has* phrases, the *with*-PP phrases in Experiment 1 and 2 were learned more like arguments, which has a privileged status in driving listeners' expectation (Boland, 2005).

In summary, the findings in the current experiment suggest a similar but weaker learning effect in 5-year-olds, when they were only provided with linguistic-conveyed event-distributional information in the training. Learning without the support of linguistic-distributional information is more difficult and the learning efficacy depends on children's language skills, which may even involve the ability to transfer the event-distributional information to linguistic distributional information. One caveat for the current conclusion is that our measurement of the learning efficacy from online testing sentences is the outcome from the combination of both the encoding during the study phase and the retrieval during the test phase. It might also be the case that children learned the combinatorial facts between the verbs and the following nouns in the current experiment less efficiently.

The findings in the current experiment partly replicated our findings in Experiment 1, providing evidence that both linguistic- and event-distributional information contributes in verb bias learning. However, the weaker training effect and the lack of the anticipatory looks suggest verb bias learning with less support from linguistic distributional information becomes more difficult. Experiment 4 further teases apart the event information from linguistic distributional information by using a different tactic. We removed the verb semantic information, as well as the

semantic relatedness between a verb and its *with* PP-object nouns from the training dialogues, and ask whether verb bias learning is possible without the rich support of the verb semantics.

2.5 Experiment 4: Verb bias learning without verb semantics

The semantic meanings of familiar verbs in the previous three experiments delivered rich event information that was tightly intertwined with the linguistic distributional information in the training dialogues. In the current experiment, we modified the training dialogues in Experiment 1 by replacing the familiar verbs with invented verbs. For example, in an instrument training dialogue, children heard “*how did Tim veeb the tiger?*” “*he veebed the tiger with the magic sword...*”. Even with some event information from the discourse and the plausible instrument nouns, the *veebing* action in this context was still very vague. In the test phase, children heard “*flom the pig with the flower*”. The design of the current experiment allowed children to participate in the test trials without knowing what *flom* means. If children are able to learn the verb bias without event knowledge of verb semantics, we expect the similar results as Experiment 1. The results will provide even stronger evidence for the critical role of linguistic distributional information in verb bias learning. If children fail to learn verb bias for the nonsense verbs, we will not observe the effect of training on children’s online sentence processing. Then the results will suggest verb semantic acquisition might precede verb bias learning.

2.5.1 Methods

2.5.1.1 Participants

Thirty-two 5-year-olds ranging in age from 4; 6 to 5; 7 participated in the study (M : 4; 11 months; 15 girls and 17 boys). One additional child was eliminated from analysis, because she declined to complete the experiment.

2.5.1.2 Materials

Two types of experimental dialogues were created for each of the four novel verbs (*veeb*, *flom*, *glim* and *moop*). For each participant, two verbs were paired with prepositional object nouns that should bias listeners toward a modifier interpretation, while the other two verbs were paired with prepositional object nouns that should bias listeners toward an instrument interpretation. For each verb, we constructed one instrument and one modifier training dialogue. The training dialogues were the same as the ones for the familiar equi-bias verbs (*point at*, *feel*, *turn over* and *drag*) in Experiments 1 and 2, except that the verbs in the sentences were replaced by the made-up verbs, as shown in example (6). Note in the example that the meaning of the unknown verbs remained obscure in the training dialogues, despite the provision of familiar words surrounding the new verb. *Red pencil* and *magic sword* were conceivable instruments for pointing in Experiment 1, but provided little constraint on the type of action meant by *flomming*. Each child also heard five distractor dialogues mixed with the experimental dialogues during the study phase in a quasi-random order.

- (6) A: Do you remember that story about the tiger? What did Tim use to flom the tiger?
B: He flommed the tiger with the red pencil.

A: Right, he flommed the tiger with the red pencil. Hmm. If I were him, I would flom the tiger with the magic sword.

B: Wow! You want to flom the tiger with the magic sword! How exciting!

In the test phase, children received 4 critical trials with ambiguous *with* prepositional phrases, one for each novel verb, paired with appropriate toys. For example, (7) shows a test sentence that was accompanied by the same visual context shown in Figure 1B with a test sentence from Experiment 1. For this sentence the objects included a target animal (a pig with a small flower), a distracter animal (an elephant with a small crayon), a target instrument (a big flower), and a distracter instrument (a big crayon). The four critical trials were intermixed with 12 filler trials and blocked by training condition. Thus, unlike in Experiment 1, children received test trials for both instrument- and modifier-trained verbs in the test phase, rather than test trials from only one training condition. This change was made to yield a reasonable number of test trials while keeping within bounds the total number of novel verbs presented in the dialogue phase. Because the critical trials all contained a novel verb whose meaning children would not know, we introduced some words into the filler trials that children might find unfamiliar-sounding as well. Half of the filler trials thus included either a proper noun for an animal (e.g., *Onyx*, *Susie*) or a low-frequency noun (e.g., *trowel*, *filly*). As described in the procedure below, children did not need to understand what we meant by these nouns in order to participate in the test trials.

(7) Flom the pig with the flower.

2.5.1.3 Procedure

Before the study phase, the first experimenter (E1) told the children that they would watch videos of people talking to each other, and that they sometimes used words the experimenter did not know. In the study phase, children watched 9 dialogues (4 experimental and 5 distracter dialogues) presented on a laptop as in Experiment 1. In order to ensure that children gave the intended interpretations to the *with* phrases in the training sentences, E1 asked a comprehension question involving this phrase immediately after each experimental dialogue. For example, (8a) and (8b) show example questions for an instrument and modifier training dialogue. Children were highly tolerant about questions with novel verbs and were accurate in providing short answers, suggesting the discourse context of the dialogues was easy to understand.

(8a) Instrument training: What would she use to flom the tiger?

(8b) Modifier training: Which tiger did Tim flom?

In the test phase, children were instructed to collaborate with the second experimenter (E2) in this game of following audio instructions. The participant took turns with E2 in the roles of “actor” and “helper”. When E1 played the audio instruction, the actor always closed his/her eyes, while the helper listened attentively to the sentence and looked for the toys required for the action. E1 then rang the bell to signal the start of the action. The helper provided the actor with the toys for him/her to complete the action. The first critical trial was arranged after three demo trials. Children showed no difficulty in understanding the game rules. Importantly, E2 always performed the actor’s role for the critical trials, so that it would be the child’s turn to search for the needed toys, as the instruction sentence unfolded. This was done so to enable the online eye-movement measures as well as children’s offline interpretation of the globally ambiguous sentences. If the child considered a modifier interpretation, he/she would only hand the target

animal to E2. But if the child considered an instrument interpretation, he/she would hand E2 both the target animal and the target instrument. The knowledge of the verb's meaning was not necessary for children to fulfill the helper role. Depending on the toys the participant chose, E2 then conducted one of four actions (tap, poke, rub, and circle above the animal's head) either by using her finger or the instrument toy provided by the child.

Each child also completed the Peabody Picture Vocabulary Test—Third Edition (PPVT-III) at the end of the study.

2.5.1.4 Design

The 4 novel verbs were divided into two lists. Each child heard instrument-training dialogues for 2 verbs and modifier-training dialogues for the other 2, intermixed with filler dialogues in a quasi-random order. Each verb was presented in a modifier-training dialogue to half of the children, and an instrument-training dialogue to the other half. In the test phase, children received all the 4 critical test trials, one test trial for each verb, blocked by training condition. Therefore, different from Experiment 1 and 3, training condition in the test phase of the current experiment was manipulated within subjects.

2.5.1.5 Coding

Children's actions were categorized as: Instrument actions (the child picked up both the target animal and the target instrument), Modifier actions (the child only picked up the target animal), and Mini-instrument actions (only one child picked up the target animal but commented about the mini-instrument interpretation). There was one caveat however. Because children did not conduct complete actions in the current experiment, it is likely that some trials with mini-

instrument interpretation were masked by the modifier-like action. Eye-movement data were coded in the same way as Experiment 1. However, because the task was more difficult in the current experiment, we included a 4.5-sec coding window as opposed to 3.5-sec in the previous experiments in order to capture children's late responses. One trial coded as mini-instrument actions was removed from the eye-movement analyses. Four trials with fewer than 1/3 of frames coded as fixations to a quadrant of the podium were removed from analysis. In total, 5 trials (3.9 %) were eliminated from eye-movement analyses. Coding reliability was assessed for 8 children and yielded agreement on 93% of coded video frames.

2.5.2 Results and Discussion

2.5.2.1 Behavioral measures

The overall mean proportion of instrument actions was 0.74. Children had a strong bias to conduct instrument actions for the unknown verbs, which was consistent with the corpus analysis reported in Snedeker and Trueswell, 2004. Verb-attachment uses of *with* ambiguous phrases were more common than noun-attachment in children's speech as well as parents' speech (Snedeker & Trueswell, 2004). Children's offline interpretation did not differ for the sentences with instrument and modifier-trained verbs (0.73 vs. 0.75, $\chi^2(1) = 0.09$, $p > 0.7$). A significant age effect suggests older children tended to conduct more instrument actions ($\chi^2(1) = 5.44$, $p = 0.02$), but no interaction was revealed between age and training ($\chi^2(1) = 0.31$, $p > 0.5$). Analysis with sex and vocabulary was attempted but did not yield significant effect on children's action.

2.5.2.2 Online eye-movement measures

The utterance tokens of the testing sentences in the current experiment had very similar duration (average: 2.3 seconds) as the previous experiment. Therefore, we adopted the same window lengths for statistical analysis, with one exception for the Noun-2 region. Considering the difficulty of processing novel verbs, in addition to the early and late Noun-2 regions, we also included a third 600-ms time window to evaluate the later training effect.

Figure 10 plots eye fixations in the three critical analysis windows, as children heard the verb (e.g. *point*), the direct object noun (Noun-1), and the prepositional phrase object noun (Noun-2). Figure 10A shows the proportion of fixations to animals (either target or distracter) and instruments (either target or distracter) in the Verb region. As in Experiment 1, during the Verb region, children tended to look more towards either instrument and less towards either animal after hearing an instrument-trained verb. Statistical analysis revealed a marginal training effect on fixations to animals ($\chi^2(1) = 2.89, p = 0.09$) and a non-significant effect on fixation to instruments ($\chi^2(1) = 1.83, p = 0.18$). Similar to what we found in Experiment 2, the effect on the anticipatory looks was mainly carried by girls, not boys. Girls' looks to either animal were significantly affected by training in the Verb region ($\chi^2(1) = 5.58, p = 0.02$). These results are consistent with the findings in Experiment 1 and 2, suggesting children rapidly retrieved the newly learned verb bias information soon after they encountered the same verb.

Figure 10B shows the proportion of fixation to animals and instruments as the target animal name unfolded. Children's fixation pattern was not affected by training during the Noun-1 region ($\chi^2(1) < 0.8, p > 0.3$).

Figure 10C shows the proportion of fixations to the target animal (TA) and target instrument (TI) in the Noun-2 window. In the First Noun-2 window, children tended to look

more at the target animal in trials with a modifier- rather than an instrument-trained verb. Though the overall training effect was not significant ($\chi^2(1) < 0.80, p > 0.3$), the effect size reliably correlated with children's age, which was entered as a continuous variable ($\chi^2s(1) = 5.91, p = 0.015$). Figure 11 plots the mean proportion of fixation to target animal by median-split age group. Older children (mean age: 61.4 months) showed a larger training effect than younger children (mean age: 56.1 months). The training dialogues did not significantly affect children's fixations in either the Second or the Third Noun-2 window ($\chi^2s(1) < 0.81, p > 0.3$). Analysis with vocabulary and gender showed no effect on children's fixation during Noun-2 regions.

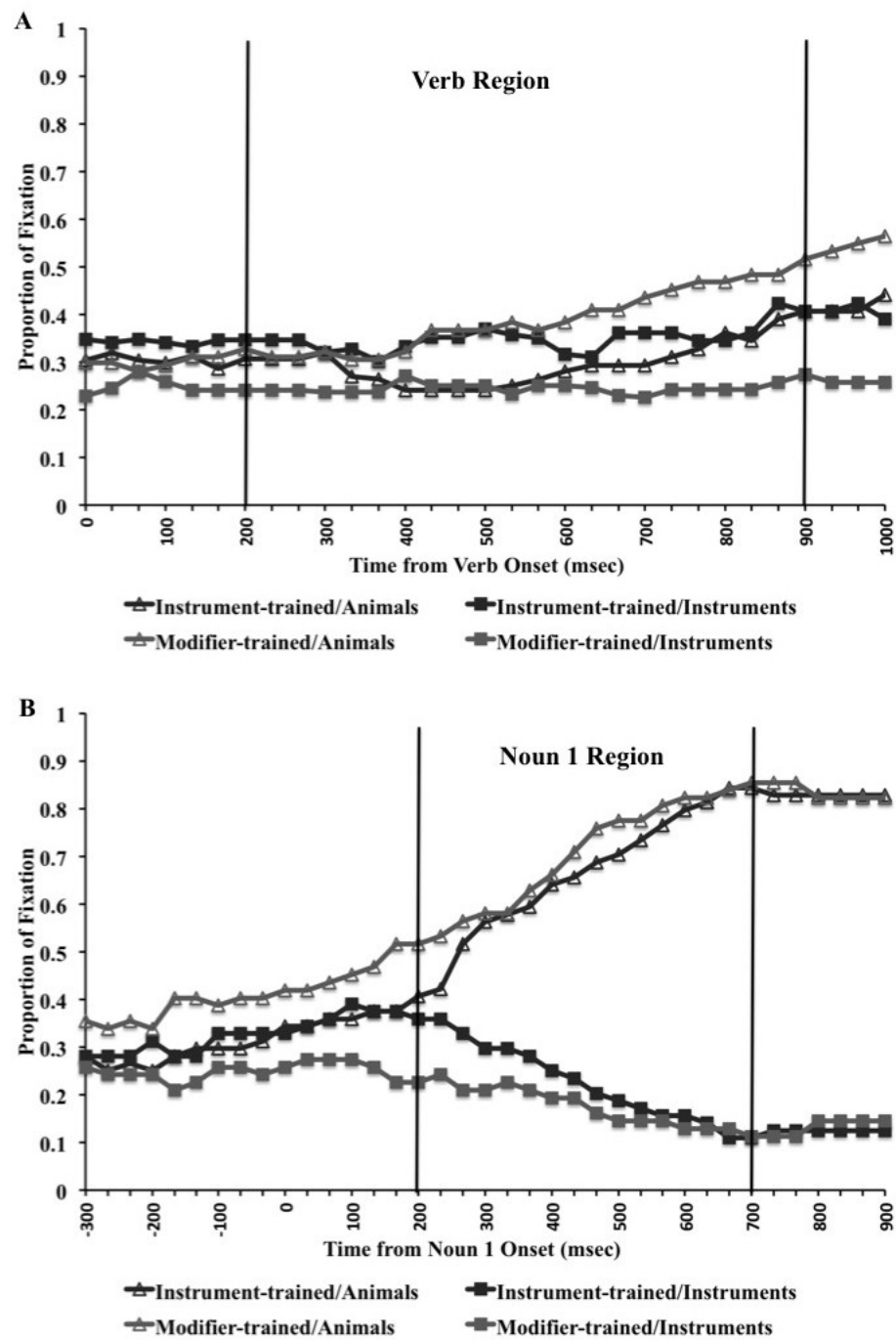


Figure 10 (cont. on next page)

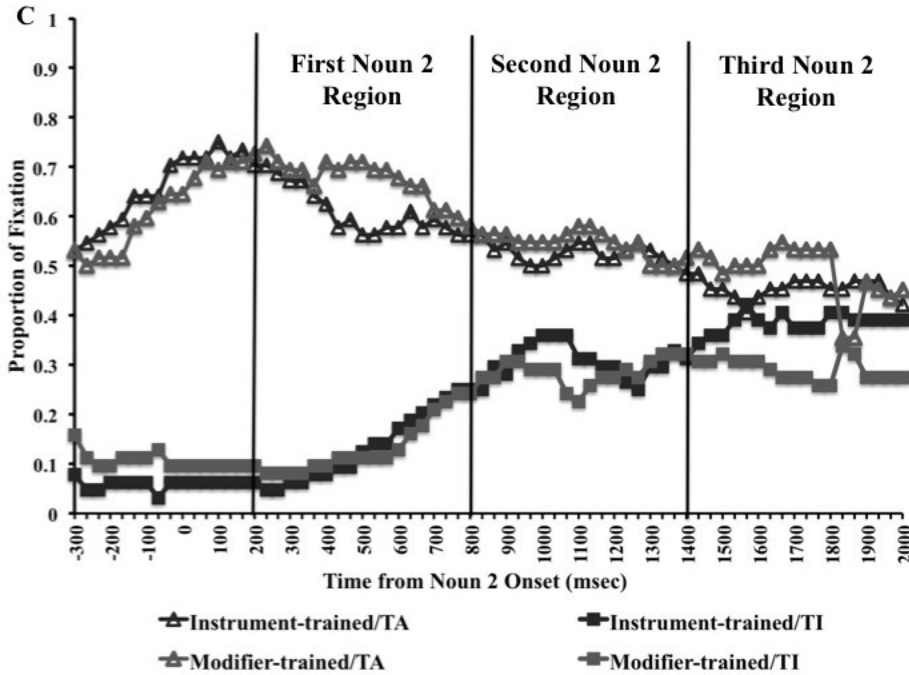


Figure 10: Experiment 4 eye-movement results. 5-year-olds' proportion of fixation to objects, time-locked at the onset of the verb (A), direct object noun (B) and the prepositional phrase object noun (C). The dark grey lines represent test trials containing instrument-trained verbs. The light grey lines represent test trials containing modifier-trained verbs. The vertical lines in each plot define the width of the analysis window. TA: target animal; TI: target instrument.

The overall pattern of the training effect was similar to Experiment 1 and 2, except that the anticipatory looks during the verb region was not different between the first half versus the second half of the experimental trials. This was possibly because the training dialogues were children's only experience with these novel verbs. Children had only heard the novel verbs in the context of ambiguous *with* phrases and therefore the high co-occurrence frequency of the novel verbs and the *with* structures was rapidly retrieved upon hearing the same verbs.

Mixed-effect analysis with 2 levels of training conditions (instrument vs. modifier) x 5 sentence positions (Verb, Noun-1, the First Noun-2, the Second Noun-2 and the Third Noun-2) repeated measures was conducted. Participants' fixations to animals and instruments changed across the sentence, yielding a main effect of sentence position ($\chi^2_s(4) > 27.6, p < 0.0001$).

There is a marginal training effect on fixations to animals ($\chi^2(1) = 3.16, p = 0.075$) and a non-significant training effect on fixations to instruments ($\chi^2(1) = 2.26, p = 0.13$).

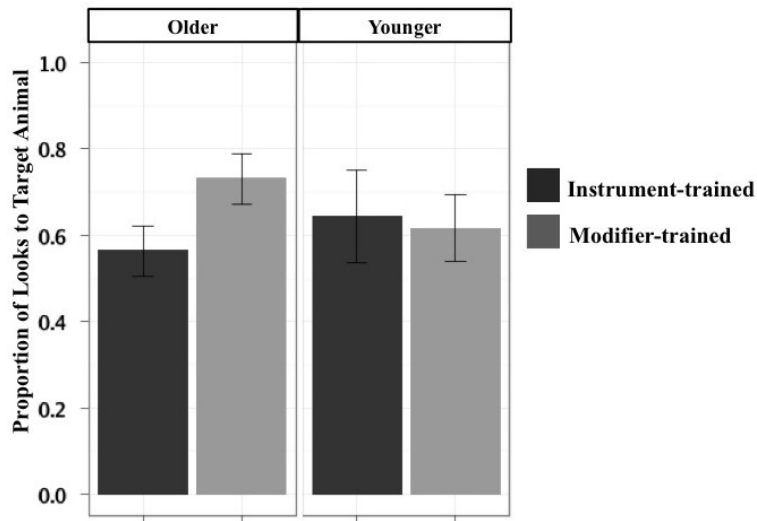


Figure 11: Experiment 4 eye-movement results by age. Mean proportion of fixation to target instrument during the First Noun-2 region. Older children looked more to the target instrument in instrument-trained trials and less in the modifier-trained trials. Error bars represent standard errors.

These results provided evidence of verb bias learning in 5-year-olds, which was independent of the event information from the verb semantics. Children, particularly girls looked less at the animals upon hearing an instrument-trained verb early during the Verb region, which suggested a rapid retrieval of verb-specific linguistic information. When the sentence was fully unfolded at the Noun-2 region, older children looked more at the target animal for modifier-trained trials as opposed to instrument-trained trials, indicating children's capability of using the newly learned bias to resolve ambiguity depended on their age.

2.5.3 Summary

The findings of the current experiment suggested 5-year-olds were capable of attaching abstract linguistic-distributional information to individual verbs with limited event knowledge

from the verb semantics. Similar as found in Experiment 1 and 2, newly learned verb bias guided children's anticipatory looks soon after they heard the same verbs. After hearing PP-object noun region, older children were able to use the verb bias information to resolve ambiguity. The training effect was seemingly smaller compared to that in Experiment 1. However, considering three important aspects of the experiment design, we found the revealed training effect highly impressive. Firstly, processing sentences with unknown verbs in both the study phase and the test phase is cognitively more challenging. The age effect observed in eye-movement data demonstrated that children with more advanced development were more capable of learning novel verb bias. Secondly, children's task in the test phase did not require explicit interpretation for the test sentences. The collaborative style of the task might have slowed down online processing of the target sentences. Therefore the difference between instrument and modifier-trained verbs further supported the crucial role of linguistic-distributional information in verb bias learning.

However, our results did not thoroughly rule out the influence of event-distributional information in verb bias learning. The discourse context (e.g. *What did Tim use to flom the tiger?*) and the semantic feature of the PP-object nouns (e.g. *magic sword*) both indicated instrument use in an unknown event. An alternative explanation would be that 5-year-olds were able to learn verb-specific instrument preference, relying on the co-occurrence frequency of a novel verb and the likelihood of instrument use, without verb semantics. As a matter of fact, even in isolated phrases, 2-year-olds learned the co-occurrence frequency about the individual verb and the semantic category (e.g. *stipe the pig* vs. *nerk the fork*) of its direct object nouns (Yuan, Fisher, Kandhadai, & Fernald, 2011). However, children did not learn to associate a category of nouns with a neighboring novel noun, indicating the verb-argument structure

motivated the encoding of the combinatorial fact about the verb and the semantic category of its argument. Therefore in the current experiment, the event-distributional information, though a lot more abstract than the previous experiments, could still be extracted from the semantic category of the *with*-PP object nouns and further contributed into verb bias learning.

2.6 General Discussion

The present study explored whether 5-year-olds and adults can track verb-specific bias of either instrument PP-attachment or modifier PP-attachment from a mix of language experiences with multiple verbs, by examining the process of ambiguity resolution of sentences containing ambiguous *with* phrases after training. Across four experiments, the results all pointed toward a rapid statistical learning mechanism in verb bias acquisition. Importantly, our findings suggest though semantic information of the action event is greatly involved in learning verb bias in reality, linguistic distributional information, i.e. the likelihood a specific verb co-occurs with a structure, plays a crucial role as an independent information source for verb bias learning. Experiment 1 and 2 provided listeners with short dialogues for each of the eight equi-bias verbs containing ambiguous *with* PP phrases. We found that the structural preferences of familiar equi-bias verbs were reshaped by a brief verb-specific listening experience in both 5-year-olds and adults. During the study phase, listeners attached each familiar equi-bias verb with its co-occurrence frequency with either the modifier PP phrases or instrument PP phrases. During the test phase, listeners retrieved the updated verb bias information when encountering the same verbs. The fine-grained eye-movement data revealed listeners' interpretation of the prepositional object nouns (Noun-2, Figure 2C & Figure 5C) was affected by whether the particular verb in the sentence was heard in an instrument or a modifier training dialogue. In addition, the training

effect shifted early in both children and adults during the second half of the experiments (Figure 3 & Figure 5), indicating information about the verb became active immediately with the participant encountered the verb, once they were more experienced with the testing sentences. Similar evidence was hinted in Snedeker and Trueswell (2004) for the biased verbs. During the 300 ms before the onset of the prepositional object noun, children appeared to look more at the animals in sentences with modifier-bias verbs as opposed the sentences with instrument-bias verbs. However it is not clear whether such bias effect emerged earlier in the direct object noun region and even earlier in the verb region.

Given the fact that the training dialogues in Experiment 1 and 2 also delivered rich event information regarding the use of an instrument in the action, it is curious whether the retrieval of the updated verb bias was mainly derived from the linguistic co-occurrence frequency or from the semantic association between the verb and the instruments. On one hand, as discussed earlier, verb-structure co-occurrence frequency plays an important role in guiding people's online comprehension (Boland, 2005; Boland et al., 1990; Garnsey et al., 1997; Trueswell et al., 1993). In a recent sentence-processing model trained by a large set of corpus data, verb information substantially increased the performance of predicting an upcoming structure. Verbs with stronger structural bias has the strongest ability in predicting the subcategorization structure (Roland, Elman, & Ferreira, 2006).

On the other hand, the tight relationship between the verb and its semantic associations during online sentence processing became another focus of research. Altmann and Kamide (1999) found that upon hearing "eat", listeners rapidly looked at the edible items on the display (Altmann & Kamide, 1999). Sussman (2006) found people's anticipation of an instrument differed as a function of whether an instrument is semantically required or semantically optional

for the verb action. For example, people were more likely look at a possible instrument (e.g. a pencil) soon after hearing a verb with instrument bias (e.g. *poke* in *poke the dolphin*), than hearing a verb that does not require an instrument (e.g. *touch* in *touch the dolphin*). The instrument bias was normed by the frequency of a tool being used in the action. Therefore Sussman (2006) interpreted the anticipatory looks as supportive evidence for the view that semantic access of the verb immediately activates its likely instrument (Sussman, 2006). Same findings were replicated in a later study using sentences containing VP-attached *with* phrases. For example, participants rapidly looked at the plausible instruments (e.g. *sword*) in a visual world paradigm during the verb region, direct object noun region upon hearing an instrument-obligatory verb (e.g. *hack*). In contrast, listeners did not make saccades to the instruments until it was mentioned in the sentence containing an instrument-optional verb (e.g. *injure*) (Bienvenue, Maurer, & Koenig, 2007). If the semantic association between the verb and its instrument argument was the key, training with event-distributional information in the current study is expected to elicit the anticipation of instruments to a similar degree as the training with both the event and linguistic-distributional information.

The findings of Experiment 3 revealed a weaker training effect without reduced involvement of the linguistic co-occurrence frequency information in the training. There was a tendency that children looked more at the target instrument in the sentences with instrument-trained trials than in those with modifier-trained trials. However the marginal effect emerged only after the prepositional object noun was mentioned in the sentence (Figure 8), consistent with the late instrument looks for the instrument-optional verbs found in Bienvenue et al. (2007). These results suggested the semantic link between the verb and the critical nouns was not strong enough to guide listeners' anticipation.

Experiment 4 provided further support for the view that linguistic distributional information is sufficient to guide listeners' expectation toward instruments even without the verb semantics. For each verb, children heard two ladies talking about a non-depicted event using the combination of a novel verb and either a *with* instrument PP object or a *with* modifier PP object. The co-occurrence frequency of an unknown verb and the type of the *with* phrases was sufficient to activate or suppress children's anticipation of an instrument (Figure 10A). Arguably, the use of an instrument in the unknown action is part of the semantics of the action. However, the association of a verb and any arbitrary instrument is much more abstract without the verb meaning. Taken together, children and adults learn and retrieve the abstract co-occurrence frequency in the linguistic distributional information upon hearing the verb.

How about the use of the newly learned verb bias for ambiguity resolution? The training manipulations in all four experiments (only 4 sentences of exposure per verb) were too subtle to affect people's proportion of instrument actions as a measure for their final interpretation. However, fine-grained eye-movement analysis showed the retrieval of verb bias during the first 600-ms time window after the onset of the prepositional object noun across experiments, indicating people's initial consideration of either instrument or modifier interpretation of the ambiguous *with* phrases was affected by their training experiences (Figures 2C, 5C, 8C and 10C). One important note needs to be made about the training effect revealed in Experiment 3, especially in the high-vocabulary children. The results suggest event-distributional information might also play a facilitating role in ambiguity resolution. However, based on the evidence from children's production data during training, we could not entirely eliminate the possibility that this effect was carried by some children who were able to transfer the event-distributional information to the linguistic-distributional information either during the study phase or the test

phase. Experiment 4 adopted a different design and procedure from the previous three experiments in order to examine children's responses to the sentences containing unknown verbs. The results showed older children showed greater training sensitivity in their online fixation. However the age effect was mainly due to the complexity of the task. Moreover, the older children who carried the interaction between age and training in Experiment 4 were not much older compared to the average age of participants in Experiment 1 and 3 (61.4 months vs. 60.6 months). Therefore Experiment 4 provided equally strong evidence for the capability of 5-year-olds in applying the verb bias, which was learned without support of verb semantics. In summary, these data indicate the newly learned verb bias participated in online ambiguity resolution in the way that is similar as familiar verb bias.

The current study provides strong evidence for the critical role of linguistic distributional information and proposes the statistical learning mechanism for verb bias learning that is independent of verb semantics. Meanwhile, our findings also indicate the involvement of both linguistic and event distributional information in verb bias learning, because reducing the contribution of either information source resulted in a relatively weaker training effect.

However, there are several limitations of the present study, which will be addressed in the future direction. First of all, the challenges of dissociating the linguistic from event distributional information are rooted in the asymmetry between instrument and modifier attachment of the *with* PP phrases. Instrument is considered as a semantic argument for the verb (Koenig, Mauner, & Bienvenue, 2003), while modifier object noun barely participates in the semantics of the verb action. Therefore learning the PP-attachment preference of the verb inevitably involves the interaction between semantics and syntax, which is perhaps the reality for learning most types of verb bias at the first place. However, inspired by the adult studies such as the miniature language

learning task (Wonnacott, Newport, & Tanenhaus, 2008) and the dative syntactic priming (Coyle & Kaschak, 2008), we are interested in exploring the learning mechanism of verb bias where the alternative structures deliver same or similar meaning.

Another limitation of the study is inherited from the all-or-none design of the training paradigm. In natural language, verb bias is probabilistic rather than discrete categories. The current study does not directly address the mechanisms of probabilistic learning, given its extremely short study session. However in Experiment 1-3 at least the updated verb bias at the test session was probabilistic, because listeners had baseline verb bias knowledge of these familiar equi-bias verbs.

Finally, the present project has the methodological limitations in monitoring real-time learning process during the study phase. The results revealed in the test phase reflected both the retrieval efficacy and the encoding efficacy. From the perspective of the function of learning, these two aspects lead to the same outcome during online language processing. However, encoding and retrieval might involve completely different sources of information. For example, a weaker training effect in the Experiment 3 might be derived from the more difficult retrieval without the syntactic similarity between the training and testing sentences. Alternatively children might have encoded the distributional information less efficiently under the unambiguous training context. In order to tease these two aspects apart, future experiments using different techniques will explore the fine-grained learning process during verb bias training.

In sum, we conclude that children and adults are able to rapidly learn verb-specific structural preference via a strikingly minimal language exposure either in the context where event-distributional information was richly provided by verb semantics or in the context where event-distributional information was greatly reduced. Learners soon apply the new statistical

information about particular verbs in making prediction during the earlier stage of sentence processing as well as resolving ambiguity as the sentences unfolded.

CHAPTER 3 ELECTROPHYSIOLOGICAL PROCESS UNDERLYING NOVEL VERB BIAS LEARNING

The likelihood of structural alternatives for verbs (verb bias) guides online sentence comprehension (Garnsey, Pearlmutter, Myers, & Lotocky, 1997; MacDonald, Pearlmutter, & Seidenberg, 1994; Osterhout, Holcomb, & Swinney, 1994; Trueswell, Tanenhaus, & Kello, 1993). The current research focuses on how people learn frequency-based verb biases. In particular we are interested in determining whether such biases can be learned without support from verb semantics. Three EEG experiments were conducted to investigate verb-bias learning when participants read sentences containing novel verbs. Sentences contained prepositional phrases that could be either verb instruments or direct-object modifiers and which role they took was disambiguated by critical noun meaning. The results suggest a highly dynamic learning system that continuously collects statistical information about words in sentences and applies the newly learned information during online comprehension. Over the course of training, there was a transition from an event-related brain potential (ERP) component associated primarily with meaning processing (N400) to another component associated primarily with structure processing (P600), suggesting a change in the nature of the processing as experience with the novel verbs increased. A similar transition has been observed in previous ERP studies of second language grammar acquisition (McLaughlin et al., 2010; Osterhout et al., 2008; Osterhout, McLaughlin, Pitkanen, Frenck-Mestre, & Molinaro, 2006). We also found that biases were learned better for verbs that were trained in ambiguous sentence structures than for those trained in unambiguous structures, suggesting that resolving ambiguity might be a crucial component of verb bias learning. When there were cues other than the verb making the training sentences unambiguous, learning of verb biases was less effective. Finally, our results also revealed large variability

across participants. Some learners were more sensitive to statistical information in the language input than others, reflected by higher efficiency in both verb bias learning and garden-path sentence comprehension.

3.1 Introduction

Adults' knowledge about verb-specific structural preferences, called verb bias, plays a central role among many other constraints in guiding parsing as sentences unfold over time. Numerous studies have shown that language users develop expectations about the upcoming words in a sentence based on verb bias, and find sentences that violate those expectations difficult to process, reflected as longer reading times (e.g., Boland, Tanenhaus, & Garnsey, 1990; Garnsey et al., 1997; Trueswell, 1996) in studies of written sentences or as fixations to an incorrect destination location in a visual world paradigm with spoken sentences (Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995; Trueswell et al., 1999). In addition, when sentences are globally ambiguous, language users prefer interpretations that are consistent with their knowledge about the bias of the verb (Snedeker & Trueswell, 2004; Snedeker & Yuan, 2008).

As shown in the studies reported in Chapter 2, language users appear to continuously collect verb-specific distributional information from language input and update their knowledge of the biases of particular verbs. Both children and adults were found to be highly sensitive to linguistic distributional information about particular verbs and to apply the newly learned information during online parsing with strikingly minimal exposure. Experiment 4 showed that preschoolers were able to retrieve the linguistic distributional information even without knowledge of verb semantics. The question of interest in the current study concerns the neural

systems underlying rapid adaptation to the dynamic changes in the linguistic distributional information as learners gradually establish associations between a verb and its preferred structure. We recorded event-related potentials (ERPs) from the scalp to take advantage of the fact that the electroencephalogram (EEG) is highly sensitive to transient events in the brain in order to explore the verb bias learning process across training. A few previous studies investigated the neural markers of the learning of artificial languages or natural second languages and those will be reviewed below after a brief introduction to some relevant ERP components.

Two ERP components have been found to change over the course of language learning. The first is the N400, which is a negative deflection peaking around 400 ms after stimulus onset that reflects primarily meaning processing. N400 amplitude is sensitive to several kinds of manipulations of lexical and contextual features, including words' cloze probability (DeLong, Urbach, & Kutas, 2005; Kutas & Hillyard, 1984), semantic fitness with preceding context (Federmeier, Wlotko, De Ochoa-Dewald, & Kutas, 2007; Kutas, 1993), semantic priming (Bentin, McCarthy, & Wood, 1985; Kutas, 1993), word frequency (Van Petten & Kutas, 1990) and word neighborhood size (Laszlo & Federmeier, 2010). In contrast, the P600 is a positive deflection starting around 500 ms after stimulus onset and continuing for a few hundred milliseconds. P600 amplitude reflects aspects of structure processing and its amplitude increases in response to various types of morpho-syntactic violation (Hagoort, Brown, & Groothusen, 1993; Osterhout & Holcomb, 1992), more complicated syntax (Friederici, Hahne, & Saddy, 2002; Kaan, Harris, Gibson, & Holcomb, 2000), and less preferred syntactic structure, such as garden-path sentences with sentence endings that are inconsistent with the expectation based on verb bias (Itzhak, Pauker, Drury, Baum, & Steinhauer, 2010; Osterhout, Holcomb, et al., 1994). In contrast, others argue that P600 is not specific to syntactic processing in language, but is

instead an instance of P3b, which is sensitive to the perceived probability of many kinds of stimuli (Coulson, King, & Kutas, 1998; Patel, Gibson, Ratner, Besson, & Holcomb, 1998). For present purposes, it is not necessary to settle questions about the language-specificity of P600, nor will the results speak to that issue. It is only necessary that it differs from N400.

In word/sequence-learning paradigms, N400 has been found to be sensitive to the amount of exposure to pseudowords. McLaughlin and colleagues studied N400 changes during French L2 word learning. The N400 amplitude difference continued to increase across three learning sessions when comparing learners' ERP responses to pseudowords relative to the related word targets, indicating a gradual process of learning word meanings without requiring any overt lexicality judgment (McLaughlin et al., 2004). Another cross-sectional word-learning study showed that the N400 elicited by words in a second language was smaller than that elicited by words in the reader's native language. More importantly, this difference in N400 amplitude was smaller in participants with higher proficiency in the second language (Midgley, Holcomb & Grainger, 2009). A recent study has provided evidence for a similar word learning effect across training within only fourteen minutes of passive exposure to new spoken pseudo-words (Shtyrov et al., 2010). As meanings for novel words were learned from sentence contexts, N400 amplitude increased and was comparable to the N400 elicited by real words after just three trials (Mestres-Missé, Rodriguez-Fornells, & Münte, 2007). N400 amplitude has also been found to respond to the degree of learning in a study using nonlinguistic auditory sequences. Participants with high behavioral accuracy in a post-training recall task showed a larger N400 effect to the onset of familiar sound sequences compared to those with low behavioral accuracy (Abla, Katahira, & Okanoya, 2008). In sum, N400 has been found to change as new words, pseudowords, or new sequences of sounds are learned.

P600 has also been found to change as a new artificial grammar is learned in studies using statistical training paradigms. The response to ungrammatical artificial structures became indistinguishable from the effect found in natural language processing (Christiansen et al., in press; Friederici, Steinhauer, et al., 2002; Hsu, 2009), suggesting that similar neural processes are shared between statistical learning in artificial sequences and grammar learning in natural language. These studies have provided support for the importance of statistical learning in syntactic acquisition. However, little is known about how learners acquire knowledge about grammatical structure during the course of training. A series of second language studies recently done by Osterhout and colleagues tracked learning longitudinally to explore when and how learners started to categorize the information from the language input they were receiving and apply the grammatical knowledge they were acquiring about gender, number and person agreement. When participants were tested on a grammaticality judgment task, an N400 effect was observed at early learning stages, but that changed to a P600 effect at later stages of learning. The two discrete neural responses were explained as reflecting a transition from rote-memorized lexical processing to rule-based grammatical processing (McLaughlin et al., 2010; Osterhout et al., 2008, 2006).

The present study took a longitudinal approach similar to McLaughlin et al. (2010) to examine the dynamics of neural responses while participants were learning verb-specific structural biases in a brief training paradigm. Participants read training sentences containing one of four novel verbs for 30-75 minutes during training. Half of the sentences contained ambiguous prepositional *with* phrases that could be either verb instruments or direct-object modifiers. The meaning of the noun in the prepositional phrase (PP) disambiguated the role of the PP in the sentence, as illustrated by comparing *tractor* and *stalks* in example (1a) below. The other half of

the sentences contained unambiguous *using* and *that has* phrases in the place of *with* phrases so that the disambiguating information arrived earlier before the critical nouns (see example 1b). Each participant saw each verb used in only one of the four structures illustrated in (1) across multiple sentences during training.

(1). a. *The suntanned farmer dakked the corn...*

with the large tractor / with the high stalks.

b. *The suntanned farmer dakked the corn...*

using the large tractor / that has the high stalks.

Three major questions were addressed:

1). How did participants learn to use newly learned verb biases to guide online prediction about upcoming words and structures, without any support from verb semantics?

We monitored participants' EEG continuously during training with the goal of observing training effects in real time. Before each sentence containing the novel verbs, participants were shown two pictures that depicted words that would be mentioned somewhere in the upcoming sentence. One depicted an object that would be the direct object in the upcoming sentence and the other depicted an object that would be an instrument mentioned somewhere in the sentence. Participants were encouraged to try to predict when the pictures would be mentioned while reading the sentences. An intentional asymmetry was built into the design with respect to the predictability of the instrument and modifier words in the sentences. Once participants learn the verbs' structural preferences, the position of the instrument nouns should be much more predictable in sentences with instrument-bias verbs than in those with modifier-bias verbs. That is, *tractor* should be more predictable after the words *with the* or *using the* in sentences with instrument-bias-trained verbs than after the words *with the* or *that has the* in sentences with

modifier-bias-trained verbs. Two things were uncertain after modifier-bias verbs: 1) the sentence position of the possible instrument noun, and 2) which of the many possible modifiers of the direct object noun would be mentioned (e.g. corn with *high stalks*, corn with *sweet taste*, etc). The goal was to exploit this asymmetry to allow us to observe whether participants developed different learning strategies for the two kinds verbs in a passive reading task.

Another goal was to see whether we would observe the transition from N400 effects to P600 effects previously found by Osterhout and colleagues for second language learning (McLaughlin et al., 2010) and by Morgan-Short and colleagues (2010) for artificial grammar learning (Morgan-Short, Sanz, Steinhauer, & Ullman, 2010). Participants may start out focusing on predicting semantic features of the critical nouns but then transition to predicting a syntactic role instead, especially given that the sentences were constructed to prevent participants from learning a coherent meaning for the novel verbs.

2). What is the role of ambiguity resolution in verb bias learning?

The rationale for comparing ambiguous and unambiguous training was to examine whether active ambiguity resolution facilitates the learning of structural biases. Previous neuroimaging studies have shown that resolving syntactic ambiguity requires extra processing resources, as shown by larger P600 amplitude and greater hemodynamic activation in Broca's area and Wernicke's area (Hagoort, Brown, & Osterhout, 1999; Mason, Just, Keller, & Carpenter, 2003; Osterhout, Holcomb, et al., 1994). It is possible that the extra processing required by ambiguity resolution would strengthen the encoding of verb bias information in the ambiguous condition, because it is a reliable cue that can reduce such cost. In contrast, in the unambiguous conditions, verb bias is redundant with other stronger disambiguating cues. Error-

based learning models (e.g, Chang, Dell, & Bock, 2006) would predict better verb bias learning in the ambiguous conditions.

3) Are there individual differences between people that affect the efficacy of verb bias learning?

The large variability across learners observed in the L2 learning literature (see Grosjean, 1998; Osterhout et al., 2006 for reviews) raises questions about the characteristics of good learners. In response to verb agreement violations tested after a second learning session, McLaughlin et al. (2010) found one subgroup of learners showed primarily N400 effects while another subgroup instead showed primarily P600 responses. Both groups showed P600 responses at the end of a third session. The amplitude of P600 responses correlated with behavioral accuracy, indicating that learners who showed greater behavioral sensitivity to training also progressed more quickly from N400 to P600 responses. It was unclear why learning rate differed between participants.

The current study considered two kinds of individual differences, based on findings from previous comprehension studies. The first was familial left-handedness. A previous ERP study (Qi, Jackson, & Garnsey, 2010) compared mixed right-handed participants, who have left-handed relatives, with pure right-handed participants in a garden-path sentence-reading task. Critical words were lateralized to one visual field or the other in that study and the results showed that when critical words were presented the left hemisphere, pure right-handers displayed a larger P600 verb bias effect than mixed right-handers. There was no familial handedness effect when critical words were presented to the right hemisphere. These findings suggest that pure right-handers were more sensitive to verb bias than mixed right-handers. We therefore predicted that pure right-handers would have a higher sensitivity to verb bias and

would show better learning in our novel verb bias training. Such a hypothesis leads to the second kind of individual differences tested in the current study. Verb bias sensitivity has been proposed to reflect a general ability to use lexical constraints during online parsing (Novick, Thompson-Schill, & Trueswell, 2008), so we tested whether participants' sensitivity to the biases of familiar verbs predicted their learning of novel verb biases.

3.2 Experiment 5: Encoding novel verb bias

The purpose of the experiment was to establish a short learning paradigm for novel verb bias learning without verb semantics. We monitored the online electrophysiological responses to the continuous input of verb-specific linguistic distributional information. We were interested in identifying neural responses that are sensitive to the degree of distributional information exposure and how readers learn to actively integrate the newly learned verb bias information into online reading.

3.2.1 Methods

3.2.1.1 Participants

32 participants (16 males and 16 females, ages 18-22) from the participant pool in the Psychology Department at the University of Illinois participated in the study after providing informed consent. All participants were right-handed with normal hearing and normal or corrected-to-normal vision. None of the participants had been exposed to languages other than English before the age of 5. Experimental participation was compensated either with 8 dollars/hour or course credit.

3.2.1.2 Materials and Design

The stimuli consisted of 64 sets of four sentences created by crossing two attachment types (instrument vs. modifier) by two levels of ambiguity (ambiguous vs. unambiguous). For each of the four novel verbs, sixteen different sentence sets were created. The sentences within a set had the same beginning but then diverged after the novel verb with one including an instrument phrase and the other a modifier phrase. For example, in the sentence *The suntanned farmer dakked the corn with the [big tractor / high stalks]*, the noun *tractor* serves as an instrument for the unknown verb (e.g. *dak with the big tractor*) and the noun *stalks* serves as a modifier for the direct object noun (e.g. *corn with the high stalks*). Adjectives were chosen to be plausible for both instrument nouns and modifier nouns. Therefore, in the instrument and modifier ambiguous conditions, sentences remained temporarily ambiguous until the critical nouns appeared (e.g. *tractor* and *stalks*). In the instrument and modifier unambiguous conditions, sentences were rendered unambiguous earlier in the sentence by substituting either *using* or *that has* in place of *with* (e.g. *dakked the corn [using the big tractor / that has the high stalks]*). See more examples in Table 2. The choices of critical nouns were determined by a norming study described below.

In order to evaluate the strength of the disambiguation provided by the critical nouns toward instrument or modifier interpretations, a norming study was conducted on 237 pairs of sentences containing ambiguous *with* phrases. Fifty native English speakers who did not participate in the main study judged sentences in the norming study. The two pairmates of each sentence pair were presented in two different lists so that no participant saw both the instrument and modifier versions of a sentence pair (e.g. *dak the corn with the big truck* and *dak the corn with the high stalks*). Participants provided two ratings for 237 sentences on a 7-point scale. They

rated how likely it was for the noun in the *with*-phrase to be 1) an instrument of the verb and 2) a property of the direct object noun. They gave both kinds of ratings for each of the sentences they saw. For each sentence, the difference between the average instrument and average property ratings provided a measure of how strongly disambiguating the critical nouns were toward one or the other structure. Sixty-four pairs of items that had absolute differences greater than 2 for both members of the sentence pair were chosen for the main study.

Subordinate clauses were added to the sentences following the critical nouns so that the instrument nouns could be mentioned after the end of the main clause in the Modifier sentences (e.g. *as soon as he needed to use the tractor*). This was done so that 1) the instrument noun appeared in both the Instrument and Modifier conditions in the *dakking* event, and 2) any priming from simply seeing the pictures and their text labels was controlled in the responses to those same words when they appeared in the sentence. The same conjunctions (e.g. *as soon as*, *though*, *even if*) were used to connect the main and subordinate clauses in the Instrument and Modifier versions of a sentence set.

Four counterbalanced lists of 64 sentences each were constructed such that across lists there were equal numbers of sentences of each type in each list. Each participant read all sixty-four sentences with the modifier structure for two of the verbs and with the instrument structure for the other two verbs, and within each of those, with the ambiguous structure for one verb and the unambiguous structure for the other. As listed in Table 3, lexical properties of the critical nouns were approximately balanced for word length, frequency (Brysbaert & New, 2009), concreteness and imageability (Wilson, 1988) between instrument and modifier conditions.

Table 2: Experiment 5 Examples of the training stimuli. The underlined words are critical nouns. The italic words were pictured in the picture-viewing session at the beginning of each trial. I: instrument; M: modifier; AM: ambiguous; UN: unambiguous.

Condition	Sentence Frames	Critical Words	Sentence Endings
I-AM / I-UN	The suntanned farmer	with/using the big <u>tractor</u>	... as soon as he needed to harvest the crop.
M-AM / M-UN	daked the <i>corn</i> ...	with/that has the high <u>stalks</u>	... as soon as he needed to use the <i>tractor</i> .
I-AM / I-UN	The nervous student	with / using the long <u>mail</u>	... since she had missed a class meeting.
M-AM / M-UN	glimmed the <i>teacher</i> ...	with / that has the good <u>reputation</u>	... since she lost the important long <i>mail</i> .

Table 3: Experiment 5 & 6 Lexical properties of critical nouns.

Training Condition	Length (letters)	Frequency (words/million)	Concreteness (100~700)	Imagability (100~700)
I-AM / I-UN (e.g. <i>tractor</i>)	5.73 (0.25)	25.36 (6.90)	345.31 (42.89)	351.26 (43.33)
M-AM / M-UN (e.g. <i>stalks</i>)	6.28 (0.24)	30.62 (5.89)	283.44 (46.58)	281.39 (46.00)

Mean of the length of words, word frequency per million words (retrieved from American English subtitles, SUBTLEX_{WF}), ratings for concreteness and imageability are reported with standard errors in parenthesis.

3.2.1.3 Procedures

Participants were seated in a dimly lit sound-attenuating booth in front of a 20-inch LCD monitor. They responded to post-sentence comprehension questions by pushing one of two buttons on a Cedrus RB-830 response pad and their responses and response times were recorded. Stimulus presentation and behavioral response recording were performed using the Presentation[®] software package (Version 0.70, www.neurobs.com).

The sequence of events for a trial is illustrated in Figure 12. Each trial began with 4 crosses at the center of the screen, which stayed on the screen for a randomly varied duration of 1360-5000 ms. There was also a varied interval of black screen of 0-1480 ms after the warning cue and before any pictures appeared on the screen. Next, two labeled clipart images were presented side by side on the screen for 5 seconds. One pictured the direct object noun (e.g. corn) and the other pictured the instrument noun (e.g. tractor) and words labeling them appeared below them. Participants were told to pay attention to the pictures because both of them would be mentioned somewhere in the sentence they would read next.

Sentence reading started with a fixation asterisk at the center of the screen for 500 ms followed by word-by-word central presentation of the sentence. Words were presented in Helvetica 22-point white font on a black background above the fixation asterisk. Each word remained on the screen for 300 ms followed by 200 ms of blank screen, for an SOA of 500 ms. Participants were asked to minimize muscle activity and to fixate centrally while reading the sentences.

In order to maintain participants' attention and to encourage them to disambiguate the ambiguous sentences, a comprehension question was presented one second after the last word of

the sentence along with two answers to choose between. For example, the participant would see the question and response options in example (3a) in an instrument training trial and those in example (3b) in a modifier training trial. Participants had a 4-second response deadline and received feedback on their accuracy after each trial. The next trial started 1800-2500 ms after the feedback. Participants were given the opportunity to take breaks between blocks and this portion of the experimental session lasted about 30 minutes.

(3a) Instrument Training Trial: *How did the suntanned farmer dak the corn?*

A. *Using the tractor.* B. *Using the wagon.*

(3b) Modifier Training Trial: *What did the suntanned farmer dak?*

A. *The corn that has the sweet taste.* B. *The corn that has the high stalks.*

3.2.1.4 Electroencephalogram recording and analysis

The electroencephalogram (EEG) was recorded from 26 scalp positions, using Ag/AgCl electrodes attached to an elastic cap (Easy Cap, see Figure 13). All electrodes were referenced online to the left mastoid and then digitally re-referenced offline to the average of the left and right mastoids. To detect blinks and lateral eye-movements for later correction, additional electrodes were placed above and beneath the right eye and at the outer canthi of both eyes. All electrode impedances were kept below 5 k Ω . The EEG was amplified and analog filtered by Grass Model 12 amplifiers with bandpass of 0.01 to 30 Hz. Data were digitized at a sampling rate of 200 Hz and saved to a computer along with event timing information using the IWave software package (Version 6.5, InstEP Systems).

Data were analyzed using the MATLAB EEGLAB toolbox (Delorme & Makeig, 2004) and its ERPLAB plugin (Luck & Lopez-Calderon, 2011). Ocular artifacts were identified and

removed using independent component analysis (Jung et al., 2000). Trials were removed from analysis if the peak-to-peak voltage between 100 ms pre-stimulus and 200 ms post-stimulus exceeded 150 μ V for either of the EOG channels, or between 100 ms pre-stimulus and 1500 ms post-stimulus for any of the 28 EEG channels, resulting in an average loss of 4.6% of the data per participant. For each trial, event-related potentials were computed at each electrode time-locked to the onset of the novel verb and the onset of the disambiguating noun. Averages were baselined on the 100 ms before critical word onset.

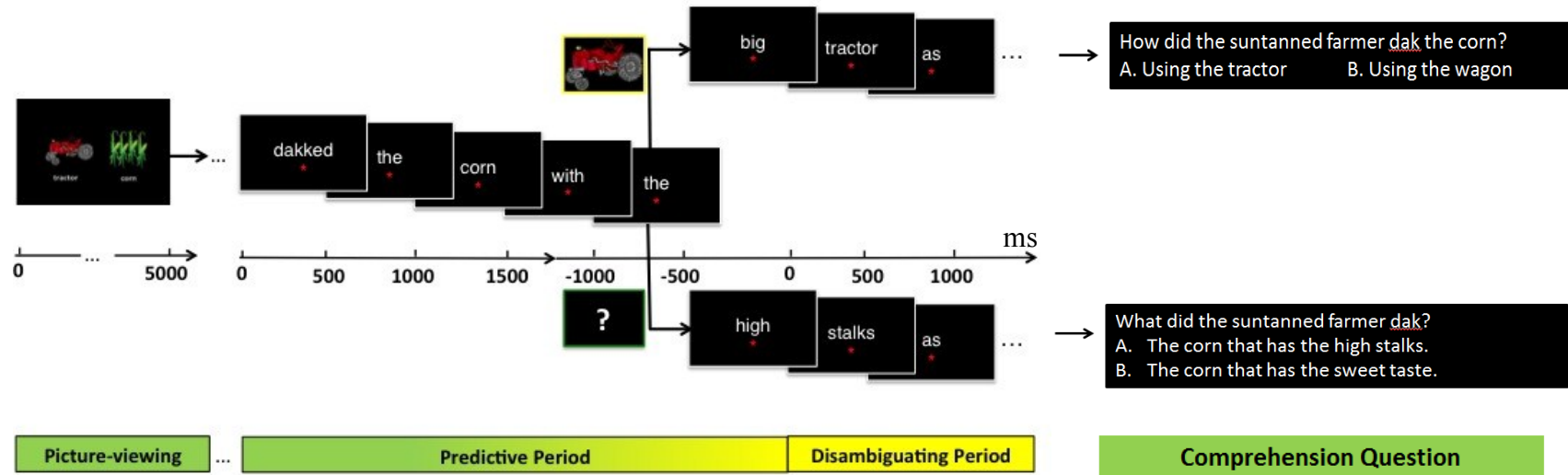


Figure 12: Experiment 5 & 6 Example of an ambiguous training trial.

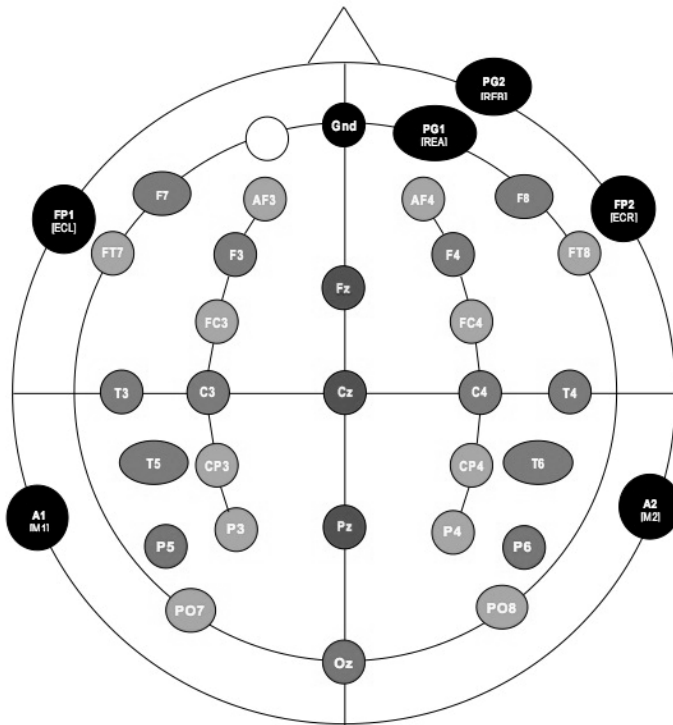


Figure 13: Schematic map of electrodes layout. Electrodes are divided into clusters according to 1). anteriority: frontal: AF3/4, F3/4, F7/8, FT7/8 and FZ; central: FC3/4, C3/4, T3/4, CP3/4 and CZ; 2). posterior: T7/8, P3/4, P5/6, PO7/8, PZ and OZ; 2) laterality: midline: FZ, CZ, PZ and OZ; left: AF3, F3, F7, FT7, FC3, C3, CP3, T3, T4, P3, P5, PO7; right: AF4, F4, F8, FT8, FC4, C4, CP4, T4, P4, P6 and PO8.

3.2.2 Results

3.2.2.1 Behavior

Participants' accuracy on comprehension questions was high overall. The mean accuracy for each of the four training conditions was above 90%, suggesting that participants were paying attention to the critical nouns despite the unknown verbs. A two-way Analysis of Variance (ANOVA) with the factors Attachment (modifier vs. instrument) and Ambiguity (ambiguous vs. unambiguous) showed that participants were more accurate for instrument sentences (97%) than for modifier sentences (92%, $F(1,31) = 28.2, p < 0.001$). There was no main effect of ambiguity ($F < 1$) or interaction between the two factors ($F(1,31) = 1.82, p > 0.1$). The fact that participants were more accurate in answering questions about items with instrument-trained verbs than ones with modifier-trained verbs suggests that their short-term memory for the instrument critical nouns was better than for the modifier critical nouns. Participants' behavioral performance did not differ across the two training blocks ($F < 1$). Therefore, any differences in the ERP responses across blocks were not likely to be due to changes in attention.

3.2.2.2 Post-experimental debriefing

All participants were naïve about the statistical training purposes of the experiment at the beginning of their training. In the post-experiment questionnaire, none of the participants reported that they noticed that the sentences were ambiguous except for the unknown verb meaning. Most participants (25 out of 32) reported that the picture-viewing session helped their sentence comprehension and the majority of these participants (20 out of 25) reported that they had tried to predict when the pictured objects would be mentioned during the sentence. Only 6 (out of 32) participants claimed they noticed any patterns in the linking of particular verbs with

particular sentence structures, suggesting that changes in the ERPs across the blocks (reported below) occurred mostly without participants' conscious awareness of the training distributions.

3.2.2.3 Event-related potentials

ERP components were analyzed with repeated measures of ANOVA on mean amplitude values calculated across particular time windows. N400 was measured during the 300-500 ms after the onset of the critical nouns and late positivity was measured during 500-1300 ms time window. For electrode effects involving more than two levels, the Greenhouse–Geisser correction (Greenhouse & Geisser, 1959) was applied, to avoid Type I errors due to violation of the sphericity assumption.

(A). ERPs at critical nouns

The critical noun in the PP disambiguated the sentence structure in the ambiguous conditions, so its ERP responses should reflect whether the disambiguation matched any expectations participants had developed based on the novel verb earlier in the sentence.

(a). Ambiguity ERP effects overall

Ambiguous sentences required participants to actively resolve their structure based on the meaning of the critical noun, while the unambiguous sentences provided other disambiguating cues earlier in the sentence. ERPs (N=32) to the critical nouns in ambiguous and unambiguous sentences collapsing across the two training blocks and the two attachment types are shown in Figure 14. The critical nouns in the ambiguous conditions elicited a larger late positivity in the ambiguous conditions than in the unambiguous conditions, which was quantified as the mean between 500 and 1300 ms after the onset of the critical noun. A repeated-measures ANOVA with

two levels of ambiguity (ambiguous vs. unambiguous) and 17 levels of centro-posterior scalp electrodes found a reliable ambiguity effect ($F(1,31) = 5.36, p < 0.05$), suggesting that the effect was a modulation of P600 amplitude due to more difficulty parsing the relationships between words in the ambiguous sentences, consistent with previous results showing that P600 is sensitive to parsing difficulty (Friederici, Hahne, et al., 2002; Kaan et al., 2000; Kaan & Swaab, 2003).

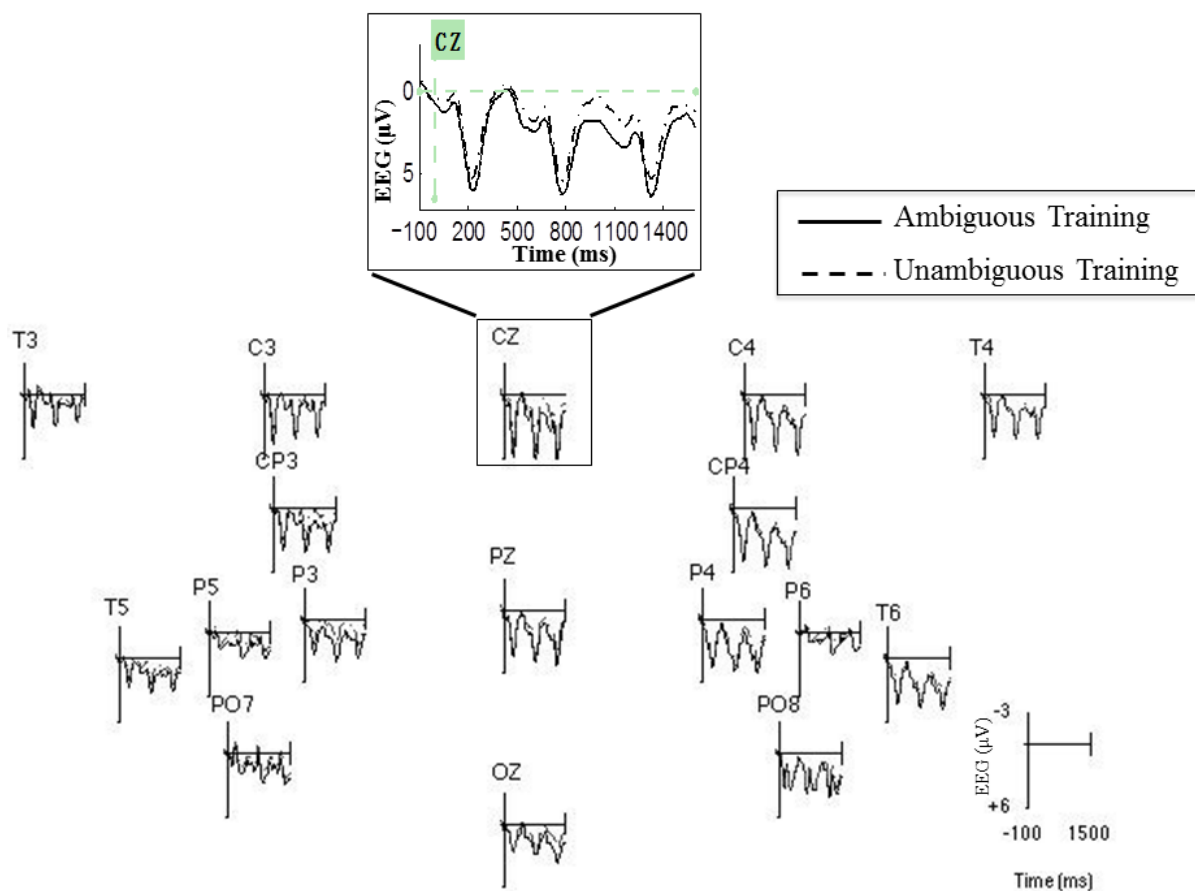


Figure 14: Experiment 5 Ambiguity Effect. Grandmean ERPs elicited by the critical nouns (e.g. *tractor* and *stalks*) at the 17 centro-parietal electrode sites in the ambiguous (solid line) and the unambiguous (dashed line) training conditions in Experiment 5. The waveform at CZ channel was enlarged, which showed the critical nouns in the ambiguous *with* structures elicited a larger P600 than those in the unambiguous *using* or *that has* structures.

(b). Attachment ERP effects overall

N400 amplitude has been found to be smaller for words that are highly predictable from a strongly constraining sentence context (Federmeier & Kutas, 1999; Federmeier et al., 2007). In our experiment, if participants actively predicted an instrument based on either the newly learned instrument bias for a particular novel verb or on the presence of the disambiguating cue *using* in the sentence, at the critical noun position they would expect the particular instrument word they had seen associated with the pictured instrument that was presented before the sentence.. In contrast, in ambiguous sentences with modifier-trained novel verbs and in unambiguous sentences with *that has*, participants should not be able to make any strong prediction about the critical noun. Therefore, instrument critical nouns were predicted to elicit smaller N400s than critical modifier nouns. There was another less interesting reason for instrument critical nouns to elicit smaller N400s than modifier critical nouns, which is that instrument words were repeated from the picture phase to the sentence-reading phase of a trial. N400 amplitude reduction effects for both of these reasons should be equivalent in ambiguous and unambiguous conditions, since the instrument word is repeated in both and instruments should be more predictable in both, based on newly learned verb bias in the ambiguous conditions and on the inclusion of *using* as well as possibly newly learned instrument verb bias in the unambiguous conditions.

ERP brain potentials to the critical nouns averaged across both blocks are plotted in Figure 15. The instrument nouns (e.g., *truck*) elicited reduced N400 amplitude relative to the modifier nouns (e.g., *stalks*). N400 amplitude was measured from 300 ms to 500 ms after the onset of the critical nouns. ANOVA with the factors of attachment (modifier *vs.* instrument), ambiguity (ambiguous *vs.* unambiguous) and 28 levels of electrode sites revealed a main effect

of attachment ($F(1,31) = 4.59, p < 0.05$) but no main effect of ambiguity ($F < 1$), nor any interaction ($F = 1$).

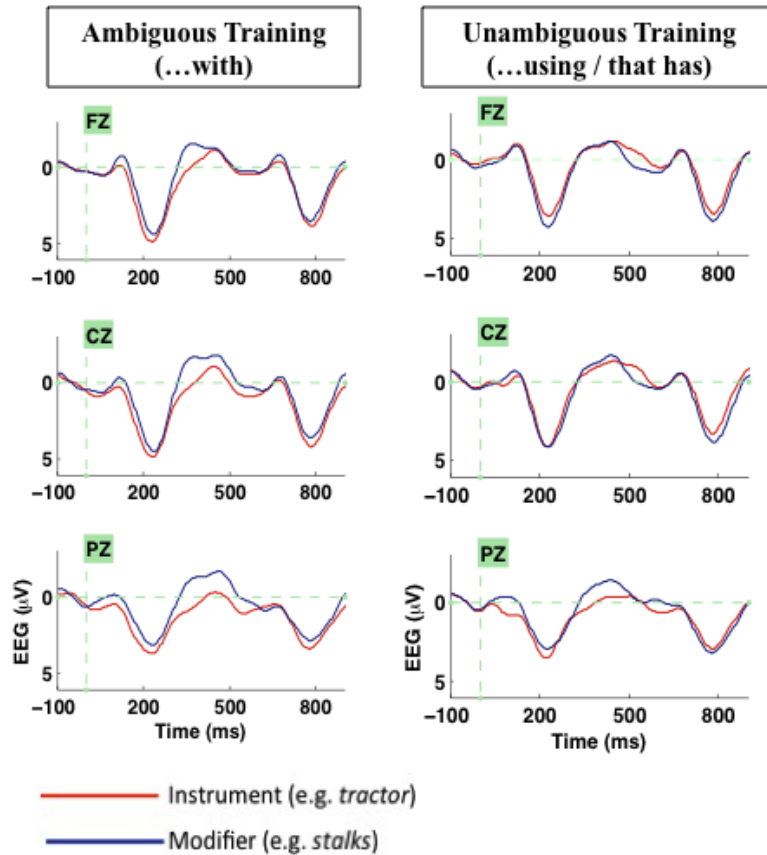


Figure 15 : Experiment 5 Attachment effect. Grandmean of ERPs elicited by the instrument (red line) and modifier (blue line) critical nouns at the midline electrode sites in the ambiguous and the unambiguous training conditions. Analysis over 28 scalp electrodes indicate an overall reduction of N400 effect elicited by the instrument critical nouns in relative to the modifier critical nouns. The modulation of N400 effect is mainly observed in ambiguous training conditions, rather than the unambiguous training conditions.

(c). Attachment ERP effects across the blocks

We examined changes in attachment effects from the first block to the second block of training. Figure 16 shows a reduced N400 effect of attachment type in the second block compared to the first, and consistent with that an ANOVA revealed a significant interaction

between attachment type and block ($F(1,31) = 6.37, p < 0.05$). This interaction suggests that the smaller N400 observed for instrument nouns than for modifier nouns when collapsing across blocks cannot be entirely due to effects of word repetition on N400 amplitude. In both blocks, the instrument nouns in the sentences repeated picture label words from the pre-sentence pictures, so if the N400 reduction in response to those nouns was entirely due to repetition effects, it should have been observed in both blocks. The fact that the difference diminished in the second block suggests that it was related to increasing experience with the types of sentences the verbs appeared in. It should be easier to associate instrument-bias with verbs because instruments are arguments of the verb while modifiers are not. Instruments quickly became highly predictable in sentences with instrument-trained verbs, and by the second training block modifiers grew less unexpected after modifier-trained verbs.

There is another difference between the waveforms in the first and second blocks that may contribute to the apparent diminishment of the N400 effect in the second block. In block 2, modifier nouns elicited a small long-lasting positivity compared to instrument nouns, which was visible predominantly at frontal sites. However, it is possible that it was not actually limited to frontal sites but rather that it overlapped in time with the N400 effect so that the two components tended to cancel each other out in the waveforms at centro-posterior sites. When the frontal positivity was quantified as the mean amplitude 500-1300 ms after the onset of the critical noun, it showed no reliable effect of attachment type ($F(1,31) = 2.26, p > 0.1$).

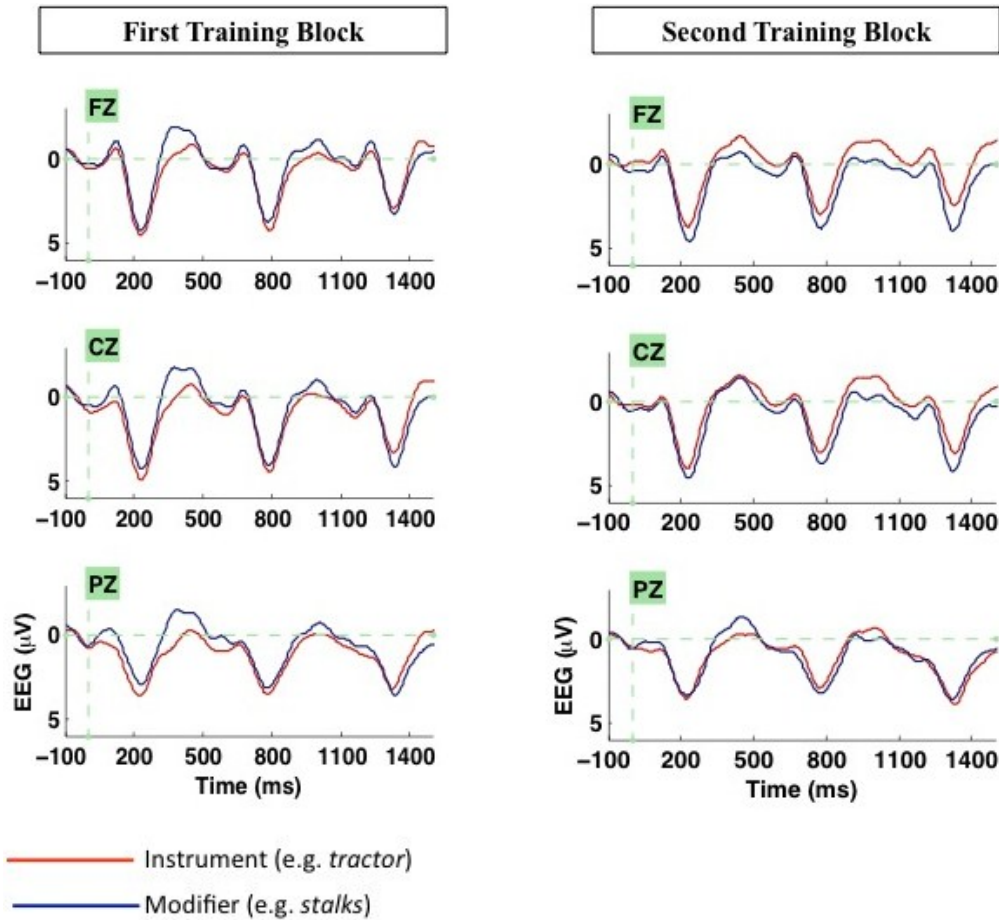


Figure 16: Experiment 5 Attachment effect shifted across blocks. Grand mean ERPs elicited by the instrument (red line) and modifier (blue line) critical nouns at the midline electrode sites in the first block and the second training blocks. The N400 attachment effect was larger in the first block than in the second block. Note that the decrease in the N400 effect that the second block is replaced by the emergence of a centro-frontal positivity.

(B). Predictive effects in the ERPs before the arrival of disambiguation

The monitoring of EEG throughout the sentence makes it possible to examine the responses to words before the critical noun to look for evidence of the development of predictions before disambiguation arrives. As people learned the structural preference of an instrument-biased verb, they should have started predicting that the instrument noun they saw labeling a picture just before the sentence would follow the verb in an instrument phrase. In

contrast, when they saw a modifier-biased verb they could not make any such specific prediction. The sentences in the instrument and modifier training conditions were identical until the adjective modifying the critical noun appeared, and those were designed to be uninformative about whether the noun would be an instrument or a modifier. Therefore the only difference between the instrument and modifier conditions before the arrival of the critical nouns was people's previous experience with the kinds of sentences particular verbs had appeared in.

The greater predictability of the critical noun after instrument-trained verbs was true for both the ambiguous and unambiguous conditions, but it was not possible to make the unambiguous sentences identical to one another before the critical noun since different words were required to accomplish the disambiguation toward the two structures (i.e., *using* vs *that has*). Thus, different responses to these different kinds of words would be completely confounded with any differences in prediction based on newly acquired verb biases, so only the ambiguous conditions were examined at earlier word positions.

The responses to the words before the critical noun are shown in Figure 17 for the two training blocks, time-locked at the onset of the determiner *the*, which was 1000 ms before the onset of the critical noun. In the second block, sentences with instrument-trained verbs elicited a positivity starting about 500 ms before the onset of the critical noun, which was at the onset of the adjective, and peaking around 200 ms later (-300 ms on the X-axis in Figure 17. Mean amplitude was averaged across the 500 ms before the onset of the critical nouns (i.e., from -500 ms to 0 on the X-axis in Figure 17). An ANOVA with the factors attachment type (instrument vs. modifier), training block (first vs. second), and 17 posterior channels revealed a significant interaction between attachment type and block ($F(1,31) = 4.83, p < 0.05$), and followup comparisons found a reliable attachment type effect ($F(1,31) = 4.12, p = 0.05$) only in the

second block, suggesting that participants gradually learned to make specific predictions only for instrument-trained verbs.

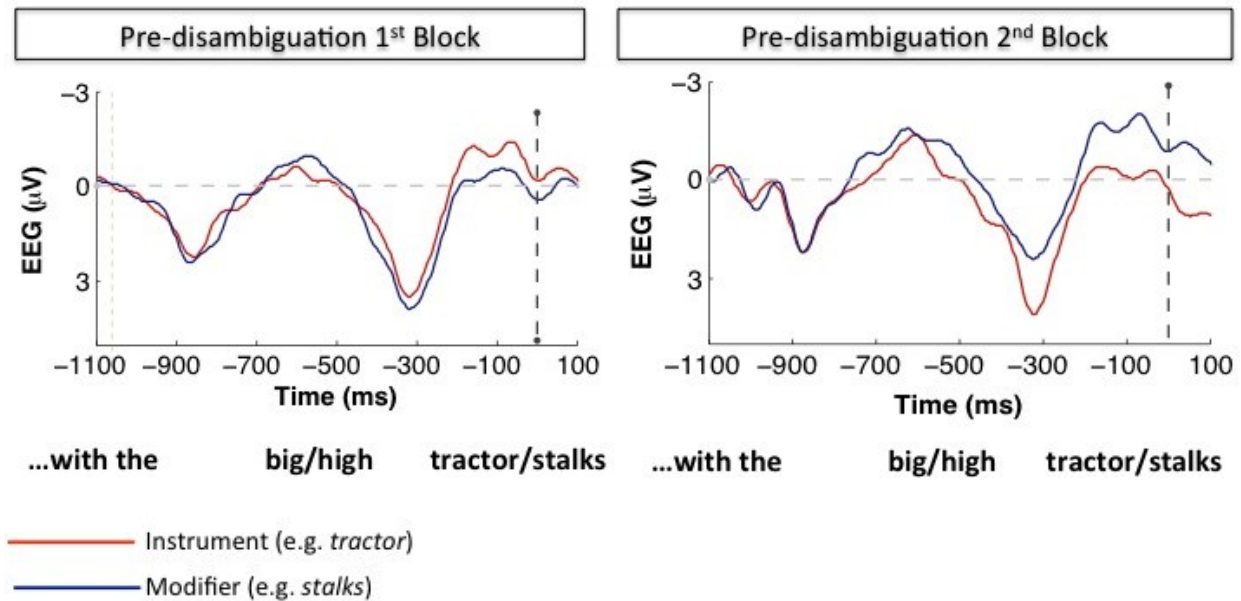


Figure 17: Experiment 5 Predictive effect. Grand mean ERPs averaged over the three central-posterior electrodes (P3, PZ and P4) 1000 ms before the onset of the critical nouns (the zero point on the x-axis) in instrument (red line) and modifier (blue line) training. EEG was baselined 100 ms before the onset of the determiner *the* (-1000 ms on the x-axis). The adjectives were presented 500 ms after the determiner. The difference between the two conditions grew larger over the course of training with a larger positivity elicited by the instrument-trained structure than the modifier-trained structure in the second block.

3.2.3 Discussion

The results provide electrophysiological evidence for rapid learning of verb bias from the distribution of nonsense verbs in sentences of different types without support from verb semantics. This conclusion is supported by two key findings: 1) greater predictability of instrument critical nouns after instrument-trained verbs led to smaller N400 amplitude for instrument nouns than for modifier nouns, and 2) there were differences in the waveforms before

the disambiguating critical nouns that suggested more prediction in sentences with instrument-trained verbs by the second training block.

First of all, the instrument critical nouns elicited a smaller N400 compared to the modifier critical nouns, probably reflecting an expectancy effect. In the unambiguous condition, the transparent *using* and *that has* were sufficient cues for an upcoming instrument or modifier noun, but in the ambiguous condition, readers made predictions relying on just the particular verbs' newly learned biases. The picture-viewing at the beginning of each trial presented pictures of the instrument and the direct object, and therefore provided a more constraining context for predictions about specific instrument nouns while modifier nouns were much less predictable. The reduced amplitude in N400 to the instrument nouns probably reflected easier semantic processing of words that confirmed readers' expectations. The findings are consistent with previous studies showing that N400 is reduced for predictable words in strongly-constraining sentences contexts compared to less predictable words in weakly-constraining sentence contexts (Federmeier et al., 2007; Kutas & Hillyard, 1984; Wlotko & Federmeier, 2007).

As mentioned above, the difference in degree of contextual constraint provided by instrument- and modifier-training conditions was confounded with the fact that only instrument critical nouns were explicitly presented visually during picture-viewing, while the modifier critical nouns were more implicit embedded features of the direct object nouns in the pictures. N400 has been found to be modulated by both word repetition (Van Petten et al., 1991) and by semantic priming for both words and pictures (Federmeier & Kutas, 2001; Ganis, Kutas, & Sereno, 1996), so it is possible that the N400 reduction for instrument nouns in the sentences was a result of repetition and/or semantic priming from the earlier pictures (Hamm, Johnson, & Kirk, 2002; Holcomb & McPherson, 1994). However, the fact that the differences in N400 amplitude

between instrument and modifier nouns decreased across the two training blocks suggests that more was going on. If the effects were due entirely to repetition and/or semantic priming from the pictures and their labels, it should have just as strong in the second block. The fact that it was not suggests that the N400 effect was due at least in part to learning the biases of the verbs.

The emergence of a late positivity at frontal sites in the response to the modifier nouns in the second block suggested a shift in the way the sentences were processed as learning proceeded. Previous studies on L2 learning (McLaughlin et al., 2010; Osterhout et al., 2008, 2006) and artificial language learning (Morgan-Short et al., 2010) have observed shifts across learning from N400 effects to P600 effect in response to morphosyntactic violations. The late positivity seen here did not have the usual centro-parietal scalp distribution of P600 effects so it may be another ERP component, but it is also possible that the observed scalp distribution resulted from overlapping N400 and P600 effects canceled each other out at posterior sites. Further investigation is necessary to determine the nature of this effect.

Another intriguing finding was the emergence in the second block of an ERP effect beginning 500 ms before the critical noun was presented, suggesting that participants were learning to make predictions after the instrument-trained verbs. The temporal distribution in relative to the onset of the adjectives resembles the P2 response, which has been linked to effects of target expectancy in selective attention tasks (e.g. Luck & Hillyard, 1994). In language processing, P2 amplitude was modulated by sentential contextual constraints and semantic concreteness. Larger P2 amplitude was elicited by stronger expectations for the upcoming words based on the context either provided by more constraining sentential context or a more concrete adjectives (Huang, Lee, & Federmeier, 2010; Wlotko & Federmeier, 2007). The timing and posterior scalp distribution of this effect was also consistent with some preliminary MEG

findings (Dikker, submitted) about activity during the pre-stimulus interval when participants were anticipating a particular word. Electromagnetic oscillations in the theta range (4-7 Hz) were enhanced in posterior brain regions 150 ms before a predictable target word following a strongly constraining context, compared to words following less constraining contexts. In the present study, readers learned to predict upcoming nouns when their previous experience with the particular verb in the sentence taught them that particular kinds of nouns (i.e., instruments) were predictable after particular kinds of verbs.

3.3 Experiment 6: Neural stages of novel verb bias learning

Experiment 6 doubled the number of training trials, which allowed us to try to replicate the major findings in Experiment 5 for learning after two blocks as well as to examine verb bias learning after more exposure. A final test phase using the trained verbs in globally ambiguous sentences was also added, to determine whether participants would use their newly learned verb biases to resolve ambiguities.

Another question of interest in this experiment concerned possible individual differences in learning and applying newly learned verb biases. Preliminary findings from a previous ERP study (Qi et al., 2010) revealed potentially relevant differences between pure right-handers and mixed right-handers in their sensitivity to the biases of familiar verbs during sentence reading, particularly when critical words were initially presented to the left hemisphere in a visual half-field paradigm. Experiment 6 tested whether differences in familial handedness also influenced the acquisition and application of biases for novel verbs, during both online reading and an offline behavioral task.

3.3.1 Methods

3.3.1.1 Participants

Twenty-four right-handed native English speakers (13 females and 11 males) at the University of Illinois, Urbana-Champaign participated in the experiment. Their mean age was 21 years, ranging from 18 to 25 years. Thirteen participants reported having left-handed family members. Their participation was compensated with \$8/hour.

3.3.1.2 Materials and Design

The experiment consisted of a training session with EEG recording followed immediately by a behavioral testing session using a sentence-picture matching task. The training stimuli were the same as in Experiment 5 but each sentence was presented twice to double the number of training trials. The third and fourth training blocks repeated the items from the first two blocks but presented them in reversed order.

The picture-matching task consisted of 24 sentences and 24 pairs of corresponding pictures. The sentences were carefully constructed so that they were globally ambiguous, i.e., the critical noun could be either an instrument or a modifier. For example, in (4), the *with*-PP object noun *low score* could either attach to the verb *dak* as an instrument or to the direct object noun *student* as a modifier.

(4) *The angry teacher daked the student with the low score.*

Two new novel verbs were introduced in the behavioral test in addition to the four trained verbs. Each participant read 24 test sentences and was tested on all four trained verbs plus the two untrained control verbs.

3.3.1.3 Procedure, electroencephalogram recording and analysis

The setup was similar to Experiment 5 during the training session, except that a new 22-inch wide-screen LCD monitor was used, and words were presented in black on a grey background to eliminate glare on that monitor. During the picture-matching test, participants first viewed a pair of pictures side-by-side until they either pushed a button to indicate they were ready to see the sentence or 5 seconds had elapsed, after which a globally ambiguous sentence appeared below the two pictures, as illustrated in Figure 18. A response was required within the next 5 seconds and feedback was provided regarding whether the choice was consistent with the training pattern for that verb⁴. The screen position of the pictures representing instrument and modifier interpretations was counterbalanced across items.

The EEG recording and analysis procedures were the same as for Experiment 5. Artifact rejection resulted in an average loss of 5.9% of trials per participant. The experiment lasted about 75 minutes.

⁴ Preliminary (N=16) results in a separate experiment with verbs that were trained in an equi-bias training paradigm (i.e., they appeared in an equal number of sentences with the four kinds of structure) showed that providing feedback during the picture-matching task did not induce any immediate verb bias learning effect. In addition, participants' response to sentences containing equi-trained verbs were not different from those containing untrained verbs.

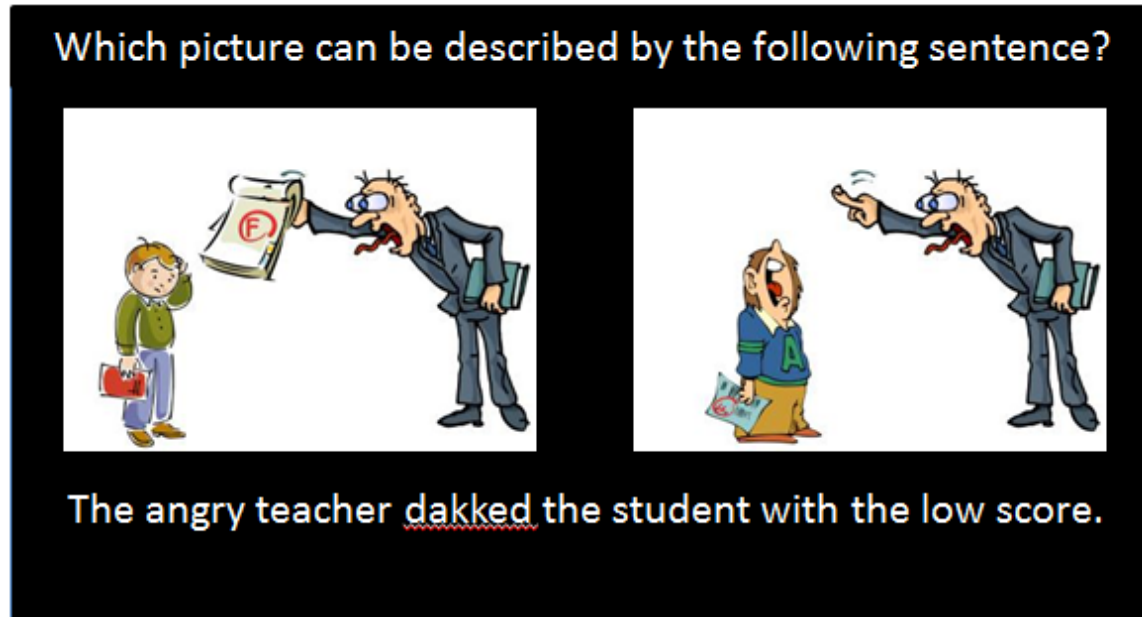


Figure 18: Experiment 6 Example of picture-matching task. The left picture represents an instrument interpretation and the right picture represents a modifier interpretation of the sentence.

3.3.2 Results

3.3.2.1 Behavioral results

(A). Comprehension accuracy in training

Accuracy on the comprehension questions was high overall, with accuracy above 95% in all training conditions. As in Experiment 5, accuracy was slightly higher for instrument-trained verbs (97%) than for modifier-trained verbs (95%, $F(1,23) = 5.60, p < 0.05$). There were no differences in accuracy between the first (95%) and second (96%) halves of training ($F(1,23) = 1.94, p > 0.1$).

(B). Behavior in the post-training picture-matching task

Table 4 shows the proportion of instrument choices for each training condition for the tested verbs. Mean proportions of instrument choices were calculated for each participant for instrument-trained, modifier-trained, and untrained control verbs and analyzed across participants using paired-sample *t*-tests. There were no differences between the untrained control verbs and any of the trained verbs in any of the training conditions (all *t*'s (23) < 1.2, *p*'s > 0.2). Thus, the training was not strong enough to influence readers' explicit decisions about globally ambiguous sentences. However, when participants were divided based on familial left-handedness, pure and mixed right-handers showed different patterns, illustrated in Figure 19. Pure right-handers (N=13) showed a reliable preference for instrument interpretations in the trials with verbs that were trained in ambiguous instrument sentences compared to untrained control verbs (*t* (12) = 2.62, *p* < 0.05) and a similar trend in trials with verbs that were trained in unambiguous instrument sentences (*t* (12) = 1.87, *p* = 0.09). In contrast, none of the trained verbs differed from the control verbs for mixed right-handers (N=11) (*ts* < 1.69, *ps* > 0.1). The only effect that approached significance in mixed right-handers was a trend toward more instrument choices in sentences with verbs that were trained in ambiguous modifier sentences than in sentences with verbs trained in unambiguous modifier sentences (*t* (10) = 1.88, *p* = 0.09). A between-group two-sample *t*-test confirmed that the difference between the two groups was driven by participants' behavior for the verbs trained in ambiguous instrument sentences (*t* (21.8) = 1.8, *p* = 0.09).

Table 4: Experiment 6 Picture-matching task result. Mean proportions of instrument choices (standard errors in parenthesis).

Training Conditions of the Verb	Proportion of Instrument Choices
I-AM	0.60 (0.06)
M-AM	0.59 (0.04)
I-UN	0.51 (0.06)
M-UN	0.52 (0.05)
Control	0.54 (0.05)

Training conditions: I: instrument; M: modifier; AM: ambiguous; UN: unambiguous

3.3.2.2 Post-experimental Debriefing

As in Experiment 5, none of the participants reported noticing that the training sentences were ambiguous other than having a novel verb. Seventeen participants (out of 24) reported that the picture-viewing was helpful for their comprehension of the subsequent sentences, and 12 of those claimed they tried to predict when the pictures would be mentioned in the sentence. After four blocks of training, one-third of the participants (8 of 24) reported noticing that particular verbs co-occurred with particular sentence structures. Thus, doubling the amount of training slightly increased the proportion of participants who became aware of some contingencies in the training (8 out of 24 vs. 6 out 32).

When asked to describe their experience in the end-of-session picture-matching task, all 24 participants noticed the ambiguity of the testing sentences, and 11 claimed they tried to recall the training pattern in resolving the ambiguity.

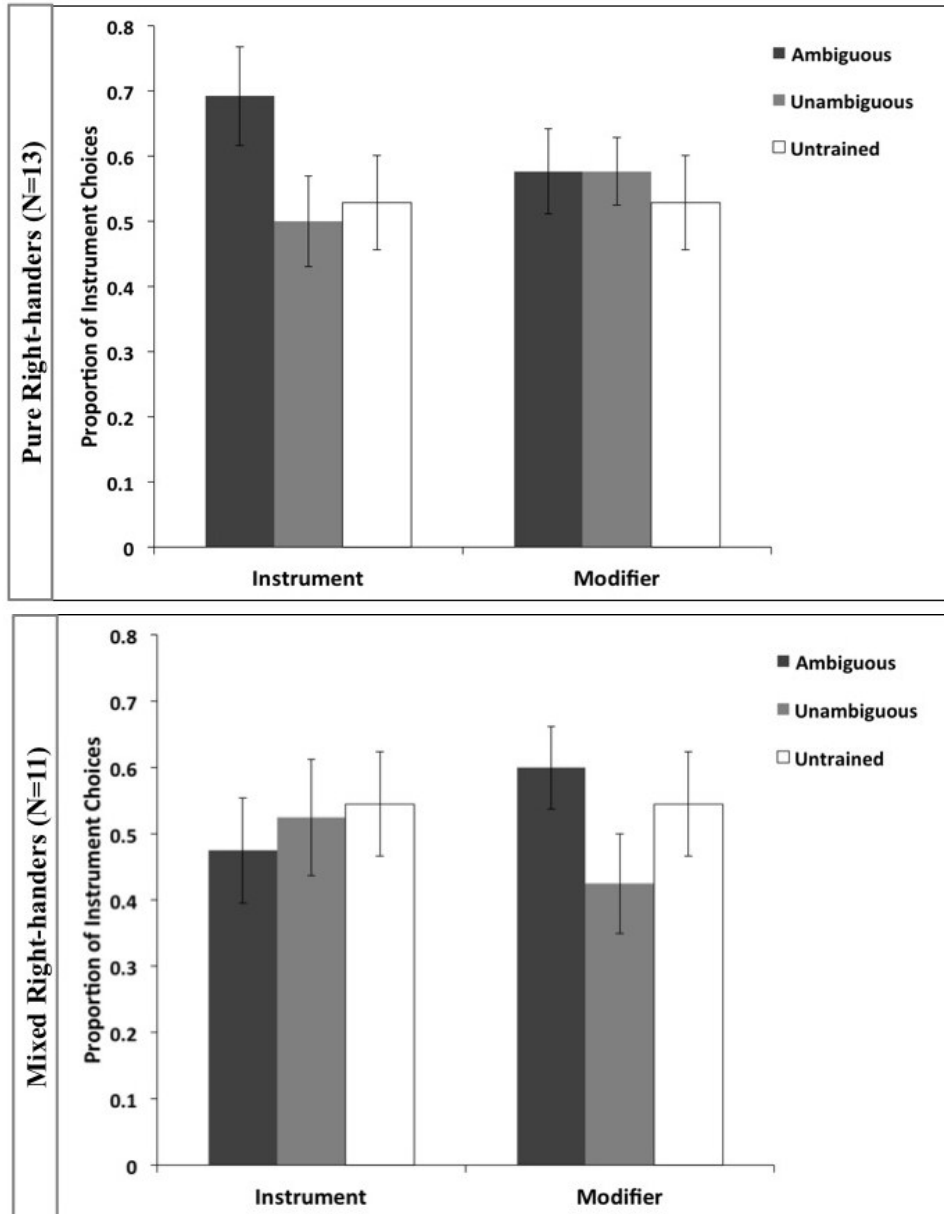


Figure 19: Experiment 6 Individual differences in behavioral results for pure and mixed right-handers. Mean proportions of instrument choices in the picture-matching task. The results for untrained verbs (white bars) are repeated in each half of the figure for each group for ease of comparison. Error bars represent standard errors. Only pure right-handers showed a reliable training effect and only in the instrument-ambiguous condition.

3.3.2.3 Event-related potentials

(A). ERPs in the first half of training

Results will first be presented for just the first two blocks of training, since that part of the experiment was identical to Experiment 5, in order to determine which patterns replicated across studies with the same amount of training. Visual inspection of the ERP waveforms shown in Figures 20, 21 and 23, compared to Figures 14, 16, and 17 from Experiment 5, reveals considerable similarity across studies, so the data from the first two training blocks were combined with the data from Experiment 5 to increase statistical power ($n=56$, $32+24$). The same time windows were used to quantify ERP components. Whenever an interaction was found between an experimental factor and electrode site, follow-up analyses included laterality (left, midline, and right) and anteriority (frontal, central and posterior) factors to investigate the scalp distribution of the effect.

(a). Ambiguity ERP effects for the critical nouns

As in Experiment 5 (compare Figures 14 and 20), the critical nouns (e.g., *tractor* and *stalks*) elicited a larger centro-posterior positivity in ambiguous sentences than in unambiguous sentence, reflected in a main effect of ambiguity in an ANOVA including the 17 centro-posterior scalp electrodes ($N=56$, $F(1,55) = 4.07$, $p < 0.05$). This is most likely a P600 effect reflecting greater difficulty parsing the relationships between words in the ambiguous structure.

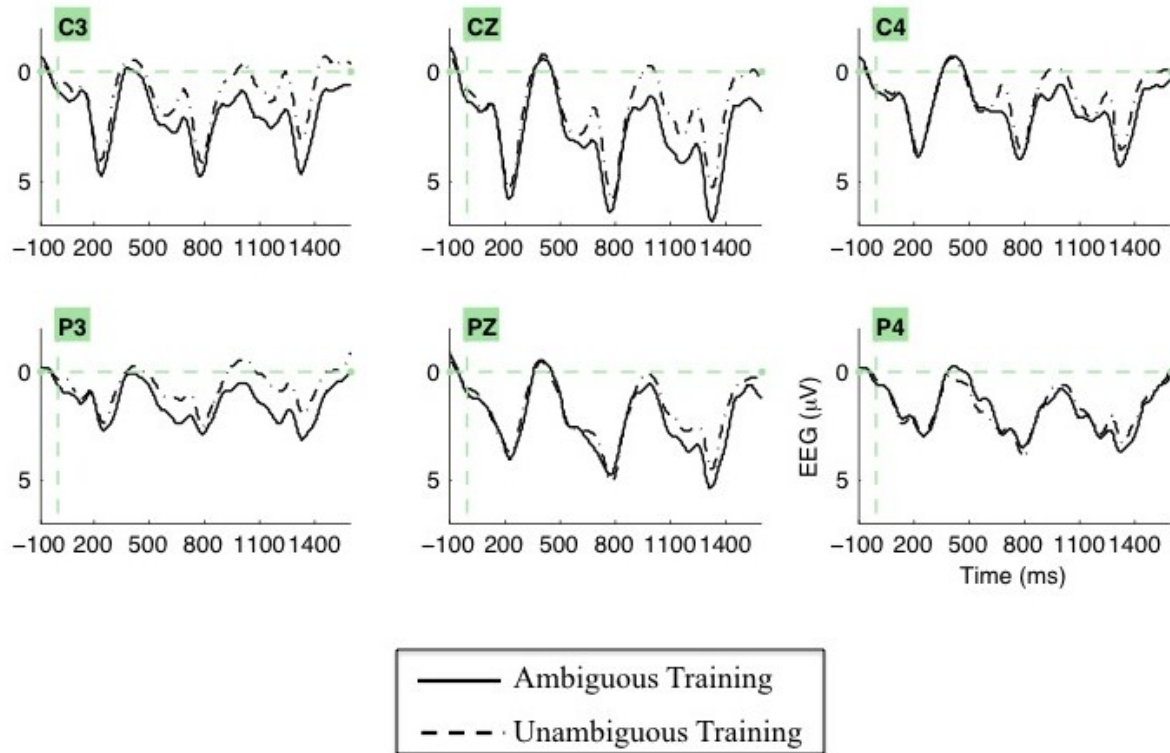


Figure 20: Experiment 6 Ambiguity effect. Grand mean ERPs elicited by the critical nouns (e.g. *tractor* or *stalks*) at the six centro-parietal electrode sites in the ambiguous (solid line) and the unambiguous (dashed line) training conditions during the first two training blocks. The critical nouns in the ambiguous *with* structures elicited a larger P600 than those in the unambiguous *using* or *that has* structures.

(b). Attachment ERP effects for the critical nouns

Figure 21A plots the overall attachment effect that the instrument critical nouns (e.g. *tractor*) elicited a smaller N400 compared to the modifier critical nouns (e.g. *stalks*). Statistical assessment of the N400 amplitude on 28 scalp channels confirmed the main effect of attachment on N400 amplitude ($F(1,55) = 11.397, p < 0.01$) with no interaction with ambiguity ($F(1,55) = 0.49, p > 0.4$). Attachment N400 effect significantly interacts with electrodes laterality ($F(2,110) = 5.88, p < 0.01$), reflecting the largest effect driven by the midline channels.

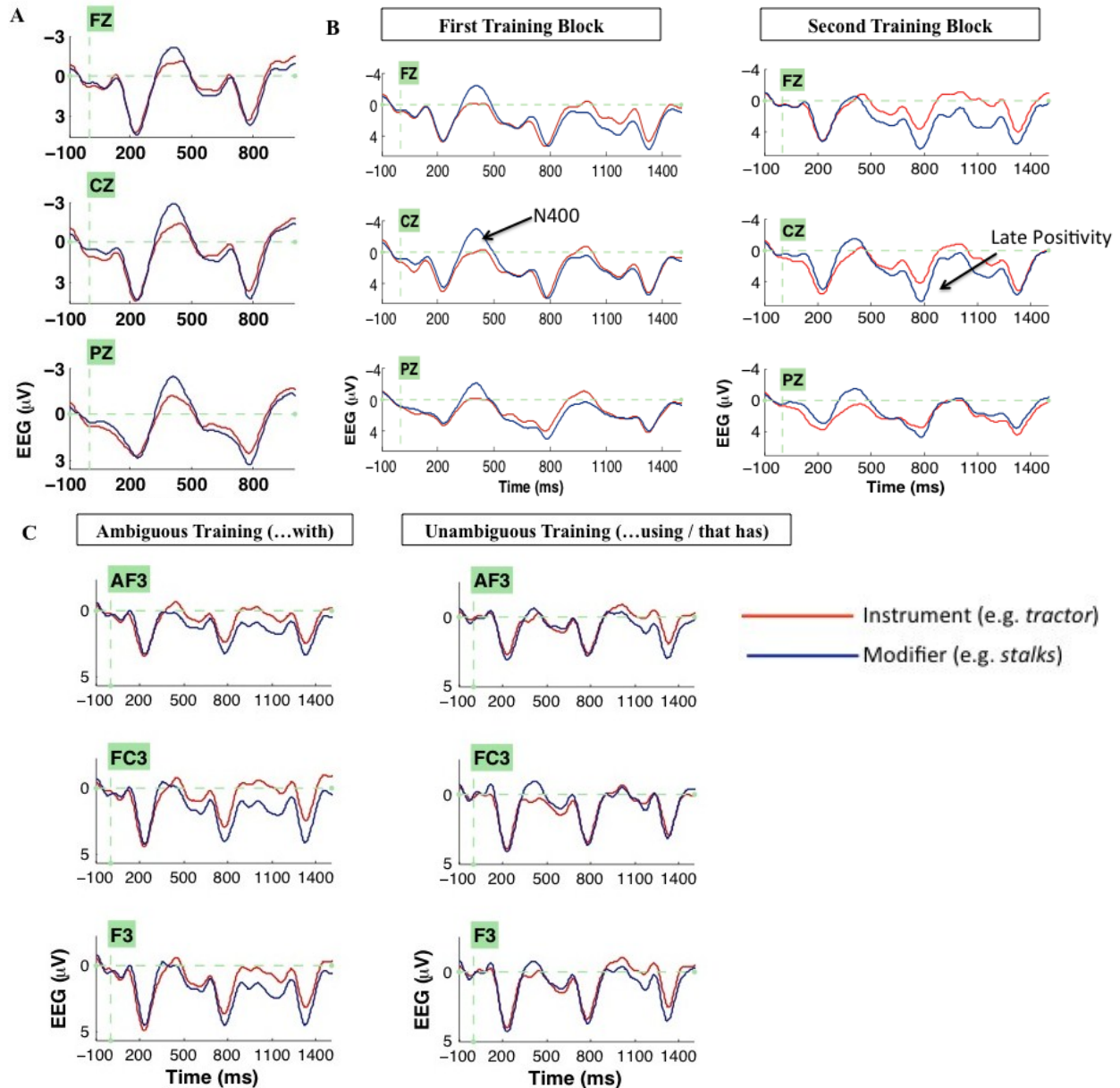


Figure 21: Experiment 6 Attachment effect. Grand mean ERPs to the critical nouns in the instrument (red line) and modifier (blue line) training conditions during the first two training blocks. (A) ERPs to the instrument and modifier critical nouns at the midline electrode sites averaged across two training blocks. The instrument critical nouns elicited a smaller N400 than the modifier critical nouns. (B). ERPs to the instrument and modifier critical nouns at the midline electrode sites separately for the first and second training blocks. The attachment effect shifted from N400 to a fronto-centrally distributed positivity. (C). ERPs to the instrument and modifier critical nouns at the left anterior sites in the second training block. Ambiguous training produced a larger attachment effect than unambiguous training.

Figure 21B shows the ERPs to the instrument and modifier critical nouns split by two training blocks. The N400 attachment effect was only reliable in the first block ($F(1, 55) = 11.37, p < 0.01$), but not in the second block ($F(1, 55) = 1.29, p > 0.1$). Comparing the attachment N400 effect in the first block with the second block yielded a non-significant interaction ($F(1, 55) = 1.91, p > 0.1$).

As in Experiment 5, the decrease in the N400 attachment effect in the second block was accompanied by the emergence of a late positivity at the central-frontal sites, with the instrument critical nouns inducing a smaller positivity than the modifier critical nouns. When tested across all 28 channels, this effect was marginally reliable only in the second block ($F(1, 55) = 2.93, p = 0.09$), with a significant interaction with site anteriority ($F(2, 110) = 4.05, p < 0.05$). When the analysis was restricted to the 18 central-frontal sites, the attachment effect was reliable ($F(1, 55) = 4.87, p < 0.05$) and there was a marginal interaction with training block ($F(1, 55) = 3.38, p = 0.07$), indicating the central-frontal positivity did not emerge until the second block. A significant three-way interaction laterality x ambiguity x attachment ($F(2, 110) = 4.75, p < 0.05$) reflected a stronger effect at the left side of the scalp in ambiguous conditions compared to the unambiguous conditions (Figure 21C). As was described for Experiment 5, it is not clear whether this late positivity was a P600, since it did not have the typical centro-parietal scalp distribution of that component. However, the scalp distribution could again be the result of a posteriorly distributed N400 effect overlapping with and cancelling out a P600 effect at posterior sites so that the positivity is only visible at anterior sites.

Topographic maps showing the scalp distributions of the N400 effect in the first block and the late positivity in the second block are provided in Figure 22A and Figure 22B,

suggesting that the subpopulations of neurons responding to the critical words changed as learning proceeded.

(c). Predictive component before the critical word

As in Experiment 5, there was again a larger positivity in the instrument-trained structure compared to the modifier-trained structure starting about 500 ms before the disambiguating information was provided by the critical nouns, as shown in Figure 23. For the 500 ms interval immediately before the onset of the critical nouns, a marginal interaction between training and block at the 17 posterior channels ($F(1,55)=3.80$, $p < 0.06$) suggested that participants' prediction strengthened across blocks. The training effect was reliable only in the second block ($F(1,55)=4.86$, $p < 0.05$; first block, $F < 1$).

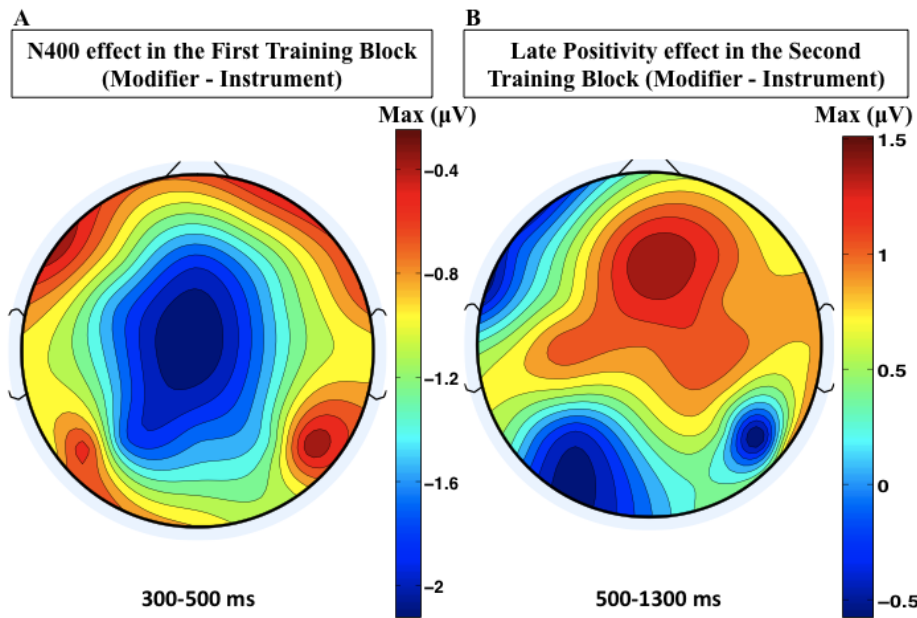


Figure 22: Experiment 5 & 6 Attachment effect shifts across training blocks. Topographic map (56 participants from both experiments) of the ERP difference between the instrument and the modifier critical nouns in the 300-500 ms interval in the first block (A) and the 500-1300 ms interval in the second block (B) with the difference amplitude values displayed by coloration.

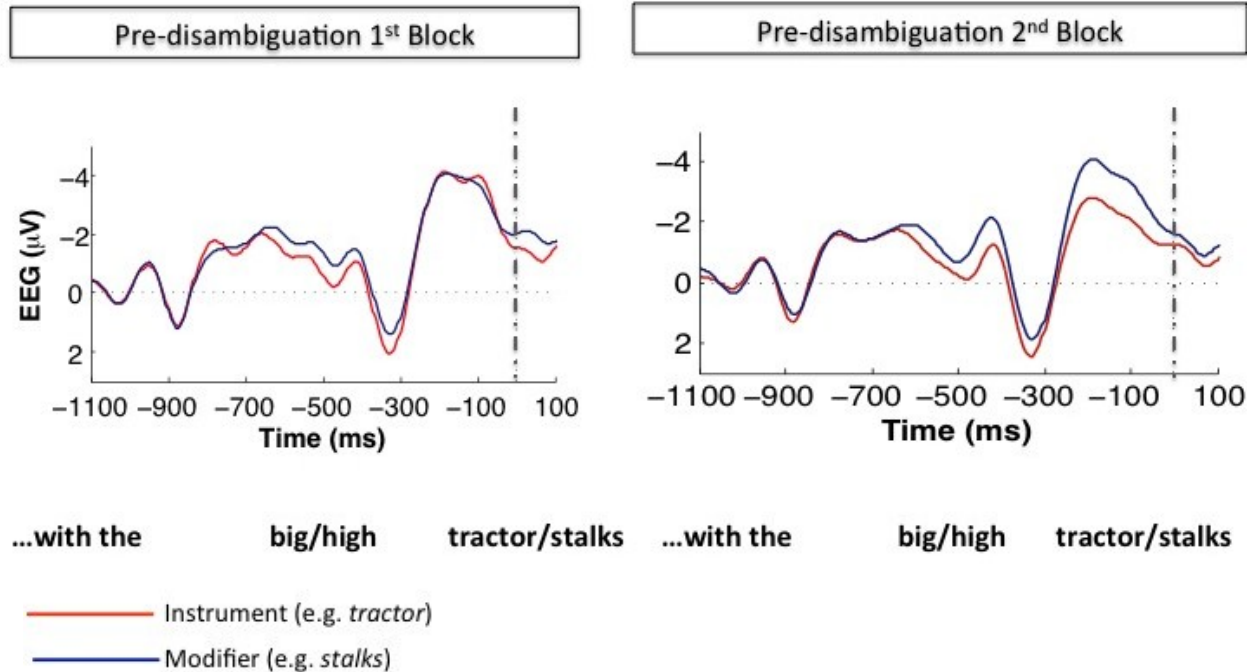


Figure 23: Experiment 6 Predictive effect. Grand mean ERPs averaged over the three central-posterior electrodes (C3, CZ and C4) 1000 ms before the onset of the critical nouns (the zero point on the x-axis) in instrument (red line) and modifier (blue line) training during the first half of training. EEG was baselined 100 ms before the onset of the determiner *the* (the -1000 ms point on the x-axis). The adjectives were presented 500 ms later after the determiner.

(B). ERPs in the second half of training

The training items presented in the first two blocks were repeated in reversed order for another two blocks of training. Thus, any additional training effects are confounded with possible repetition effects. Figure 24 shows ERPs to the critical nouns in the ambiguous and unambiguous conditions for the two repeated blocks. Instrument critical nouns elicited a smaller prolonged positivity than modifier critical nouns at most sites, which persisted beyond the 1500 ms shown in Figure 24. Analysis conducted on the 500-1300 ms interval for all 28 scalp sites revealed a significant attachment effect ($F(1,23) = 8.20, p < 0.01$) and a marginal ambiguity effect ($F(1,23) = 3.17, p = 0.09$). The latter was consistent with the ambiguity effect found in the first half of training. There was no interaction between attachment and ambiguity ($F < 1$). But

when ambiguous and unambiguous conditions were tested separately, the attachment effect was reliable only in the ambiguous training condition ($F(1,23) = 4.72, p < 0.05$) and not in the unambiguous training ($F(1,23) = 2.74, p > 0.1$).

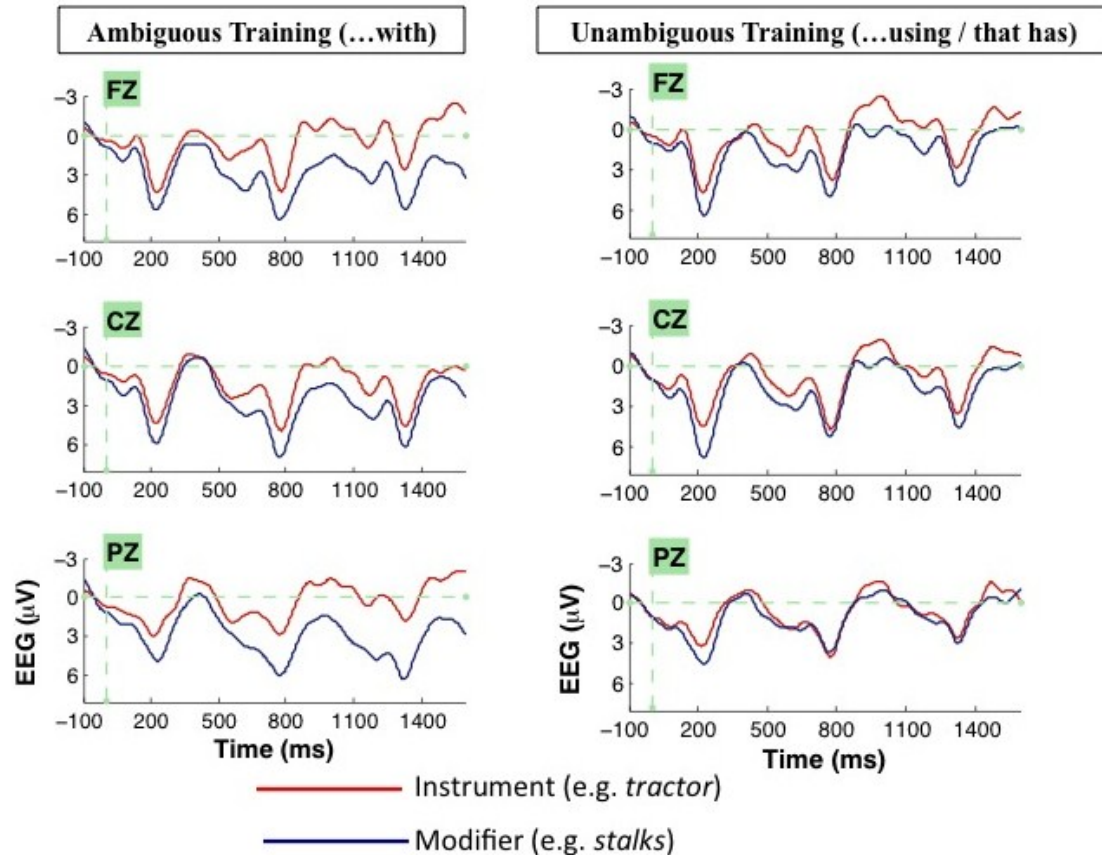


Figure 24: Experiment 6 Attachment effect during the second half of training. Grand mean ERPs elicited by the instrument (red line) and modifier (blue line) critical nouns at the midline electrode sites in the ambiguous and the unambiguous training conditions. Analysis over 28 sites indicated an overall reduction of P600-like positivity elicited by the instrument critical nouns relative to the modifier critical nouns. The modulation of the effect was mainly observed in ambiguous training conditions, rather than the unambiguous training conditions.

(C). Individual differences in familial left-handedness

Some behavioral differences in the post-training picture-matching task between the thirteen pure right-handers and the eleven mixed right-handers were described earlier, showing

better instrument-bias learning in the pure right-handers. ERPs elicited during the second half of training were also compared for the two groups.

Figure 25A shows the ERPs separately for the two familial handedness groups for the ambiguous and unambiguous conditions, collapsing across the attachment manipulation. The pure right-handers displayed a larger advantage in processing the unambiguous sentences compared to the ambiguous sentences, as reflected as a decreased P600 in response to the critical nouns in unambiguous condition. In comparison, the mixed right-handers showed no difference in processing two types of sentences. The effect of ambiguity was explored by comparing the grand mean ERPs averaged across 28 scalp channels between the pure and mixed right-handers (Figure 25B). The ambiguity effect significantly interacted with familial handedness ($F(1,22) = 4.19, p = 0.05$) and was reliable only in the group of pure right-handers ($F(1,12) = 6.57, p < 0.05$). These results suggest pure right-handers might be more proficient in leveraging the disambiguating cues, such as *using* and *that has* in the sentences, to facilitate their online parsing. It may also provide part of explanation to the asymmetry in pure right-handers' ambiguity resolution performance between the sentences containing ambiguously trained verbs and those containing unambiguously trained verbs. If participants were sensitive to the earlier disambiguating cues, they learned not to attach the adjuncts with the unambiguously trained verbs.

Figure 26A shows ERPs in the instrument and modifier sentences collapsing across the ambiguity manipulation, separately for the two familial handedness groups. Both groups showed less positivity in the 500-1300 ms interval for the instrument critical nouns compared to the modifier critical nouns, reflected in the absence of any interaction between familial handedness and attachment ($F(1, 22) = 0.006, p > 0.9$). Figure 26B provides a topographic map of the

attachment effect in the two groups, showing that the effect was widely distributed at both frontal and posterior scalp sites in pure right-handers, while in the mixed right-handers the effect was restricted to frontal sites. Analysis of the spatial distribution of the attachment effect revealed a significant three-way attachment x anteriority x familial handedness interaction ($F(2, 44) = 3.41$, $p < 0.05$), indicating that verb bias learning involved different subpopulations of neurons for the two familial handedness groups.

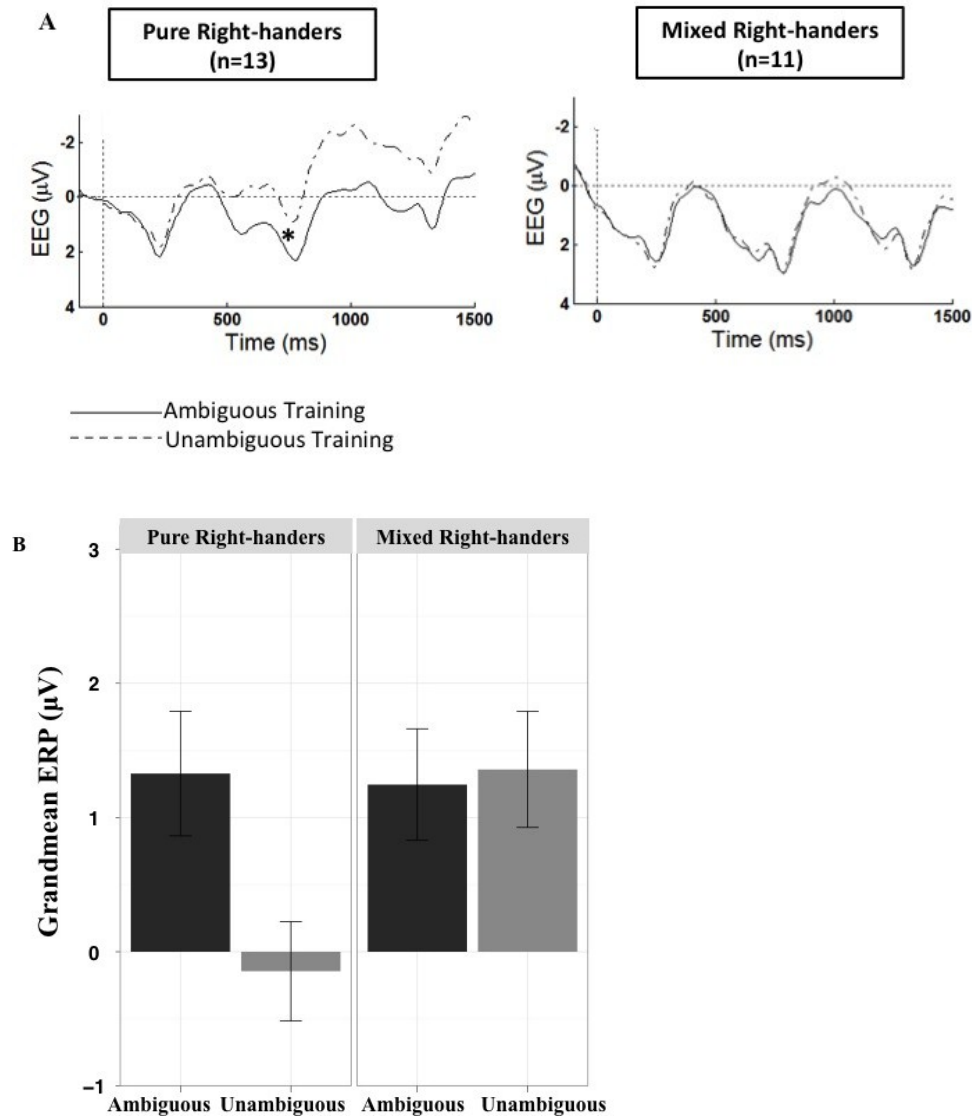


Figure 25: Experiment 6 Individual differences in familial handedness ERP results. Grand mean ERPs to the critical nouns in ambiguous and unambiguous sentences during the second half of the training. (A). ERPs averaged across the three posterior channels P3, PZ and P4. ERPs to the critical nouns in the unambiguous conditions showed a long-lasting posteriorly distributed late negativity only in pure right-handers. (B). Mean amplitude of the 500-1300 ms time window after the critical noun averaged across all 28 scalp sites for the ambiguous and unambiguous condition separately for the two familial handedness groups. Error bars represent the within-group standard errors.

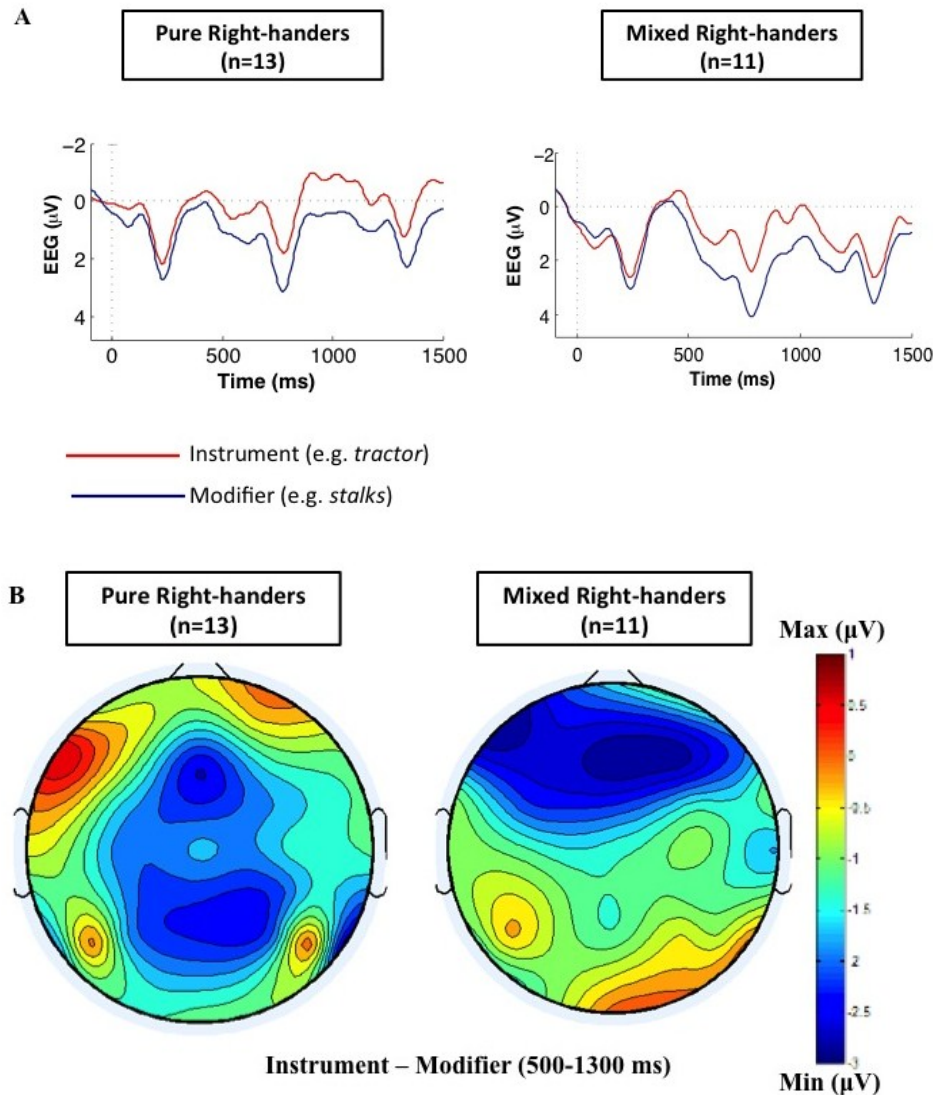


Figure 26: Experiment 6 Individual differences in familial handedness ERP results (continued). ERP responses to the instrument and modifier critical nouns during the second half of the training. (A). Grand mean ERPs averaged across all 28 scalp sites. Both groups displayed a reduction of positivity in response to instrument critical nouns compared to modifier critical nouns. (B). Topographic maps for the Attachment effect, shown as the ERP amplitude difference during the 500-1300 ms time window after the critical nouns. The attachment effect in pure right-handers was distributed widely on the scalp, while the effect in the mixed right-handers was restricted to frontal scalp sites.

3.3.3 Discussion

The results of the current experiment replicated the major findings in Experiment 5 with an extended training paradigm, and furthermore revealed group differences between pure and mixed right-handers during novel verb bias learning.

The ERP results from the first two learning blocks verified the learning effects found in Experiment 5. Mean amplitudes of N400s elicited by instrument nouns were reduced relative to those elicited by modifier nouns during the first training block, suggesting that readers developed an expectation for the instrument noun following verbs with a newly learned instrument bias, making that noun easier to process and integrate when it appeared. In the second training block, the N400 effect diminished and a late frontal positivity emerged that was smaller for instruments than for modifiers. We argued that the scalp distribution of this positivity might have been different from the typical P600 distribution because it overlapped in time with the N400, with the two partially canceling each other out at posterior sites. The centro-frontal scalp distribution of the positivity was consistent with the ERPs to syntactically anomalous words found in L2 learners, which has also been argued to be a P600 effect (Osterhout et al., 2008). If the positivity was indeed P600, it would suggest that readers' expectations for instruments were becoming more abstract and structural by the second block.

Along with the P600 effect, in the second block the instrument training sentences also elicited a larger positive predictive component beginning about 500 ms before the disambiguating word appeared, providing additional support for the claim that readers were learning about verb biases. They learned to anticipate an instrument after a particular verb followed by *with*, based on the co-occurrence of that verb and instrument *with*-phrases during training.

During the two additional training blocks in this experiment, the positivity during the 500-1300 ms time window remained smaller for instrument critical nouns than modifier critical nouns, which was reliable only in the ambiguous training conditions. The differences found between ambiguous and unambiguous training indicate that the link between specific verbs and structures was learned less well when no ambiguity resolution was required.

We also noticed in Figure 24 that during the last two blocks the reduced positivity elicited by the instrument critical nouns started earlier and lasted longer. It appeared that there was an additional earlier negativity effect (EN), which peaked at around 200ms after the instrument noun onset. The EN effect may also reflect a better pattern recognition process of instrument structure than modifier structure as reported in a number of earlier sequence learning studies (Gaillard & Verduin, 1985; Hansen & Hillyard, 1980; Mäntysalo & Gaillard, 1986).

The training outcomes not only depended on the amount of exposure to distributional information, but also differed depending on participants' familial handedness. At the behavioral level, mixed right-handers' picture choices were not reliably affected by training, while pure right-handers did choose pictures consistent with their training reliably more often than pictures inconsistent with their training, but only for verbs trained in ambiguous instrument sentences.

The electrophysiological results during the final two blocks of training were consistent with the claim that pure right-handers had learned more about the novel verbs by then than the mixed right-handers had. During the second half of training, which immediately preceded the behavioral test, a centro-posterior positivity that was argued to be P600 was reduced only in pure right-handers. Centro-posterior P600 has been linked with effortful reanalysis processes during syntactic parsing (Coulson, King, & Kutas, 1998; Hagoort, Brown, & Groothusen, 1993; Kaan & Swaab, 2003; Osterhout, Holcomb, & Swinney, 1994), so its reduction in the last two blocks of

training for the pure right-handers suggests they had learned more about what to expect and thus needed to reanalyze less often than the mixed right-handers.

For the pure right-handers', both their better behavioral performance and the reduction in P600 amplitude during the final two training blocks was restricted to verbs trained in ambiguous instrument sentences. The apparent absence of verb bias learning for verbs trained in unambiguous sentences suggests that pure right-handers took full advantage of the other disambiguating cues in unambiguous sentences (i.e., *using* and *that has*) and therefore did not link information about whether the verb could take instruments as arguments to the verbs themselves. Mixed right-handers did not show the same difference between verbs trained in ambiguous and unambiguous sentences, suggesting that they may have been failing to attend even to obvious lexical cues about sentence structure, making it unsurprising that they also did not encode verb-structure linkages. The greater sensitivity of pure right-handers to the training manipulations is consistent with previous findings showing that pure right-handers relied more on the biases of familiar verbs during garden-path sentence processing (Qi et al., 2010).

In summary, the results in both Experiment 5 and 6 suggest a highly dynamic neural system that continuously collects distributional information from the language input. During the earlier stage of verb bias learning, participants learned to predict the semantic feature of the upcoming critical nouns and the confirmation of their expectation reduced the N400 amplitude. During the later stage of learning, participants started to initiate more abstract prediction about the upcoming structure based on their previous verb-specific experiences, even before the arrival of the critical nouns. The confirmation of their structural prediction reduced the late positivity amplitude. In addition, ambiguity of the training sentences seemed to motivate efficient verb bias learning. In the instrument and modifier unambiguous conditions, there was lack of evidence for

the transition from N400 to late positive complex, suggesting learning the association between individual verbs and unambiguous structures was less efficient. Moreover, pure right-handers who showed stronger sensitivity to the disambiguating cues in the unambiguous sentences also learned less about the verb bias in the unambiguous context, possibly because *using* and *that has* phrases were sufficient parsing cues that could actually prevent efficient learning about the redundant verb-structure association.

3.4 Experiment 7: Newly-learned verb bias in conflict detection

The final experiment was designed to test verb bias knowledge one day after Experiment 6's training session. The purpose of the experiment was two-fold: 1) to examine how readers' verb bias training experience affected their processing of garden-path sentences, where the critical nouns violate the training pattern; 2) to collect information about the participants' sensitivity to the biases of familiar verbs in order to investigate any correlations between that and the acquisition and use of newly learned verb bias. A set of sentences with familiar direct-object/sentential-complement bias verbs was included in the stimuli for measuring verb bias sensitivity.

3.4.1 Methods

3.4.1.1 Participants

15 participants from Experiment 6 came back to the lab on the second day and participated in the study. Their participation was compensated with 8 dollars/hour.

3.4.1.2 Materials

Following the same selection criteria as in Experiment 5, another sixty-four sentence pairs containing the 4 trained verbs and ambiguous *with* phrases were chosen from the same norming study conducted prior to Experiment 5. In addition, sixteen sentence pairs from the training session in Experiment 6 were also included. Each trained verb was tested in 4 old items from the previous day and 16 new items, within which half of the critical nouns disambiguated the sentences in the way that was consistent with the participants' training experience on the first day (Bias-matching condition), while the other half violated the training pattern (Bias-mismatching condition). Some example sentences are shown Table 5. The procedure for counterbalancing the conditions across lists and controlling lexical characteristics of critical nouns was the same as Experiment 5 (see Table 6).

The filler sentences included 60 sentence pairs containing direct-object/sentence-complement (DO/SC) ambiguity (taken from Garnsey et al., 1997), as well as 40 sentence pairs with half of them exhibiting either semantic anomalies or grammatical violations (taken from Tse et al., 2007).

The filler stimuli with the DO/SC ambiguity were developed from Garnsey et al. (1997). For example, the noun phrases (e.g. *the rejection* in (5a) and *the fighter* in (5b)) are temporarily ambiguous between being the direct object of the main verb and the subject noun for the sentential complement, and the auxiliary verbs following them (e.g., *would* or *had*) disambiguate towards a sentence complement structure. Twenty sentences contained DO-biased verbs (e.g., *expect*), which were followed by direct objects at least twice as often as it was followed by sentential complements in the norming study conducted by Garnsey et al (1997). Another 20 sentences contained SC-biased verbs (e.g., *suspect*), which were followed by sentential

complements at least twice as often as it was followed by direct objects in the norming study. Half of these stimuli had the complementizer *that* included before the sentence complement, rendering them unambiguous. The frequency and length of the disambiguating words (underlined in example 5) were matched across conditions.

(5a). DO-SC: *The anxious applicant expected (that) the rejection would arrive in the mail.*

(5b). SC-SC: *The boxing referee suspected (that) the fighter had broken some of his bones.*

In order to prevent participants from always expecting a sentence complement, another 20 sentences with DO endings in the main clause were included, as in (6). The pictures chosen for the picture-viewing phase in trials with DO/SC sentences were always the subject noun (e.g. *referee*) and the post-verbal noun phrase (e.g. *fighter*).

(6a). DO-DO: *The bank worker forgot the combination when she was asked to open the safe.*

(6b). SC-DO: *The careless customer realized the trap after she signed for the credit card.*

The remaining filler sentences consisted of 40 sentence pairs taken from Tse et al. (2007), illustrated in (7), split over 2 lists. Each participant saw 40 of these sentences (20 acceptable and 20 unacceptable). Among the unacceptable sentences, 10 had a semantically incongruent ending at the final word position of the sentence as in (7a, congruent version in parenthesis), and the other 10 had a grammatical violation of pronoun case at word positions in the middle of the sentence as in (7b, grammatical version in parentheses). The 20 acceptable sentences were the matched controls for the semantic and syntactic violation sentences from the other list. The pictures for these filler sentences were chosen for the randomly-selected nouns shared between two members in each sentence pair.

(7a). The rider helped put the saddle on the pool (horse).

(7b). The silver plane took we (us) to paradise and back.

Each participant read 180 sentences randomly ordered and evenly distributed into 4 blocks.

3.4.1.3 Procedure, electroencephalogram recording and analysis

Participants provided informed consent for the multiple-session experiment before their participation in Experiment 6 and were scheduled to come back to the lab at the same time on the next day. The procedure was similar to Experiments 5 and 6, with each trial starting with two pictures followed by a sentence presented word-by-word at the center of the computer screen. Instead of answering a comprehension question at the end of each trial, however, participants were instructed to make a judgment within four seconds after each sentence about whether it was both syntactically correct and semantically congruent.

The experiment started with a brief reminding phase. Participants read 16 sentences containing the 4 trained verbs. Half of the sentences were the same as the training items seen on the previous day, which should get “Yes” judgment responses. The other half had the critical nouns from a different training list, which violated the verb’s trained bias and so should get “No” judgment responses. After each trial, participants received feedback regarding the correctness of their judgment.

After the reminding phase, target sentences were intermixed randomly with fillers and acceptability judgment task continued but participants were told that instead of feedback after each trial, they would be told about their overall accuracy after each block.

EEG recording and analysis procedures were the same as in Experiment 6. The experiment lasted about 1.5 hours.

Table 5: Experiment 7 Examples of the test trials. The underlined words are critical nouns. The italic words were pictured in the picture-viewing session at the beginning of each trial. I: instrument; M: modifier; AM: ambiguous; UN: unambiguous.

Trained Condition	Type of Trials	Sentence Frames	Critical Words	Sentence Endings
I-AM / I-UN	Mismatching	The strong gladiator norged the <i>lion</i> ...	with the sharp <u>teeth</u>	... as the <i>spear</i> broke in half.
M-AM / M-UN	Mismatching		with the sharp <u>spear</u>	...as the spectators cheered with excitement.
I-AM / I-UN	Matching	The young girl veebed the <i>plant</i> ...	with the green <u>scissors</u>	... in order to get rid of the yellow leaves.
M-AM / M-UN	Matching		with the small <u>leaves</u>	... in order to try out her new scissors.

Table 6: Experiment 7 Lexical properties of critical nouns.

Testing Condition	Length (letters)	Frequency (words/million)	Concreteness (100~700)	Imagability (100~700)
Instrument endings (e.g. <i>tractor</i>)	6.09 (0.26)	37.32 (11.41)	363.40 (41.23)	367.78 (41.46)
Modifier endings (e.g. <i>stalks</i>)	5.95 (0.22)	26.47 (3.76)	309.39 (47.58)	310.55 (47.95)

Mean of the absolute values of length of words, word frequency per million words (retrieved from American English subtitles, SUBTLEX_{WF}), ratings for concreteness and imagability are reported with standard errors in parenthesis.

3.4.2 Results

3.4.2.1 Behavior results

Behavioral performance was measured as the proportion of accurate judgments in the acceptability judgment task. People were accurate overall in detecting the grammatical and semantic violations in the fillers, with 92% correct for pronoun case mismatch violations and 93% correct for semantic violations.

(A). Acceptability judgment for the DO/SC sentential complement sentences

All the DO/SC temporarily ambiguous sentences were grammatically and semantically correct, so the correct acceptability judgment for them was “yes”. The mean proportion of these sentences judged acceptable was 82%, ranging from 52 - 99% across participants. Figure 27 shows the mean proportion of acceptable judgments in each of the four conditions. The sentences with DO-bias verbs were judged unacceptable more often than those with SC-bias verbs, but only in the ambiguous conditions. A two-way ANOVA with the factors verb bias (DO-verb or SC-verb) and ambiguity (whether or not the sentence contained the complementizer *that*) revealed a main effect of verb bias ($F(1,14) = 8.03, p < 0.05$) and a significant interaction between verb bias and ambiguity ($F(1,14) = 23.05, p < 0.001$). The significance pattern was the same when proportions were normalized using an arcsine transformation. These results mirror the findings from a number of previous studies using the similar material (Garnsey et al., 1997; Novick et al., 2008; Osterhout, Holcomb, & Swinney, 1994). Because readers have experienced greater processing difficulty in the sentences containing a DO-bias verb followed by a sentential complement, they are more likely to judge them unacceptable than when a SC-bias verb is followed by a sentential complement.

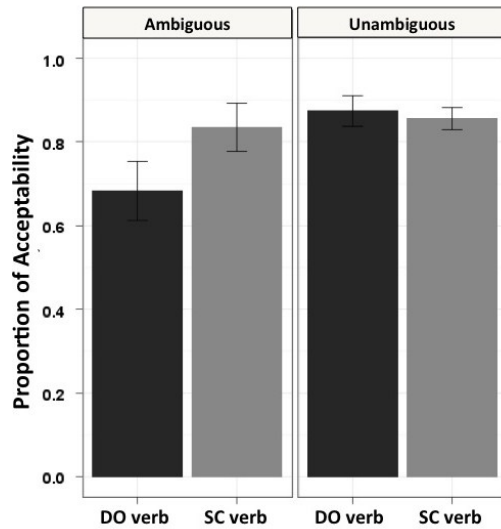


Figure 27: Experiment 7 Behavioral results for DO/SC sentences. Mean proportion of acceptability judgment. Error bars represent the standard errors.

(B). Acceptability judgment for the target sentences with the trained novel verbs

Figure 28 shows the mean proportion accuracy in judging whether the sentences containing one of the four trained verbs were acceptable or not. Overall mean accuracy was 49%, which was not different from chance (50%). The performance in judging the sentences with verbs trained in unambiguous modifier sentences was the worst among the four conditions. A two-way ANOVA with the factors attachment (instrument or modifier) and ambiguity (ambiguous or unambiguous) revealed a significant interaction between attachment and ambiguity ($F(1,14) = 8.12, p < 0.05$), with marginal main effects of attachment ($F(1,14) = 4.32, p = 0.06$) and ambiguity ($F(1,14) = 3.84, p = 0.07$).

Table 7 shows participants' accuracy in each condition separated into hits and correct rejections. Participants judged the sentences that consistent with prior verb bias training acceptable more often than chance except in the modifier-unambiguous condition. For trials that violated their previous verb bias training, participants tended to make more false alarm errors

(judging the sentence acceptable when they should not) for the modifier-trained verbs than the instrument-trained verbs. The same main effects and interaction were found when the behavioral responses were transformed into d-prime scores. The results suggest two types of learning asymmetry: 1) it was more difficult to learn modifier bias than instrument bias, and 2) it was more difficult to learn from the unambiguous than from the ambiguous sentence structure.

It is worth noting at this point that participants had a total of just 75 minutes of exposure to the four novel verbs in sentences that were constructed to make it difficult to draw inferences from them about the verbs' meanings. The newly-learned verb biases were no doubt still quite fragile. That together with the fact that the only thing wrong with the critical sentences that violated the trained verb biases was that violation means that the forced-choice task may have increased the chance of engaging in guessing or other behavioral strategies. The EEG responses to the test sentences may provide a better measure of verb bias knowledge.

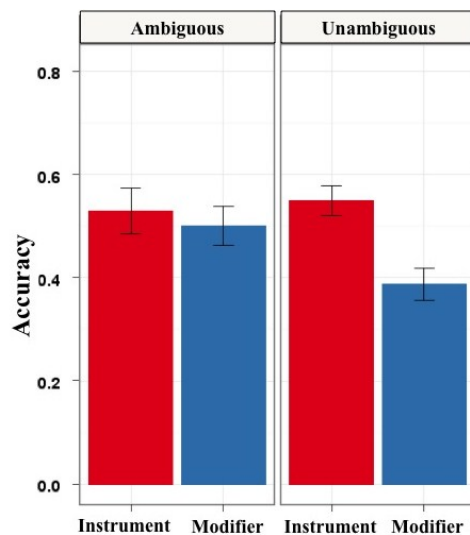


Figure 28: Experiment 7 Behavioral results for target sentences. Mean proportion of judgment that was consistent with training. Error bars represent the standard errors.

Table 7: Experiment 7 Behavioral results for target sentences by participants' responses. Mean proportion of accuracy (standard errors in parenthesis).

Condition	Hits	Correct Rejections	Overall Accuracy
I-AM	0.65 (0.06)	0.42 (0.08)	0.53 (0.04)
I-UN	0.64 (0.05)	0.47 (0.07)	0.55 (0.03)
M-AM	0.65 (0.06)	0.36 (0.05)	0.50 (0.04)
M-UN	0.54 (0.06)	0.25 (0.04)	0.40 (0.03)

Training conditions: I: instrument; M: modifier; AM: ambiguous; UN: unambiguous. Hits represent the proportion of correctly accepting the sentences that were consistent with previous verb bias training. Correct rejection rates represent the proportion of correctly rejecting the sentences that were inconsistent with previous verb bias training.

3.4.2.2 Event-related potentials

Previous studies have found the dominant ERP response to verb bias violations for familiar DO/SC verbs to be P600, so the 500-1300 ms time window following the onset of the disambiguating auxiliary verb (e.g., *would* and *has*) was analyzed for those sentences. It was also expected that P600 would be the dominant response to critical nouns that disambiguated toward the wrong structure (e.g., *tractor* and *teeth*) in the sentences containing trained novel verbs, so the 500-1300 ms time window following the onset of those nouns was also analyzed for those target sentences.

The number of participants in this study was small (15) given the number of trials/condition/participant, with the unfortunate result that the ANOVAs testing differences between conditions yielded few significant effects. However, the data still allowed an investigation of possible effects of individual differences on verb bias learning and its behavioral and electrophysiological manifestations.

(A). ERPs to the disambiguation words in the DO/SC sentential complement sentences

Figure 29A shows grand mean ERPs at the disambiguating auxiliary verbs in all four conditions for the DO/SC sentences. It is apparent in the figure that the ambiguous sentences with DO-bias verbs elicited a larger P600 than all of the other sentence types. However, in an ANOVA with the factors verb bias and ambiguity, there was only a non-significant trend toward an interaction between the two factors ($F(1, 14) = 2.12, p = 0.17$), presumably because of low power. Figure 29B plots the P600-acceptability correlation within each sentence type. The correlation is stronger in the ambiguous than the unambiguous conditions. ANOVA with ambiguity and P600 amplitude as independent variables showed that P600 interacted with ambiguity in predicting how well the sentences were accepted by the participant ($F(1, 28) = 6.03, p < 0.05$). These results suggest P600 is sensitive to how well a sentence's structure matches its verb's bias during online sentence processing. An index of an individual participants' sensitivity to the biases of familiar verbs was calculated as a difference of differences, by subtracting the P600 amplitude in unambiguous from that in the ambiguous conditions for each verb type and then subtracting one of those difference measures for DO-verb from the difference measure for SC-verb. More positive index values indicate greater sensitivity to the biases of familiar verbs⁵. The goal of calculating each participants' verb bias sensitivity index was to determine whether participants who were better at using the biases of familiar verbs would also be better at using the newly learned verb biases of novel verbs.

⁵ Perhaps due to small sample size, pure and mixed right-handers did not differ in their sensitivity to the bias of familiar verbs: $t(10.51) = 0.39, p > 0.70$.

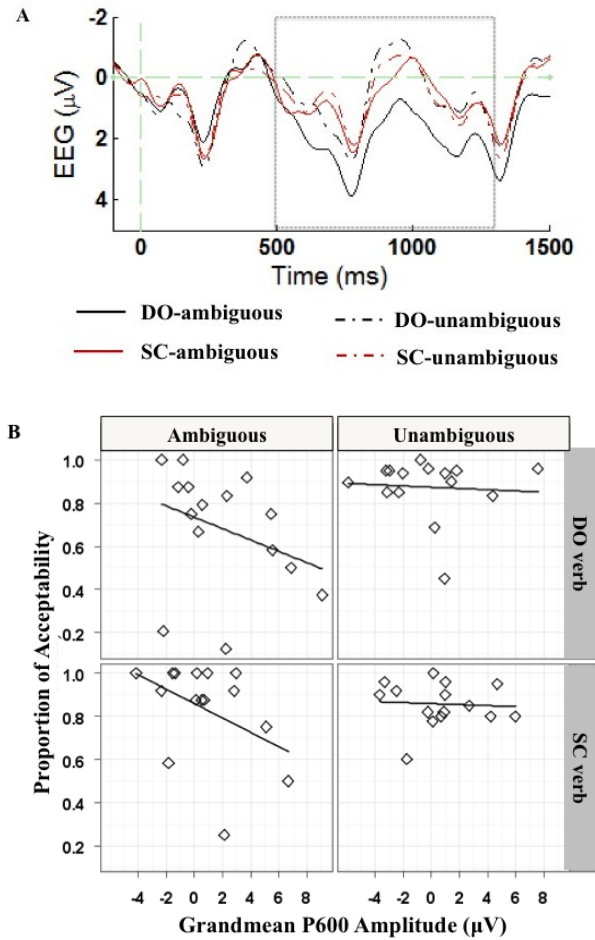


Figure 29: Experiment 7 ERP results for DO/SC sentences. A. Grand mean ERPs to the disambiguating auxiliary verbs in the sentential complement sentences with DO-bias (black lines) and SC-bias (red lines) verbs averaged across all 28 scalp sites. DO-bias-ambiguous sentences elicited the largest P600 compared to the other conditions. B. The correlation between the P600 amplitude and participants' acceptability judgments by sentence conditions, showing stronger correlation between P600 and acceptability judgments in the ambiguous sentences than in the unambiguous sentences. Lines represent linear regressions.

(B). ERPs to the critical disambiguating nouns in target sentences with trained novel verbs

Although the behavioral task in Experiment 7 required detecting violations in both the syntactic and semantic domains, which imposed different demands than the reading task used in Experiments 5 and 6, it was still expected that differences between conditions would appear as differences in P600 amplitude. Previous work has shown P600 amplitude to increase when

structural revision must be made when garden paths are encountered and also when complexity increases the difficulty of syntactic integration (Coulson et al., 1998; Hagoort et al., 1993; Kaan & Swaab, 2003; Osterhout, Holcomb, et al., 1994). Figure 30 shows the ERP waveforms to the critical nouns in each condition, averaged across three posterior sites (P3, PZ and P4). In the bias-matching trials, where the sentences were consistent with the training pattern, ANOVA with 17 levels of posterior scalp channels, ambiguity and attachment training conditions showed no reliable main effects of these factors nor any interaction between them ($F_s(1, 14) < 1$). In the mismatching trials, the nouns that violated prior instrument-ambiguous training elicited a larger positivity during the 500-1300 ms time window, but an ANOVA found only a non-significant trend toward an interaction between ambiguity and attachment ($F(1, 14) = 2.57, p = 0.13$). We computed the P600 amplitude at the critical nouns averaged across 17 central-posterior scalp channels in each target sentence condition to diagnose the degree of syntactic conflict detection.

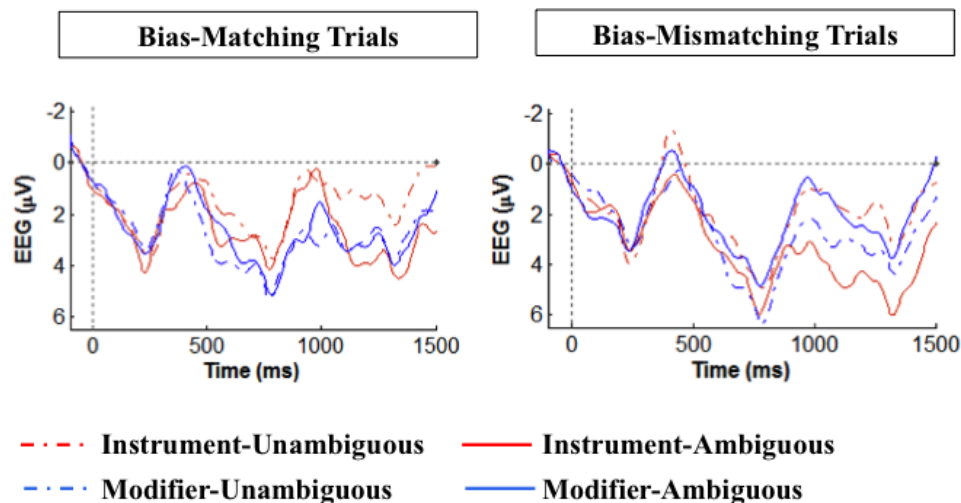


Figure 30: Experiment 7 ERP results for target sentences. Grand mean ERPs to the disambiguating critical nouns by condition in the instrument (red lines) and modifier (blue lines) sentences, averaged across three centro-posterior channels (P3, PZ and P4). In the matching trials, where the critical nouns confirmed previously learned verb biases, no obvious pattern of either attachment or ambiguity effect was observed. In the mismatching trials, where the critical nouns contradicted previous training, instrument-ambiguous trials elicit a larger P600 compared to all the other conditions.

(a). Correlations with the predictive component during training

Experiments 5 and 6 both found an apparent predictive effect in the ERPs beginning about 500 ms before the critical disambiguating noun appeared. The waveforms were more positive in the highly predictable instrument training conditions than in the less predictable modifier conditions (see Figures 17 and 23). The amplitude of this predictive effect grew across training blocks, suggesting that participants were keeping track of the distributional information about the verbs and actively predicting upcoming words based on those distributions. We examined the relationship between the occurrence of prediction during training and the degree of conflict detection during testing. The amplitude of the prediction effect was measured across the 500 ms window before the onset of the critical noun, averaged across 28 scalp electrodes and across four training blocks. Figure 31A and B show the relationship between the amplitude of the ERP prediction effect in the training sentences and the P600 amplitude in the matching and mismatching testing sentences. In the bias-matching testing condition, the ANOVA with factors of predictive component, ambiguity and attachment as independent variables yielded no significant main effects or interactions between them (F 's (1,47) < 1.33, p 's > 0.25) on the P600 amplitude during the testing. In contrast, in the mismatching trials, amplitude of predictive component marginally correlates with P600 amplitude (F (1,47) = 3.68, p = 0.06). Ambiguity and attachment did not interact with the correlation between these two ERP components (F 's (1,47) < 0.70, p 's > 0.41) in the mismatching testing condition. The results suggest that participants who showed more evidence that they were making predictions during training had more trouble with violations during testing.

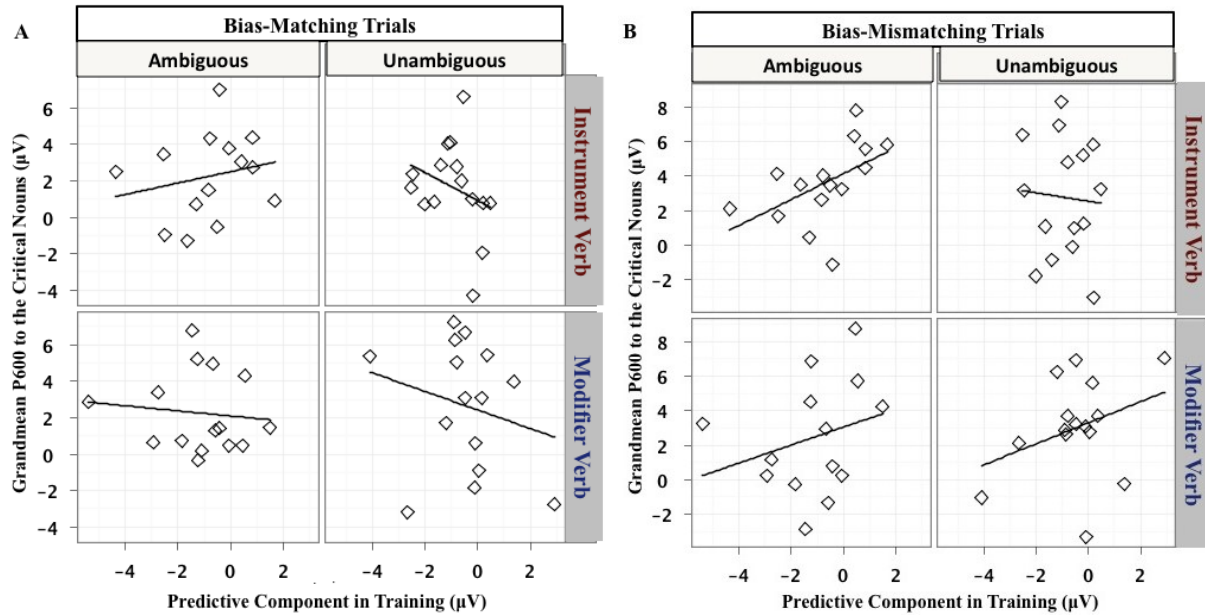


Figure 31: Experiment 7 Relationship between ERPs during training and ERPs during testing. X-axis is the mean amplitude of individual's predictive component in -500-0 ms before the onset of the critical nouns averaged across 28 scalp channels during training. Y-axis is the mean amplitude of individual's P600 to the critical nouns averaged across 17 posterior channels in the test trials. (A). In the matching trials, the degree how an individual predicted the upcoming words had no effect on how they processed an expected critical noun in the test trials. (B). In the mismatching trials, participants who showed more evidence of making predictions during training showed larger P600 to violations during testing. Lines represent linear regressions.

(b). Correlations with sensitivity to the biases of familiar verbs

We examined the relationship between individual participants' sensitivity to the biases of familiar verbs and their ERP responses to violations of newly learned biases for novel verbs⁶. Figure 32A shows relationship between the verb bias sensitivity score measured from DO/SC sentences and P600 amplitude in the matching target trials. ANOVA with verb bias sensitivity as

⁶ Two outlier data points were dropped from this analysis because their values were more than 2.5 standard deviations away from the mean (1 each in the congruent instrument-ambiguous condition and the incongruent modifier-ambiguous condition).

independent variables showed no significant effect on the P600 amplitude in response to the critical nouns in the matching sentences ($F(1, 12) = 0.39, p = 0.54$). Ambiguity and attachment do not interact with the effect of verb bias sensitivity (F 's $(1, 12) < 2.09, p$'s > 0.17). Figure 32B shows the relationship between the verb bias sensitivity score and P600 amplitude in the incongruent trials. For verbs trained in ambiguous sentences, the more sensitive participants were to familiar verb biases, the larger their P600 responses to critical nouns that contradicted the biases they had learned for novel verbs. In contrast, for verbs trained in unambiguous sentences, participants with greater familiar verb bias sensitivity produced smaller P600s in response to critical nouns that disambiguated the sentence toward the modifier interpretation. That is, when a test sentence used a verb that had been trained in unambiguous sentences to have instrument bias (i.e., *The X veebed the Y using the Z*) and the critical noun's meaning meant it had to be a modifier rather than an instrument, participants who were more sensitive to the biases of familiar verbs were less bothered. Recall that verb bias learning was only demonstrated for verbs trained in ambiguous sentences. Perhaps that was especially true for participants with good familiar verb bias sensitivity. If verbs trained in unambiguous sentences did not develop biases for them, it didn't matter which kind of critical noun appeared in the test sentence. An ANOVA revealed that verb bias sensitivity significantly interacted with ambiguity in predicting P600 amplitude in incongruent trials during testing ($F(1, 12) = 6.39, p < 0.05$). The positive correlation between familiar verb bias sensitivity and P600 amplitude in response to incongruent test trials for verbs trained in ambiguous sentences suggests that newly acquired verb biases behave much like biases for familiar verbs during sentence processing. The absence of such correlations for verbs trained in unambiguous sentences fits well with other evidence described earlier that unambiguous sentences did not lead to verb bias learning.

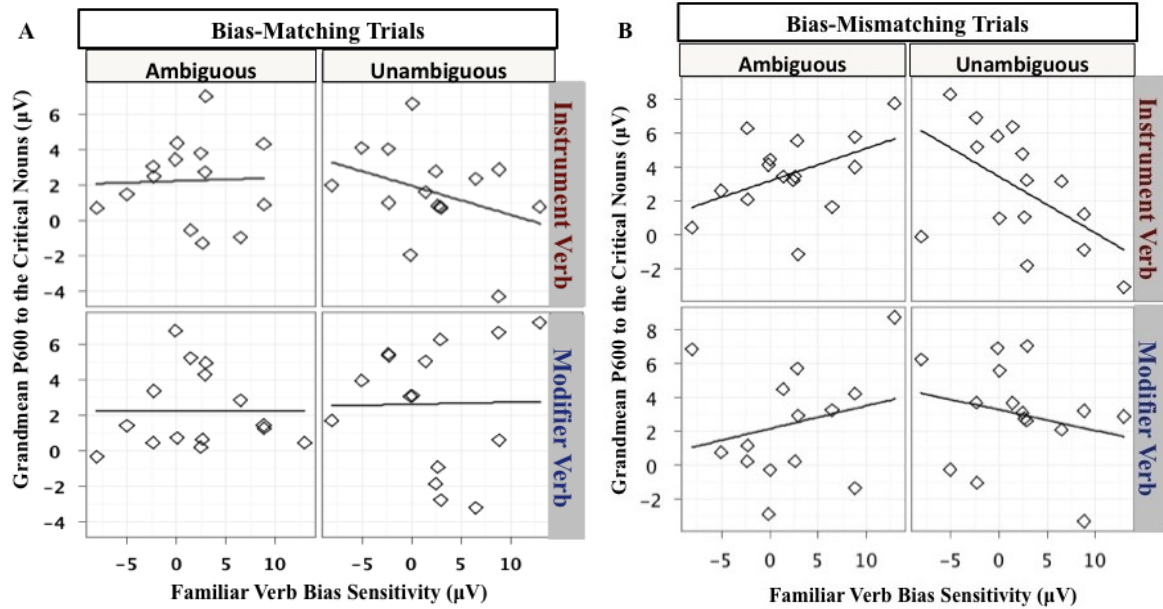


Figure 32: Experiment 7 Relationship between sensitivity to biases of familiar verbs and ERPs in test trials. Correlations between individual participants' verb bias sensitivity scores and P600 amplitude to the critical nouns in the bias-matching (A) and bias-mismatching (B) target trials. In matching trials there was no correlation, but in mismatching trials, greater sensitivity to the biases of familiar verbs predicted larger P600 amplitude in response to incongruent critical words for verbs trained in ambiguous sentences and smaller P600 amplitude for verbs trained in the unambiguous condition. Lines represent linear regressions.

3.4.3 Discussion

The overall behavioral results in the current experiment suggested participants had little explicit knowledge of the trained verb biases. In spite of that, participants showed differences in their performance across four tested conditions. Consistent with the findings in Experiments 5 and 6, the acceptability judgment data also showed better training for instrument structures and for ambiguous training sentences. Most errors were made in the modifier-unambiguous condition, confirming that learning verb bias from modifier-unambiguous training sentences was the least effective.

The ERPs in the incongruent trials, whose structure did not match the training sentences,, suggest that newly learned verb bias guided participants' expectations about upcoming sentence structure and led them to garden path during the incongruent trials. The stronger the predictions were about sentence structure during training, as evidenced by the amplitude of the prediction effect in the ERPs, the stronger the responses were to disambiguating nouns that mismatched the predicted structure at test, as evidenced by P600 amplitude.

Participants were found to vary in their sensitivity to familiar verb biases, and those individual differences also influenced their structural bias learning for novel verbs. Learning the biases of familiar verbs results from long-term exposure to sentences in natural environments, and verb meaning probably contributes to that learning. Here, however, care was taken to prevent participants from inferring meanings for the novel verbs. The fact that participants who showed greater sensitivity to the biases of familiar verbs were also better at learning the biases of novel verbs without support from the verb's meaning suggests that such people may be better at keeping track of statistical regularities in the input and learning abstract structural patterns from them.

In addition to showing better learning of the biases of novel verbs trained in ambiguous sentences, participants with more sensitivity to the biases of familiar verbs also showed worse learning of the biases of verbs trained in unambiguous. This suggests that they were generally more effective users of all of the available cues. Perhaps because words other than the verb determined the structure of the unambiguous sentences (i.e., *using* or *that has*), the more bias-sensitive participants did not link the statistical regularities in the sentence structures to the verbs. This would be consistent with a well-known general learning phenomenon called "blocking" or "overshadowing", in which a cue that is correlated with an already-learned highly

reliable cue tends not to be learned well (Ellis, 2006; Pearce et al., 2006). The first-learned cue is said to block the learning of the second cue. This would provide an explanation for why verb bias was successfully learned only from ambiguous training sentences. Perhaps the participants who were better at picking up verb biases from ambiguous training sentences were worse at doing so in unambiguous training trials because they were better at attributing sentence structure to the other cues in those sentences and thus did not associate it with the verbs themselves. Notice that this may have contributed to their higher sensitivity to the biases of familiar verbs as well, since that measure was derived by subtracting responses in unambiguous trials from those in ambiguous trials, which would be maximized for those who were better at using the complementizer cue in unambiguous DO/SC sentences.

3.5 General Discussion

The series of studies presented here explored behavioral and neural changes that took place during and after the learning of PP-attachment biases for novel verbs. In Experiments 5 and 6, participants were exposed to four novel verbs, each of which was embedded in one of four kinds of sentence structures. The relationship between a particular verb and a particular structure was fixed throughout training. The results showed first of all that it is possible to learn such biases without support from verb meaning. The ERP results also showed a shift from N400 effects early in training to late positive effect during the later stage of training, suggesting a change in the nature of the processing as learning proceeded, which is consistent with the results of other ERP studies of the learning of natural second languages (McLaughlin et al., 2010; Osterhout et al., 2008) or artificial languages (Morgan-short et al., 2010). The ERP transition over the course of learning was observed mainly for verbs trained in ambiguous sentences,

suggesting that the absence of other reliable cues about structure facilitated the efficient establishing of verb-structure linkages.

Another ERP effect emerged later in training in the instrument training trials, which was a prediction effect that began in the waveforms about 500 ms before the disambiguating critical noun was presented, suggesting participants gradually strengthen their prediction about the instrument attachment during learning. The correlational analysis in Experiment 7 further verified the amplitude of predictive component is related with the degree of parsing commitment readers made during online reading. The larger the component is, the greater violation effect the individual encountered when reading sentences contradicted the training bias.

Across Experiments 6 and 7, the contribution of individual differences to learning efficacy was analyzed. In Experiment 6, compared to mixed right-handers, pure right-handers appeared to be more sensitive to the ambiguity manipulation and showed a larger behavioral learning effect in the picture-matching task. The smaller number of subjects in Experiment 7 did not allow a comparison for pure and mixed right-handers, but another individual difference did influence learning. Participants who showed greater sensitivity to the biases of familiar verbs also learned the biases of novel verbs trained in ambiguous sentences better but learned less about the biases of novel verbs trained in unambiguous sentences, suggesting that they were generally more effective users of the best available cues about sentence structure. In sum, these results highlighted the rapid use of newly learned verb bias during online parsing and demonstrated that ambiguity greatly motivates efficient verb bias learning.

Across a number of different measurements in the study, learners appeared to acquire instrument verb bias more efficiently than modifier verb bias, reflected as reduced amplitude in N400 and P600 responding to the instrument critical nouns, as well as higher picture-matching

and acceptability judgment accuracy in instrument-trained trials. The asymmetry can be due to two possible reasons.

First of all, people's overall preference to instrument attachment may have led to easier processing and more efficient learning in instrument training sentences. According to the corpus analysis reported by Snedeker & Trueswell (2004), the verb-phrase attachment use of *with* phrases is far more common than noun-phrase attachment use in adult speech. However the error-based learning model would predict the opposite, i.e. that the infrequent modifier attachment should have been learned more efficiently, given people's overall preference for instrument attachment. Studies of syntactic priming have also found stronger priming effects if the primed structure was inconsistent with the verb bias, which has been called an effect of inversed preference (Bernolet & Hartsuiker, 2010; Chang et al., 2006).

Secondly, instrument *with*-PP phrases differ from modifier *with*-PP phrases in several ways. Although instruments are usually optional and thus could be considered adjuncts rather than arguments, they do usually contribute to the verb's meaning, making them more argument-like. Koenig, et al. (2003) proposed that argument roles are semantically obligatory and part of the meaning of only a small class of verbs. Participant roles that are semantically optional or that are obligatory but do not individuate the meanings of verbs are considered to be adjuncts (Koenig, Mauner, & Bienvenue, 2003). Thus, instruments are more argument-like, while modifiers are definitely not verb arguments. The link between a particular verb and the modifier phrase in modifier-training sentences is thus much weaker than the link between instrument phrases and the verb in instrument-training sentences. Such asymmetry leads to intrinsic differences in the predictability of particular kinds of phrases in sentences with instrument-trained and modifier-trained verbs. It is more likely that information that a verb takes instruments

and the kinds of instruments it takes will be linked to the verb itself. Participants may have learned much less about the argument structure preferences of verbs trained in the modifier structures because the modifiers were not arguments of the verb.

Thirdly, the training situation was designed to encourage prediction. We purposefully built in a picture-viewing session to allow direct observation of the asymmetric predictive process. Two nouns that would be mentioned somewhere in the sentence were presented in labeled pictures before the sentence and participants were encouraged to try to predict when they would appear in the subsequent sentence. The possible instrument noun was one of the two pictured nouns but the modifier noun was not. The other labeled picture was instead the direct object noun. The direct object picture included the property that would be mentioned if the sentence had the modifier structure, but the modifier noun itself was not presented as the picture label. Thus, instrument nouns were more predictable than modifier nouns. Crucially, though, the position of the instrument noun in the sentence was predictable only for verbs trained in the instrument structure. For verbs trained in modifier structures, the instrument noun was mentioned somewhere else later in the sentence rather than right after the direct object noun. Thus, any effects of making specific predictions before the critical noun should be observable in the ERP waveforms in instrument-training trials but not in modifier-training trials.

The predictive component showed learners gradually strengthen their prediction about the instrument attachment during learning, because of its high predictability in both semantics and syntax. The correlational analysis in Experiment 7 further verified the amplitude of predictive component is related with the degree of parsing commitment readers made during online reading. The larger the component is, the greater violation effect the individual encountered when reading sentences contradicted the training bias. The temporal feature of the predictive component to the

adjective resembles that of P2 effect, previously reported to be sensitive to the predictability of sentential context (Wlotko & Federmeier, 2007). Two recent experiments revealed intriguing evidence for the dynamic relationship between prediction, statistical learning, and sentence processing. On one hand, individual's predictive performance in a mouse-tracking paradigm after a visual pseudoword sequential training is correlated with participants' reading time of complex sentences in natural language (Misyak, Christiansen, & Tomblin, 2010). On the other hand, individual's implicit learning ability in an artificial grammar-learning task is correlated with one's ability to predict the final word in sentences. None of the other executive function measures correlated with individual ability for word prediction (Conway, Bauernschmidt, Huang, & Pisoni, 2010). These empirical findings showed individual's ability to rely on top-down knowledge for prediction is recruited in both efficient sentence processing and successful statistical learning. Taken together, our results provided additional evidence for the involvement of predictive process in using the newly learned statistical feature of a particular verb during sentence processing.

Another important finding in the current study suggests that verb bias was learned more efficiently in ambiguous sentences than in unambiguous sentences. In Experiment 6, though the size of N400 attachment effect at the earliest stage of training was similar between ambiguous and unambiguous sentences, the P600 attachment effect during the second half of training was mainly observed in the ambiguous sentences, but not in the unambiguous sentences. The lack of N400-P600 shift in the unambiguous conditions indicates that despite the predictable instrument critical nouns, the unambiguous *using* had prevented participants from effectively generating the association between the verb and the structure. Importantly, evidence from individual difference analysis further verifies the earlier disambiguation as a barrier for verb bias learning in

unambiguous condition. Pure right-handers, who were sensitive to ambiguity manipulation during training, showed a stronger learning effect in their picture-matching behavior in the ambiguous condition than the unambiguous condition. In Experiment 7, we measured individual's verb bias sensitivity from the difference of P600 ambiguity effect between DO condition and SC condition, which by definition also reflected individual's sensitivity to ambiguity manipulation. Learners with higher sensitivity score showed greater P600 response to verb bias violation for ambiguously trained verbs, but smaller response for unambiguously trained verbs. Consistent with the prediction from the implicit learning model (Chang et al., 2006), ambiguity during training increased the frequency of errors and further strengthen the connection between the verb and the structure. These findings may provide a different theoretical source of information for verb bias learning, in parallel of the linguistic co-occurrence frequency.

In conclusion, the current study provides clear evidence for the highly dynamic neural system, which constantly collects linguistic distributional information from real time language input and quickly applies the newly learned information into online parsing. The results revealed distinct neural processes transiting from a lexical semantic process to a grammatical rule-based process during verb bias learning. The learning efficacy of verb bias is modulated by the involvement of predictive process, the ambiguity of the training sentences, as well as individual differences in verb bias sensitivity.

CHAPTER 4 CONCLUSION AND DISCUSSION

Frequency-sensitive verb bias serves as a reliable parsing constraint in online sentence comprehension for experienced adult language users as well as preschoolers, whose language skills are still in development. Some pioneer work has been done to explore the contribution of linguistic statistical information for verb bias learning using miniature artificial language paradigm (Wonnacott et al., 2008). However, little is known about the dynamic interplays of event and linguistic distributional information in verb bias learning and how new statistical information is encoded online. The primary purpose of the dissertation is to examine the sensitivity of children and adults to linguistic distributional information and to further explore the neurocognitive mechanisms of verb bias learning. In the current chapter, I will first summarize the key findings of the study. Then I will elaborate discussion in the following two aspects: 1) the contribution of event distributional information in verb bias learning; 2) ambiguity motivates effective verb bias learning. Finally, I will also discuss the potential application of the current findings in future works.

4.1 Summary Of Results

First of all, our experiments demonstrated across modalities that both children and adults are highly sensitive to verb-specific statistics from a minimal degree of language exposure. In Experiment 1-4, each verb only appeared four times in either the instrument PP-attachment structure or the modifier PP-attachment structure in training dialogues. When encountering the same verbs during online ambiguity resolution, listeners' eye-movement pattern reflected consideration of either modifier or instrument interpretation that is consistent with the particular

training experience with the verbs. In Experiment 5, adults showed verb-specific ERP responses both at and before the critical nouns within 30 minutes of reading experiences.

Across experiments, linguistic distributional information, i.e. the likelihood a specific verb co-occurs with a structure, plays a crucial role as an independent information source for verb bias learning. In Experiment 3, training experiences without the co-occurrence frequency information about a particular verb and *with*-PP phrases appear to induce a weaker learning effect. Children did not retrieve the verb information until the PP-object nouns in the test phase. Even at the PP-object noun region, the overall training effect on ambiguity resolution was relatively weaker compared to Experiment 1, where linguistic distributional information was also available. In both the Experiment 4 and Experiment 5-7, both children and adults were capable of learning novel verb bias with little or limited event knowledge about the verb semantics, which further supported the central role of linguistic distributional information in verb bias learning.

Another important question addressed by the study is the time course of verb bias retrieval during online comprehension. At the *with*-PP object noun region, which was usually considered as the critical region for verb bias effect, the usage of the newly learned verb bias is similar as the previous findings with familiar verb bias. In the present eye-movement experiments, listeners looked more at the target instrument and less at the target animal for the sentences containing instrument-trained verbs as opposed to those containing modifier-trained verbs. In Experiment 7, when resolving the conflict between the upcoming words and the predicted structure, participants with higher sensitivity to familiar verb bias also elicited a larger P600 response in trials containing newly learned verbs. The current study also provided converging evidence for the predictive role of verb bias in guiding online parsing commitment.

Experiment 1, 2 and 4 revealed anticipation of an instrument soon after children and adults heard an instrument-trained verb. In the ERP studies, adults learned to predict an instrument across training trials, reflected as an increase of difference between instrument and modifier-training trials before the arrival of the disambiguating *with*-PP object nouns. These results provide strong evidence of the active involvement of newly learned verb bias during online sentence comprehension.

Finally, as the first attempt to track online encoding process for verb bias learning, the current ERP studies demonstrated discrete neural stages as learners continuously collected verb-bias statistical information. During the earlier stage of learning, the N400 effect suggests the confirmation of readers' semantic prediction about an upcoming instrument. While in the later stage of learning, the effect became a reduced P600 effect, indicating the initiation of abstract structural learning. The results are consistent with the findings from second-language acquisition literature (*e.g.* McLaughlin et al., 2010). Importantly, the N400-P600 transit was mainly observed in the training sentences containing ambiguous *with*-PP, but not in those containing earlier disambiguating cues, such as *using* and *that has*. The asymmetry between ambiguous and unambiguous training provided another possible explanation for the weaker training effect observed in Experiment 3. Disambiguation occurred earlier in the sentence interfered with the establishment of verb-structure association, which further weakened the involvement of the verb-specific information during online sentence processing.

4.2 Contribution of Event-distributional Information

Our results supported the central role of linguistic distributional learning in verb bias acquisition, which is independent from the event semantics. However, this does not deny the

contribution of event-distributional information in the learning process. Semantic association between the verb and its associative structure is considered as an important parsing constraint among many others. For example, verbs, such as *eat* and *drink*, define the animacy of the verb agents and the semantic properties (e.g. edible and drinkable) of the patients. Such semantic / thematic specification of verbs is known as selectional restriction (Chomsky, 1965). Adults use the verb selectional restriction to predict the referents in the visual world during sentence processing (Altmann & Kamide, 1999; Kamide et al., 2003). Similar as verb structural preference, the knowledge of verbs' selectional restriction should also be based on previous observations of the participants in the verb events. Removing the verb semantics from training weakened the semantic association between the verb and the following structures, but it does not prevent learning about the semantic category that could be attached to the verb (S. Yuan et al., 2011). In the current study, children and adults displayed online learning effect under the novel verb circumstances. However, their behavioral performance in Experiment 4 and Experiment 7 showed the learning process was implicit and difficult.

Evidence from machine learning literature provided insights about the possible role of verb semantics in learning how to interpret *with*-PP attachment ambiguity. Parsing model is trained by the semantic relatedness between words in the corpus data. In the test sentence (e.g. *eat salad with fork*), model's interpretation of the *with*-PP phrase is based on the relative semantic association score of *eat* and *fork* compared to the score of *salad* and *fork* (Pantel & Lin, 2000). This model outperforms the earlier classifiers, which were solely built on the lexical co-occurrence of verbs and nouns (Hindle & Rooth, 1993). Based on this, semantic association between verb and the instrument might also play a facilitating role in verb bias learning.

Another approach is to manipulate the likelihood of instrument use in the event. In a recent study by Amato et al. (2009), adults listened to sentences containing novel verbs while watching concurrent video scenes. For event manipulation, some verbs always represent action events that require tools and others represent actions that are most likely to be accomplished with bare hands. Participants only heard simple transitive sentences (e.g. *Mary fleeked the horse*) for event-manipulated verbs. For linguistic manipulation, half of the verbs frequently co-occurred with an instrument *with*-PP phrase in the sentences (e.g. *Mary fleeked the horse with the tongs*) and the other half rarely did (e.g. *Mary fleeked the horse*). Participants always watched action videos involving instruments for linguistic-manipulated verbs. When encountering the same verb in a sentence containing a modifier *with*-PP phrase (e.g. *The runner fleeked the bottle with the lime gatorade*), self-paced reading time revealed a main effect of type of manipulation, that is, a larger garden-path effect for the verb items that were trained in two linguistic manipulation conditions than those trained in two event manipulation conditions (Amato, Willits, MacDonald & Sussman, 2009). If observing the instrument use in the events could effectively shaped the *with*-PP attachment verb bias, we would expect a larger garden-path effect for the verb items trained in event manipulation than those trained in linguistic manipulation, given the different distribution of instrument use in training videos (100% vs. 50%). Authors explained their results as evidence for all-or-none linguistic statistics learning. Subjects may have well learned the association of the verb and the likelihood of instrument use from training, but event-distributional information alone did not support verb bias learning.

In sum, although it is not feasible to evaluate the exact role of event information given different experiment design between Experiment 4, the ERP experiments and the rest in the current study, the event information might have played a facilitating role in verb bias learning.

Further investigation manipulating the frequency of instrument-use in the event-distributional information is necessary to further address the contribution of events in the verb bias learning.

4.3 Ambiguous Context Boosts Effective Learning

The weak training effect on sentence ambiguity resolution revealed in Experiment 3 raised a question about whether such phenomenon was derived from less effective retrieval due to the lack of syntactic similarity, or less effective encoding because the training sentences were unambiguous. The ERP evidence suggests that verb bias was learned more efficiently in ambiguous sentences than in unambiguous sentences. The lack of N400-P600 shift in the unambiguous conditions suggests that the unambiguous *using* and *that has* had prevented participants from effectively generating the association between the verbs and the structures. Individual subjects who showed stronger sensitivity to these disambiguating cues learned worse in the unambiguous training condition.

These results are consistent with the prediction from the implicit learning model (Chang et al., 2006), the existence of ambiguity during training increased the frequency of errors and further strengthen the connection between the verb and the structure. The variability of *with*-PP phrase interpretation involves active practice of ambiguity resolution. At the disambiguating PP-object noun, learners compared their intended interpretation to the actual sentence continuation. On subsequent sentences they either learned to update their prediction to compensate any error, or to reinforce their previous prediction if consistent with the sentence. On the contrary, unambiguous sentences permit less space for errors. According to the error-based learning theory, the link between the verbs and the structures in the unambiguous sentences is expected to be weaker. As a matter of fact, in psycholinguistic literature, various types of verb bias have

been investigated in the context of syntactic ambiguity, including main verb/reduced relatives (e.g. Duffy, Morris, & Rayner, 1988), direct object/sentential complement (e.g. Trueswell et al., 1993), as well as the PP-attachment ambiguities (e.g. Snedeker & Trueswell, 2004). It might be the case that the active ambiguity resolution in sentence processing interacts with the linguistic distributional information in boosting effective verb bias learning.

4.4 Future Direction

Current research can be expanded in three directions as follows.

From the psycholinguistic perspective, as discussed above in 4.2, direct manipulation of the likelihood of instrument use would be the next step to address how event-distributional information contributes to verb bias learning. In the example (1) below, the unknown *veebing* event does not involve the usage of an instrument, however the modifier *PP*-object nouns obviously have instrument semantics. The prediction is that if listeners simply associate the instrument semantic category with the verb based on the co-occurrence frequency, they will erroneously anticipate an instrument upon hearing the same verb and show the similar fixation pattern as that in instrument-trained trials. But if listeners attach the event distributional information of instrument use to the verb, we will expect the similar training effect as the modifier-training dialogues.

(1) - *Which game character did you veeb yesterday?*

- *I veebed the one with the magic sword!*

- *Really? He is too slow in speed! I would veeb the character with the sharp ax!*

From the developmental perspective, sensitivity to verb bias emerged early in development. Indeed the current study showed similar learning effect across age range in eye-

movement pattern when children and adults applied newly learned verb bias during online ambiguity resolution. However, a set of statistical learning experiments comparing children and adults' behavior in artificial language learning paradigm revealed profound differences. When learning the probabilistic usage of determiners, children tended to regularize by generalizing the most frequently syntactic rule across items, while adults' behavior fit well with the probabilistic information in the training (Hudson Kam & Newport, 2005, 2009). These data raise interesting paradoxical asymmetry between statistical learning and usage of statistical information across age ranges. By tracking children's online encoding process of linguistic statistical information, we will be able to address whether there is qualitative difference in statistical learning process between children and adults. How does generalization interact with effective statistical learning? When children only have limited amount of information at the earlier stage of learning, do they actively make abrupt prediction and learn from their errors?

From the perspective of cognitive neuroscience, I am interested in further identifying the neural measures that predict statistical learning efficacy in language. Several recent studies in the field of attention and memory revealed the more-prepared neural state *vs.* the less-prepared neural state. Activation in parahippocampal cortex predicted participants' learning efficacy of novel scenes (Yoo et al., 2012). Moreover the pre-stimulus phase of EEG alpha rhythm reliably predicted visual awareness (Mathewson, Gratton, Fabiani, Beck, & Ro, 2009). In the current study, the N400-P600 transition during the training process has the potential to serve as a diagnostic tool for effective linguistic distributional learning. Future analysis will explore the neural correlates of effective conflict detection, by comparing the ERPs in the trials where subjects correctly rejected the sentence versus the trials where subjects falsely accepted the sentence.

4.5 Conclusions

The present study tests the possibility whether children and adults' verb bias is influenced by the statistical information distributed in a very brief language experience. The results suggest a highly dynamic learning system that constantly collects statistical information and applied the newly learned information into online comprehension. Both eye-movement and ERP evidence suggest newly learned verb bias plays a similar role as the existing verb bias knowledge in affecting language users' parsing commitment and online ambiguity resolution. Importantly, these results support the central role of linguistic distributional information in verb bias learning. The co-occurrence frequency of a particular verb and its preferred structures is critical for effective retrieval of verb-specific information during online sentence parsing and such effect is independent from verb semantic knowledge.

ERP evidence during online verb bias learning suggested separate neural stages that transits from semantic prediction to syntactic rule-based processing as learners continuously collected distributional information of verb-specific structural preference. In addition, establishing the association between individual verbs and their preferred structures was more efficient if the process involved active ambiguity resolution than if the parsing role of verb bias was overridden by other disambiguating cues, which suggests resolving ambiguity might be a crucial component of verb bias learning.

Results also revealed large variability between subjects in terms of learning efficacy. Individuals who were highly sensitive to familiar verb bias showed greater use of newly learned verb bias during conflict detection as well as ambiguity resolution. These findings point toward a shared mechanism underlying verb bias learning in the lab setting and in the natural world.

BIBLIOGRAPHY

- Abla, D., Katahira, K., & Okanoya, K. (2008). On-line Assessment of Statistical Learning by Event-related Potentials. *Journal of Cognitive Neuroscience*, 20(6), 952–964.
- Agresti, A. (2007). *An introduction to categorical data analysis*. John Wiley & Sons.
- Altmann, G. T. M., & Kamide, Y. (1999). Incremental interpretation at verbs: restricting the domain of subsequent reference. *Cognition*, 73(3), 247–264.
- Altmann, G. T. M. (2002). Statistical Learning in Infants. *Proceedings of the National Academy of Sciences*, 99(24), 15250–15251
- Anderson, J. L., Morgan, J. L., & White, K. S. (2003). A Statistical Basis for Speech Sound Discrimination. *Language and Speech*, 46(2-3), 155–182.
- Argaman, V., & Pearlmuter, N. J. (2002). Lexical semantics as a basis for argument structure frequency biases. *The lexical basis of sentence processing: formal, computational, and experimental issues*. John Benjamins Publishing Company.
- Arunachalam, S., & Waxman, S. R. (2010). Meaning from syntax: Evidence from 2-year-olds. *Cognition*, 114(3), 442–446.
- 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(4), 390–412.
- Barr, D. J. (2008). Analyzing “visual world” eyetracking data using multilevel logistic regression. *Journal of Memory and Language*, 59(4), 457–474.
- Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (Submitted). Random effects structure in mixed-effects models: Keep it maximal.

- Baxter, L. C., Saykin, A. J., Flashman, L. A., Johnson, S. C., Guerin, S. J., Babcock, D. R., & Wishart, H. A. (2003). Sex differences in semantic language processing: A functional MRI study. *Brain and Language*, 84(2), 264–272.
- Bentin, S., McCarthy, G., & Wood, C. C. (1985). Event-related potentials, lexical decision and semantic priming. *Electroencephalography and Clinical Neurophysiology*, 60(4), 343–355.
- Bernolet, S., & Hartsuiker, R. J. (2010). Does verb bias modulate syntactic priming? *Cognition*, 114(3), 455–461.
- Bienvenue, B., Mauner, G., & Koenig, J. P. (2007). Anticipating instruments. Presented at the 20th Annual CUNY Conference on Human Sentence Processing, San Diego, CA.
- Bock, K., & Griffin, Z. M. (2000). The persistence of structural priming: transient activation or implicit learning? *Journal of Experimental Psychology. General*, 129(2), 177–192.
- Boland, J. E. (2005). Visual arguments. *Cognition*, 95(3), 237–274.
- Boland, J. E., Tanenhaus, M. K., & Garnsey, S. M. (1990). Evidence for the immediate use of verb control information in sentence processing. *Journal of Memory and Language*, 29(4), 413–432.
- Brysbaert, M., & New, B. (2009). Moving beyond Kučera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods*, 41(4), 977–990.
- Carlson, G. & Tanenhaus, M. (1988). Thematic Roles and Language Comprehension. In W. Wilkins, (ed.), *Syntax and Semantics*, Vol. 21: Thematic Relations, Academic Press, 263-289.

- Chambers, K. E., Onishi, K. H., & Fisher, C. (2003). Infants learn phonotactic regularities from brief auditory experience. *Cognition*, 87(2), B69–B77.
- Chang, F., Dell, G. S., Bock, K., & Griffin, Z. M. (2000). Structural priming as implicit learning: a comparison of models of sentence production. *Journal of Psycholinguistic Research*, 29(2), 217–229.
- Chang, F., Dell, G. S., & Bock, K. (2006). Becoming syntactic. *Psychological Review*, 113(2), 234–272.
- Chomsky, N. (1965). *Aspects of the theory of syntax*. MIT Press.
- Christiansen, M. H., Conway, C. M., & Onnis, L. (In Press). Similar neural correlates for language and sequential learning: Evidence from event-related brain potentials. *Language and Cognitive Processes*.
- Conway, C. M., Bauernschmidt, A., Huang, S. S., & Pisoni, D. B. (2010). Implicit statistical learning in language processing: Word predictability is the key. *Cognition*, 114(3), 356–371.
- Coulson, S., King, J. W., & Kutas, M. (1998). Expect the Unexpected: Event-related Brain Response to Morphosyntactic Violations. *Language and Cognitive Processes*, 13(1), 21–58.
- Coyle, J. M., & Kaschak, M. P. (2008). Patterns of experience with verbs affect long-term cumulative structural priming. *Psychonomic Bulletin & Review*, 15(5), 967–970.
- DeLong, K. A., Urbach, T. P., & Kutas, M. (2005). Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. *Nature Neuroscience*, 8(8), 1117–1121.

- Delorme, A., & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, 134(1), 9–21.
- Duffy, S. A., Morris, R. K., & Rayner, K. (1988). Lexical ambiguity and fixation times in reading. *Journal of Memory and Language*, 27(4), 429–446.
- Ellis, N. C. (2006). Selective Attention and Transfer Phenomena in L2 Acquisition: Contingency, Cue Competition, Salience, Interference, Overshadowing, Blocking, and Perceptual Learning. *Applied Linguistics*, 27(2), 164–194.
- Federmeier, K. D., & Kutas, M. (1999). A Rose by Any Other Name: Long-Term Memory Structure and Sentence Processing. *Journal of Memory and Language*, 41(4), 469–495.
- Federmeier, K. D., & Kutas, M. (2001). Meaning and modality: Influences of context, semantic memory organization, and perceptual predictability on picture processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 27(1), 202–224.
- Federmeier, K. D., Wlotko, E. W., De Ochoa-Dewald, E., & Kutas, M. (2007). Multiple effects of sentential constraint on word processing. *Brain research*, 1146, 75–84.
- Fernald, A., Zangl, R., Portillo, A.L., & Marchman, V.A. (2008). Looking while listening: Using eye movements to monitor spoken language comprehension by infants and young children. In Sekerina, I.A., Fernandez, E.M., & Clahsen, H. (Eds.) *Developmental Psycholinguistics: On-line methods in children's language processing*.
- Ferretti, T. R., McRae, K., & Hatherell, A. (2001). Integrating Verbs, Situation Schemas, and Thematic Role Concepts. *Journal of Memory and Language*, 44(4), 516–547.

- Fine, A. B., & Jaeger, F. T. (2011). Language comprehension is sensitive to changes in the reliability of lexical cues. *Proceedings of the 33rd Annual Meeting of the Cognitive Science Society*. 925–930.
- Fiser, J., & Aslin, R. N. (2001). Unsupervised Statistical Learning of Higher-Order Spatial Structures from Visual Scenes. *Psychological Science*, 12(6), 499–504.
- Fiser, J., & Aslin, R. N. (2002). Statistical learning of higher-order temporal structure from visual shape sequences. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, 28(3), 458–467.
- Fisher, C., Gertner, Y., Scott, R. M., & Yuan, S. (2010). Syntactic bootstrapping. *Wiley Interdisciplinary Reviews: Cognitive Science*, 1(2), 143–149.
- Francis, W. N., Kučera, H., & Mackie, A. W. (1982). *Frequency analysis of English usage: lexicon and grammar*. Houghton Mifflin.
- Friederici, A. D., Hahne, A., & Saddy, D. (2002). Distinct neurophysiological patterns reflecting aspects of syntactic complexity and syntactic repair. *Journal of Psycholinguistic Research*, 31(1), 45–63.
- Friederici, A. D., Steinhauer, K., & Pfeifer, E. (2002). Brain signatures of artificial language processing: Evidence challenging the critical period hypothesis. *Proceedings of the National Academy of Sciences*, 99(1), 529–534.
- Gaillard, A. W. K., & Verduin, C. J. (1985). *Comparisons across paradigms: An ERP study*. Attention and performance (M.I. Posner & O.S.M. Marin., Vols. 1-11). Hillsdale, NJ: Erlbaum.

- Ganis, G., Kutas, M., & Sereno, M. I. (1996). The search for common sense: An electrophysiological study of the comprehension of words and pictures in reading. *Journal of Cognitive Neuroscience*, 8(2), 89–106.
- Garnsey, S. M., Pearlmutter, N. J., Myers, E., & Lotocky, M. A. (1997). The Contributions of Verb Bias and Plausibility to the Comprehension of Temporarily Ambiguous Sentences. *Journal of Memory and Language*, 37(1), 58–93.
- Gerken, L., Wilson, R., & Lewis, W. (2005). Infants Can Use Distributional Cues to Form Syntactic Categories. *Journal of Child Language*, 32(02), 249–268.
- Gleitman, L. (1990). The Structural Sources of Verb Meanings. *Language Acquisition*, 1(1), 3–55.
- Gomez, R. L., & Gerken, L. (1999). Artificial grammar learning by 1-year-olds leads to specific and abstract knowledge. *Cognition*, 70(2), 109–135.
- Greenhouse, S. W., & Geisser, S. (1959). On methods in the analysis of profile data. *Psychometrika*, 24(2), 95–112.
- Grimshaw, J. (1979). Complement Selection and the Lexicon. *Linguistic Inquiry Amherst, Mass.*, 10(2), 279–326.
- Grosjean, F. (1998). Studying Bilinguals: Methodological and Conceptual Issues 1. *The Handbook of Bilingualism*. Blackwell Publishing Ltd. 32-63.
- Hagoort, P., Brown, C., & Groothusen, J. (1993). The syntactic positive shift (sps) as an erp measure of syntactic processing. *Language and Cognitive Processes*, 8(4), 439–483.
- Hagoort, P., Brown, C. M., & Osterhout, L. (1999). *The neurocognition of syntactic processing*.
- Hallett, P. E. (1986). Eye movements. *Handbook of Perception and Human Performance I* (K. Boff, L. Kaufman & J. Thomas., p. 10.25–10.28)

- Hansen, J. C., & Hillyard, S. A. (1980). Endogeneous brain potentials associated with selective auditory attention. *Electroencephalography and Clinical Neurophysiology*, 49(3–4), 277–290.
- Hare, M., McRae, K., & Elman, J. L. (2004). Admitting that admitting verb sense into corpus analyses makes sense. *Language and Cognitive Processes*, 19(2), 181–224.
- Hare, M., McRae, K., & Elman, J. L. (2003). Sense and structure: Meaning as a determinant of verb subcategorization preferences. *Journal of Memory and Language*, 48(2), 281–303.
- Hindle, D., & Rooth, M. (1993). Structural ambiguity and lexical relations. *Comput. Linguist.*, 19(1), 103–120.
- Hsu, H. (2009). A neurophysiological study on probabilistic grammatical learning and sentence processing. *Theses and Dissertations*.
- Huang, H. W., Lee, C. L., & Federmeier, K. D. (2010). Imagine that! ERPs provide evidence for distinct hemispheric contributions to the processing of concrete and abstract concepts. *NeuroImage*, 49(1), 1116–1123.
- Hudson Kam, C. L., & Newport, E. L. (2005). Regularizing Unpredictable Variation: The Roles of Adult and Child Learners in Language Formation and Change. *Language Learning and Development*, 1(2), 151–195.
- Hudson Kam, C. L., & Newport, E. L. (2009). Getting it right by getting it wrong: When learners change languages. *Cognitive Psychology*, 59(1), 30–66.
- Itzhak, I., Pauker, E., Drury, J. E., Baum, S. R., & Steinhauer, K. (2010). Event-related potentials show online influence of lexical biases on prosodic processing. *NeuroReport*, 21(1), 8–13.

- Jaeger, T. F. (2008). Categorical Data Analysis: Away from ANOVAs (Transformation or Not) and towards Logit Mixed Models. *Journal of Memory and Language*, 59(4), 434–446.
- Jung, T. P., Makeig, S., Humphries, C., Lee, T. W., McKEOWN, M. J., Iragui, V., & Sejnowski, T. J. (2000). Removing Electroencephalographic Artifacts by Blind Source Separation. *Psychophysiology*, 37(02), 163–178.
- Jurafsky, D. (1996). A probabilistic model of lexical and syntactic access and disambiguation. *Cognitive Science*, 20(2), 137–194.
- Kaan, E., Harris, A., Gibson, E., & Holcomb, P. J. (2000). The P600 as an index of syntactic integration difficulty. *Language and Cognitive Processes*, 15(2), 159–201.
- Kaan, E., & Swaab, T. Y. (2003). Repair, Revision, and Complexity in Syntactic Analysis: An Electrophysiological Differentiation. *Journal of Cognitive Neuroscience*, 15(1), 98–110.
- Kamide, Y., Altmann, G. T., & Haywood, S. L. (2003). The time-course of prediction in incremental sentence processing: Evidence from anticipatory eye movements. *Journal of Memory and Language*, 49(1), 133–156.
- Kansaku, K., Yamaura, A., & Kitazawa, S. (2000). Sex Differences in Lateralization Revealed in the Posterior Language Areas. *Cerebral Cortex*, 10(9), 866–872.
- Kidd, E., & Bavin, E. L. (2007). Lexical and referential influences on on-line spoken language comprehension: A comparison of adults and primary-school-age children. *First Language*, 27(1), 29–52.
- Koenig, J. P., Mauner, G., & Bienvenue, B. (2003). Arguments for adjuncts. *Cognition*, 89(2), 67–103.
- Kotz, S. A. (2009). A critical review of ERP and fMRI evidence on L2 syntactic processing. *Brain and Language*, 109(2–3), 68–74.

- Kutas, M., & Hillyard, S. A. (1984). Brain potentials during reading reflect word expectancy and semantic association. *Nature*, 307(5947), 161–163.
- Kutas, M. (1993). In the company of other words: Electrophysiological evidence for single-word and sentence context effects. *Language and Cognitive Processes*, 8(4), 533–572.
- Laszlo, S., & Federmeier, K. D. (2010). The N400 as a snapshot of interactive processing: Evidence from regression analyses of orthographic neighbor and lexical associate effects. *Psychophysiology*.
- Levin, B. (1993). *English verb classes and alternations: a preliminary investigation*. University of Chicago Press.
- Luck, S. J., & Hillyard, S. A. (1994). Electrophysiological correlates of feature analysis during visual search. *Psychophysiology*, 31(3), 291–308.
- Luck, S., & Lopez-Calderon, J. (2011). ERPLAB Toolbox.
- MacDonald, M. C., Pearlmutter, N. J., & Seidenberg, M. S. (1994). The lexical nature of syntactic ambiguity resolution. *Psychological Review*, 101(4), 676–703.
- Mäntysalo, S., & Gaillard, A. W. K. (1986). Event-related potentials (ERPs) in a learning and memory test. *Biological Psychology*, 23(1), 1–20.
- Mason, R. A., Just, M. A., Keller, T. A., & Carpenter, P. A. (2003). Ambiguity in the Brain: What Brain Imaging Reveals About the Processing of Syntactically Ambiguous Sentences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 29(6), 1319–1338.
- Mathewson, K. E., Gratton, G., Fabiani, M., Beck, D. M., & Ro, T. (2009). To See or Not to See: Prestimulus A Phase Predicts Visual Awareness. *The Journal of Neuroscience*, 29(9), 2725–2732.

- Maye, J., Weiss, D. J., & Aslin, R. N. (2008). Statistical phonetic learning in infants: facilitation and feature generalization. *Developmental Science*, 11(1), 122–134.
- Maye, J., Werker, J. F., & Gerken, L. (2002). Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition*, 82(3), B101–111.
- McLaughlin, J., Osterhout, L., & Kim, A. (2004). Neural correlates of second-language word learning: minimal instruction produces rapid change. *Nature Neuroscience*, 7(7), 703–704.
- McLaughlin, J., Tanner, D., Pitkänen, I., Frenck - Mestre, C., Inoue, K., Valentine, G., & Osterhout, L. (2010). Brain Potentials Reveal Discrete Stages of L2 Grammatical Learning. *Language Learning*, 60, 123–150.
- McNealy, K., Mazziotta, J. C., & Dapretto, M. (2006). Cracking the Language Code: Neural Mechanisms Underlying Speech Parsing. *The Journal of Neuroscience*, 26(29), 7629–7639.
- McRae, K., Ferretti, T. R., & Amyote, L. (1997). Thematic Roles as Verb-specific Concepts. *Language and Cognitive Processes*, 12(2-3), 137–176.
- McRae, K., Hare, M., Elman, J. L., & Ferretti, T. (2005). A basis for generating expectancies for verbs from nouns. *Memory & Cognition*, 33(7), 1174–1184.
- Mestres-Missé, A., Rodriguez-Fornells, A., & Münte, T. F. (2007). Watching the Brain During Meaning Acquisition. *Cerebral Cortex*, 17(8), 1858–1866.
- Midgley, K. J., Holcomb, P. J., & Grainger, J. (2009). Language effects in second language learners and proficient bilinguals investigated with event-related potentials. *Journal of Neurolinguistics*, 22(3), 281–300.

- Misyak, J. B., Christiansen, M. H., & Bruce Tomblin, J. (2010). Sequential Expectations: The Role of Prediction - Based Learning in Language. *Topics in Cognitive Science*, 2(1), 138–153.
- Morgan-Short, K., Sanz, C., Steinhauer, K., & Ullman, M. T. (2010). Second Language Acquisition of Gender Agreement in Explicit and Implicit Training Conditions: An Event-Related Potential Study. *Language Learning*, 60(1), 154–193.
- Naigles, L. (1990). Children use syntax to learn verb meanings. *Journal of Child Language*, 17(2), 357–374.
- Novick, J. M., Thompson-Schill, S. L., & Trueswell, J. C. (2008). Putting lexical constraints in context into the visual-world paradigm. *Cognition*, 107(3), 850–903.
- Osterhout, L, Holcomb, P. J., & Swinney, D. A. (1994). Brain potentials elicited by garden-path sentences: evidence of the application of verb information during parsing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20(4), 786–803.
- Osterhout, L, & Holcomb, P. J. (1992). Event-related brain potentials elicited by syntactic anomaly. *Journal of Memory and Language*, 31(6), 785–806.
- Osterhout, L, Holcomb, P. J., & Swinney, D. A. (1994). Brain potentials elicited by garden-path sentences: Evidence of the application of verb information during parsing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20(4), 786–803.
- Osterhout, L, McLaughlin, J., Pitkanen, I., Frenck-Mestre, C., & Molinaro, N. (2006). Novice Learners, Longitudinal Designs, and Event-Related Potentials: A Means for Exploring the Neurocognition of Second Language Processing. *Language Learning*, 56, 199–230.

- Osterhout, L., Poliakov, A., Inoue, K., McLaughlin, J., Valentine, G., Pitkanen, I., Frenck-Mestre, C., et al. (2008). Second-language learning and changes in the brain. *Journal of neurolinguistics*, 21(6), 509–521.
- Pantel, P., & Lin, D. (2000). An Unsupervised Approach to Prepositional Phrase Attachment using Contextually Similar Words. *Proceedings of the 38th Annual Meeting on Association for Computational Linguistics*, 101–108.
- Patel, A. D., Gibson, E., Ratner, J., Besson, M., & Holcomb, P. J. (1998). Processing Syntactic Relations in Language and Music: An Event-Related Potential Study. *Journal of Cognitive Neuroscience*, 10(6), 717–733.
- Pearce, J. M., Graham, M., Good, M. A., Jones, P. M., & McGregor, A. (2006). Potentiation, overshadowing, and blocking of spatial learning based on the shape of the environment. *Journal of Experimental Psychology: Animal Behavior Processes*, 32(3), 201-214.
- Qi, Z., Jackson, S., & Garnsey, S. (2010). The left hemisphere knows more about verbs (in most people). Presented at the Neurobiology of Language, San Diego, CA.
- Quené, H., & van den Bergh, H. (2004). On multi-level modeling of data from repeated measures designs: a tutorial. *Speech Communication*, 43(1-2), 103–121.
- Roland, D., Elman, J. L., & Ferreira, V. S. (2006). Why is that? Structural prediction and ambiguity resolution in a very large corpus of English sentences. *Cognition*, 98(3), 245–272.
- Roser, M. E., Fiser, J., Aslin, R. N., & Gazzaniga, M. S. (2010). Right Hemisphere Dominance in Visual Statistical Learning. *Journal of Cognitive Neuroscience*, 23(5), 1088–1099.

- Rüschemeyer, S. A., Fiebach, C. J., Kempe, V., & Friederici, A. D. (2005). Processing lexical semantic and syntactic information in first and second language: fMRI evidence from German and Russian. *Human Brain Mapping, 25*(2), 266–286.
- Rüschemeyer, S. A., Zysset, S., & Friederici, A. D. (2006). Native and non-native reading of sentences: An fMRI experiment. *NeuroImage, 31*(1), 354–365.
- Saffran, J. R. (2001). Words in a sea of sounds: the output of infant statistical learning. *Cognition, 81*(2), 149–169.
- Saffran, Jenny R, Aslin, R. N., & Newport, E. L. (1996). Statistical Learning by 8-Month-Old Infants. *Science, 274*(5294), 1926–1928.
- Scott, R. M., & Fisher, C. (2009). Two-year-olds use distributional cues to interpret transitivity-alternating verbs. *Language and cognitive processes, 24*(6), 777–803.
- Shaywitz, B. A., Shaywitz, S. E., Pugh, K. R., Constable, R. T., Skudlarski, P., Fulbright, R. K., Bronen, R. A., et al. (1995). Sex differences in the functional organization of the brain for language. *Nature, 373*(6515), 561–562.
- Shtyrov, Y., Nikulin, V. V., & Pulvermüller, F. (2010). Rapid Cortical Plasticity Underlying Novel Word Learning. *The Journal of Neuroscience, 30*(50), 16864–16867.
- Snedeker, J., & Trueswell, J. C. (2004). The developing constraints on parsing decisions: The role of lexical-biases and referential scenes in child and adult sentence processing. *Cognitive Psychology, 49*(3), 238–299.
- Snedeker, J., & Yuan, S. (2008). Effects of prosodic and lexical constraints on parsing in young children (and adults). *Journal of memory and language, 58*(2), 574–608.
- Sussman, R. S. (2006). Verb-instrument information during on-line processing. *Theses and Dissertations*

- Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science*, 268(5217), 1632.
- Tatsuno, Y., & Sakai, K. L. (2005). Language-Related Activations in the Left Prefrontal Regions Are Differentially Modulated by Age, Proficiency, and Task Demands. *The Journal of Neuroscience*, 25(7), 1637–1644.
- Thothathiri, M., & Snedeker, J. (2008). Give and take: Syntactic priming during spoken language comprehension. *Cognition*, 108(1), 51–68.
- Toglia, M. P., & Battig, W. F. (1978). *Handbook of semantic word norms*. Lawrence Erlbaum Associates.
- Tooley, K. M., Traxler, M. J., & Swaab, T. Y. (2009). Electrophysiological and behavioral evidence of syntactic priming in sentence comprehension. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, 35(1), 19–45.
- Trueswell, J. C., & Kim, A. E. (1998). How to prune a garden path by nipping it in the bud: Fast priming of verb argument structure. *Journal of Memory and Language*, 39, 102–123.
- Trueswell, J. C., Sekerina, I., Hill, N. M., & Logrip, M. L. (1999). The kindergarten-path effect: Studying on-line sentence processing in young children. *Cognition*, 73(2), 89–134.
- Trueswell, J. C., Tanenhaus, M. K., & Kello, C. (1993). Verb-specific constraints in sentence processing: Separating effects of lexical preference from garden-paths. *Journal Of Experimental Psychology Learning Memory And Cognition*, 19, 528–528.
- Trueswell, J. C. (1996). The Role of Lexical Frequency in Syntactic Ambiguity Resolution. *Journal of Memory and Language*, 35(4), 566–585.

- Tse, C. Y., Lee, C. L., Sullivan, J., Garnsey, S. M., Dell, G. S., Fabiani, M., & Gratton, G. (2007). Imaging cortical dynamics of language processing with the event-related optical signal. *Proceedings of the National Academy of Sciences*, 104(43), 17157–17162.
- Turk-Browne, N. B., Jungé, J. A., & Scholl, B. J. (2005). The Automaticity of Visual Statistical Learning. *Journal of Experimental Psychology: General*, 134(4), 552–564.
- Van Petten, C., & Kutas, M. (1990). Interactions between sentence context and word frequency in event-related brain potentials. *Memory & Cognition*, 18(4), 380–393.
- Van Petten, C., Kutas, M., Kluender, R., Mitchiner, M., & McIssac, H. (1991). Fractionating the Word Repetition Effect with Event-Related Potentials. *Journal of Cognitive Neuroscience*, 3(2), 131–150.
- Weighall, A. R. (2008). The kindergarten path effect revisited: Children's use of context in processing structural ambiguities. *Journal of Experimental Child Psychology*, 99(2), 75–95.
- Wells, J. B., Christiansen, M. H., Race, D. S., Acheson, D. J., & MacDonald, M. C. (2009). Experience and Sentence Processing: Statistical Learning and Relative Clause Comprehension. *Cognitive Psychology*, 58(2), 250–271.
- Wilson, M. (1988). MRC psycholinguistic database: Machine-usable dictionary, version 2.00. *Behavior Research Methods, Instruments, & Computers*, 20(1), 6–10.
- Wlotko, E. W., & Federmeier, K. D. (2007). Finding the right word: Hemispheric asymmetries in the use of sentence context information. *Neuropsychologia*, 45(13), 3001–3014.
- Wonnacott, E., Newport, E. L., & Tanenhaus, M. K. (2008). Acquiring and processing verb argument structure: Distributional learning in a miniature language. *Cognitive Psychology*, 56(3), 165–209.

- Yoo, J. J., Hinds, O., Ofen, N., Thompson, T. W., Whitfield-Gabrieli, S., Triantafyllou, C., & Gabrieli, J. D. E. (2012). When the brain is prepared to learn: enhancing human learning using real-time fMRI. *NeuroImage*, 59(1), 846–852.
- Yuan, S., & Fisher, C. (2009). “Really? She Blinked the Baby?” Two-Year-Olds Learn Combinatorial Facts About Verbs by Listening. *Psychological Science*, 20(5), 619–626.
- Yuan, S., Fisher, C., Kandhadai, P., & Fernald, A. (2011). You can stipe the pig and nerk the fork: Learning to use verbs to predict nouns. *Proceedings of the 35th Annual Boston University Conference on Language Development*, Cascadilla Press, 665-677.