TRENDS IN
LANGUAGE ACQUISITION RESEARCH 24

Semantics in Language Acquisition

Edited by Kristen Syrett and Sudha Arunachalam

JOHN BENJAMINS PUBLISHING COMPANY

Semantics in Language Acquisition

Trends in Language Acquisition Research

ISSN 1569-0644

TiLAR publishes monographs, edited volumes and text books on theoretical and methodological issues in the field of child language research. The focus of the series is on original research on all aspects of the scientific study of language behavior in children, linking different areas of research including linguistics, psychology & cognitive science.

For an overview of all books published in this series, please see
http://benjamins.com/catalog/tilar

Volume 24

Semantics in Language Acquisition
Edited by Kristen Syrett and Sudha Arunachalam

Semantics in Language Acquisition

Edited by

Kristen Syrett
Rutgers University, New Brunswick

Sudha Arunachalam
Boston University

John Benjamins Publishing Company
Amsterdam / Philadelphia

The paper used in this publication meets the minimum requirements of the American National Standard for Information Sciences – Permanence of Paper for Printed Library Materials, ANSI Z39.48-1984.

DOI 10.1075/tilar.24

Cataloging-in-Publication Data available from Library of Congress.

ISBN 978 90 272 0137 9 (HB)
ISBN 978 90 272 6360 5 (E-BOOK)

John Benjamins Publishing Company · https://benjamins.com

Table of contents

CHAPTER 1

The historical emergence and current study of semantics in acquisition

Kristen Syrett
Rutgers University

A child acquiring language – be it English, German, Tamil, Japanese, Mandarin Chinese, Navajo, American Sign Language, or any other of the thousands of languages in the world – is faced with the task of mastering multiple components of that language. Some of the most basic components may be rapidly picked up from the language to which the child is exposed: word order or mappings of certain phonological forms to possible meanings. Other components, such as the constraints or rules that guide the language acquisition process – some of which appear to apply universally across languages, and some of which are language-specific – may be part of an innately endowed language-specific system, or constructed by abstracting over multiple occurrences of patterns. The challenges faced by the young language learner, and the strategies and timeline according to which children meet these challenges, have been the focus of language acquisition research by linguists and developmental psychologists for decades now.

A major inspiration for research programs in this vein were the proposals set forth in Chomsky's foundational book *Aspects on the Theory of Syntax* (1965). In *Aspects*, Chomsky put research on language acquisition at the forefront, making it a major goal for linguistic theory to be rich enough to account for how rapidly and effortlessly the young child seems to acquire language universally and at the same time allow for cross-linguistic diversity. It was here that Chomsky proposed that children are guided by an innate *language acquisition device* (a biologically-endowed mental capacity to acquire language), a concept that later evolved into a set of universal principles to be called *Universal Grammar* (UG). The idea of a Universal Grammar was simple and compelling, because it helped to explain a great deal of both the universality and diversity of languages: every child was said to be equipped with similar assumptions and expectations about the processes and principles that govern language, but every child must also discover how these processes and principles are instantiated in the language(s) they are acquiring. In this way, there was an interplay between the innate components of grammar present from the very beginning, and the ones dependent upon experience and linguistic input from the

https://doi.org/10.1075/tilar.24.01syr

caregiver. The very concept of UG changed the way that researchers approached the study of language and language acquisition, but it would also become a matter of fierce debate in years to come.

Chomsky proposed a fundamental distinction between *competence* and *performance* – a distinction between what a language user knows about their language (without even knowing they know it) and how they actually use language. This competence/performance distinction carries clear implications for research on language acquisition and our ability to extrapolate what a child *knows* from what a child *does* when using language or performing in an experimental study. We cannot (and should not) assume that children's early linguistic production – such as the single and telegraphic multi-word utterances carefully documented in Brown (1973), Bloom (1970), and elsewhere – adequately captures their full and always developing linguistic capacity.

Around the same time that *Aspects* was published, Ross's (1967) dissertation *Constraints on variables in syntax* appeared, and served as a rich source of syntactic phenomena (identified with creative and now-standard terminology) to be studied for years to come, including islands, sluicing, pruning, scrambling, and other constraints and conditions on movement. Subsequent publications in the field of linguistics brought to light the challenges and appeal of investigating phenomena such as *wh*-movement (Chomsky, 1977), case, anaphora and binding constraints involving pronouns, reflexives, and names (Chomsky, 1980; Reinhart, 1983), null subjects/objects, pro, PRO, raising and control (Bresnan, 1982; Rizzi, 1986; Williams, 1980). Years later, in *Lectures on Government and Binding* (1981) Chomsky proposed a system of principles and parameters that constrained linguistic representations, and set forth X-bar theory, a theory governing syntactic phrase structure.

The overall result was that during this time, linguistic theory was highlighting specific structural (syntax-specific) aspects of language that every child needed to acquire, along with specific features that children acquiring certain languages would need to master. A common goal of researchers was to take linguistic theory as a starting point for targeting specific aspects of children's language acquisition process, and to arrive at a linguistic theory with explanatory adequacy that could formally account for behavioral data collected from children and children's frequent deviation from adult production, and reveal their underlying competence at various stages in development. Research in language acquisition was therefore focused on syntax during this period of time, because it was in this subfield of linguistics that Chomsky's proposals could tested, and researchers could gather evidence that spoke directly to the cross-linguistic universals that may or may not govern the structure of language, the cross-linguistic variation that it was possible to observe, and the processes by which children might acquire syntax, be it through an innate capacity or by generalizations over patterns in the input.

Over a number of years, a convincing and powerful argument was being made that formal semantics had a firm place in the growing field of cognitive science, and that the investigation of meaning was built upon an understanding of lexical and compositional semantics, syntactic structures, conceptual representations, and an understanding of the link between language and cognition. This move had started years before. Because in *Aspects*, Chomsky talked of linguistic theory as being concerned with the mental or psychological reality underlying behavior, linguistics not only found itself firmly established in the cognitive sciences, but empirical behavioral data and introspective judgments were becoming a window into more abstract linguistic representations and mechanisms. Fodor's *The Language of Thought* (1975) made an explicit link between the syntax of linguistics and mental representations in thought, both inspiring future lines of work and sparking further debate. Jackendoff's *Semantics and Cognition* (1983) and *Semantic Structures* (1990) emphasized this connection clearly, as did Barbara Partee's (1995) chapter 'Lexical semantics and compositionality' in Gleitman, Liberman, & Osherson's *An Invitation to Cognitive Science*. Adele Goldberg's *Constructions* (1995), too, reinforced this link with the novel proposal that constructions themselves conveyed meaning beyond individual words and their composition, offering an alternative to generative approaches of language acquisition and learning, which would inspire lines of research investigating how children could acquire abstract constructions and generalize over patterns via similarity and analogy.

As attention to the subfield of semantics in new areas grew, so did research in word learning. Again, this was by no means a new phenomenon. To take one well-known example, hundreds of years prior, in the 17th century, John Locke (1690/1894) had highlighted the challenges of word learning in discussions about the relation between a word, an idea, and reference. But researchers were now equipped with sophisticated linguistic and cognitive theory, data, and methodologies to move beyond Locke's early form of an associationist account of word learning and make more specific proposals about the process of acquiring meaning based on recent advances in lexical semantics and the mapping between syntactic structure, conceptual structure, and semantic representations. In the 1970s and 1980s, semantic feature theory, and the connection between linguistic and conceptual categories, played a prominent role in discussions of the acquisition of the lexicon, and the progression through stages of lexical understanding in language development (see e.g., Clark, 1973; Bowerman, 1978, 1980).

In the 1990s, the syntax-semantics interface in language acquisition had attained a central position in the field. Landau and Gleitman (1985) and Gleitman (1990) capitalized on this tight linking between syntax and semantics – specifically that the structures in which verbs appear are projections of the meaning of these verbs – to propose a theory of syntactic bootstrapping: that children could deduce something about the meaning of words (verbs specifically) by paying attention to the syntactic environments in which they appear in the input. Support for this hypothesis came from their (1985) analysis of interactions between a mother and her blind child prior to the point in development when the child was producing vision-related words. A close examination of the patterns that surfaced in this corpus revealed that it was not just the proximity of the object to be explored that distinguished lexical meaning between verbs like *look* and *see*, but rather, that the subcategorization frames and linguistic environments in which these words appeared seemed to provide reliable cues to the learner about differences in meaning – and that attention to extralinguistic information could be modulated by sensitivity to linguistic information. That such a strategy could account for how even a blind child without the ability to perceive visual events could differentiate between *look* and *see* is compelling evidence that perception need not be at the forefront of word learning, and that children recruit structural patterns that can be abstracted from caregiver input.

In a complementary vein, Pinker (1989, 1994b) suggested how this connection between syntax/semantics could allow the learner to acquire argument structure: the learner pays attention to the situations in which a verb is used, and maps event participants onto arguments with thematic roles projected from the verb meaning. Gleitman and Landau brought together a number of prominent linguists, psychologists, and developmentalists for contributions to the volume *Acquisition of the Lexicon* (1994), which laid out the conceptual and linguistic issues inherent to word learning, and verbs in particular. But many of the same concerns – the relation between conceptual and linguistic categories, for example – apply equally to other predicates, such as adjectives, as Kamp and Partee (1995) showed quite clearly.

Also in the 1990s, textbooks appeared that made it easier for students and researchers of linguistics to become acquainted with the tools of formal logic and semantics and to apply them to issues relevant to cognitive science, including language acquisition and psycholinguistics. These included the two-volume (1991) *Logic, Language, and Meaning* published by notable logicians and semanticists under the pseudonym L. T. F. Gamut, Chierchia & McConnell-Ginet's (1993) *Meaning and Grammar*, and Heim & Kratzer's (1998) *Semantics in Generative Grammar*. These texts continue to be standard textbooks used in graduate semantics courses, because they draw upon decades of theoretical advances such as Montague grammar (Montague 1973, 1974; see Partee, 2006), and not only make such formalism accessible to students of semantics, but illustrate its relevance to capturing and investigating a range of natural language phenomena.

Montague grammar served as a main starting point of formal semantics because of the merger of the techniques of logic with natural language and its own "Universal Grammar," which highlighted the mathematical foundations of language, the truth conditional components of semantics, entailment relations, the Fregean-inspired principle of compositionality, and the homomorphism between syntax and semantics. In subsequent years, key theoretical publications would come to shape the way researchers accounted for the interpretations that were available (or not) in natural language expressions, and the kinds of sentences they targeted in their research. These included, for example, May's (1977, 1985) proposal of Quantifier Raising – that a quantificational phrase moves covertly through the structure, allowing for scopal interaction; Ladusaw's (1979) appeal to downward entailing environments as licensors for Negative Polarity Items (NPIs) (e.g., *any*); Barwise & Cooper's (1981) work on generalized quantifiers and monotonicity; and Heim's (1982) idea of partitioning sentences into a tripartite structure (quantifier, restrictor, and nuclear scope) and her account of 'donkey' sentences (*Every farmer who owns a donkey feeds it*) that appealed to quantifier binding.

As a result, in the mid-nineties to early 2000s, and research on children's knowledge of quantifier meaning had begun to surge. Certainly, this was not the first time quantification had drawn interest in the area of language development; Inhelder and Piaget (1958, 1964) had documented children's interpretations of sentences involving universal quantifiers long before this, and quantifiers continued to be of interest to developmentalists for years afterwards (see, for example, Smith 1980; Freeman, Sinha, & Stedman 1980). But during this period of time, key publications on children's knowledge of quantifier meaning began to appear. What distinguished these works from earlier ones, and what fostered the move towards investigations in semantics, was that, now, acquisition researchers were equipped with the theoretical linguistic machinery and formalism that earlier developmentalists lacked. Some research focused on how children interpret universal quantifiers such as *every*, *each*, and *all* vis-à-vis quantifier spreading and entailment relations (Brooks & Braine, 1996; Brooks et al., 2001; Crain et al., 1996; Crain, Meroni, & Minai, 2004; Drozd, 2001; Philip, 1995). Other research focused on the scopal interaction between quantifiers and negation or indefinite expressions (Gualmini, 2003; Lidz & Musolino, 2002; Musolino, 1998). Both lines continue to have a profound impact on the topics and cross-linguistic range of research conducted in the field of acquisition.

Around this same time, two other instructional books were published, which contributed to the rise of acquisition work in semantics in another way. The first, McDaniel, McKee, & Smith Cairns's *Methods for Assessing Children's Syntax* (1996) featured chapters written by experts in the field, detailing in clear, step-by-step fashion how to implement both established and growing methodologies in language acquisition and development, one of which was the Truth Value Judgment Task (TVJT), which was also subsequently widely applied to the acquisition of semantics and pragmatics. This methodology was first introduced in Crain & McKee (1986), but was introduced at length with multiple illustrative and entertaining examples in Crain & Thornton's (1998) *Investigations in Universal Grammar*. Research on semantics in language acquisition continued to flourish throughout the early 2000s, with the topics under investigation broadening to include not only quantification, but also phenomena such as negation and disjunction (Goro & Akiba, 2004; Gualmini & Crain, 2005a, b), the scope of *only* (Gualmini & Schwarz, 2009), and gradable adjectives (Barner & Snedeker, 2008; Syrett 2007; Syrett, Kennedy, & Lidz, 2010).

But this period of time also saw the surge of another area of language acquisition beyond syntax: pragmatics. If semantics is the study of lexical and truth-conditional meaning, pragmatics is the study of meaning in context. Semantic meaning arises from the meaning of words and phrases, and propositional content determined compositionally, whereas pragmatics moves beyond what is said to the meaning that a speaker intends or a listener retrieves when language is used. It may not be so surprising, then, that a great deal of pragmatic meaning takes a long time to acquire.

A number of publications in the seventies served as the impetus for the study of formal pragmatics, beginning with Horn's (1972) dissertation, *On the Semantic Properties of Logical Operators in English*, which served as a springboard. Around this same time, work in semantics and pragmatics by Robert Stalnaker, Lauri Karttunen, and Stanley Peters formalized the notions of presuppositions and how they are triggered by specific lexical expressions and calculated in complex sentences (Karttunen, 1974; Karttunen & Peters, 1979; Stalnaker, 1974). David Lewis illustrated how a speaker and hearer negotiate meaning and truth conditions, and make accommodations in a context as a conversation unfolds (Lewis 1979). Grice's *Logic and Conversation* first appeared in 1975, but began to gain influence when his writings were published posthumously in 1989, the same year that Horn's *A Natural History of Negation* was published. Soon afterwards, more linguists were appealing to the Gricean maxims of quantity, quality, relation, and manner, and the Cooperative Principle. Much later, however, neo-Griceans (Horn, 2004) and Relevance theorists (Wilson & Sperber, 2004) would propose alternative principles to the Gricean Maxims. These theoretical proposals brought presupposition triggers, conversational implicatures – and specifically scalar implicatures – to the foreground, paving the way for work in acquisition and psycholinguistics to pick up on these phenomena as topics of investigation. Clark and Amaral (2010) highlight a number of ways in which lexical acquisition capitalizes upon pragmatic sources of information, including gesture and eye gaze, common ground between speaker and hearer, relevance, and discourse context.

Noveck's seminal (2001) article was perhaps the first paper in acquisition to investigate how older children and adults calculated scalar implicatures on so-called 'Horn scales' with words like *might* and *some*. He showed that children, but not adults, generally accept statements such as *There might be a parrot in the box* when there is a probability of 1.0 that there is a parrot in the box (and it is therefore more felicitous and informative to say that there *must* be a parrot in the box). Likewise, children accept statements such as *Some giraffes have long necks* when our experience in the real world tells us that *all* giraffes have long necks. Papafragou and Musolino (2003) then showed that much younger children (preschoolers) also accept such underinformative statements, but become more adult-like with training that emphasizes conventional descriptions. Soon afterwards, there was an explosion of acquisition research on monolingual – and later, bilingual – children's offline and online comprehension of scalar implicatures, as well as on whether and how methodological revisions could improve performance, shed light on the nature of the implicature, or access children's more fine-grained judgments (see, for example, Katsos & Bishop, 2011).

Well into the 2000s, then, acquisition researchers had increasingly taken up semantic and pragmatic issues in language acquisition, taking as a starting point theoretical linguistic formalism, with the goal of illuminating the process of acquisition and development, and at the same time obtaining empirical data that could bear on the aforementioned theoretical proposals. And a number of researchers who were initially trained as theoretical semanticists rather than developmentalists have turned to acquisition experiments to further probe the limits of meaning and test hypotheses that propose that some aspect of meaning depends on a particular component being present in the grammar or a particular ability to make certain kinds of inferences. Thus, whereas early investigations in acquisition and development were centered on syntactic phenomena, the field of language acquisition now has a more widespread focus, extending to investigations of semantic and pragmatic meaning.

This brings us to the current date and this volume on *Semantics in Acquisition*. We have invited established and leading researchers in language acquisition and development, along with scholars newer to the field, to contribute to this volume. Together, their work collectively represents both novel and classic investigations of semantics in acquisition that are firmly grounded in semantic theoretical traditions while advancing the field of acquisition forward into new territory.

We have divided the volume into five sections, which not only represent the research interests of these respective researchers, but also capture the major areas of investigation within this growing subfield in acquisition: the acquisition of word meaning and the structure of the lexicon, the acquisition of verb meaning and event semantics, the mapping between syntactic structure and semantic meaning and its role in the acquisition of word meaning, logical interpretations and abstract operations, and the relation between semantics and pragmatics in a speaker's meaning. The sections are sequenced in a way that roughly mirrors language development – a progression from the early acquisition of word meaning and the link between concepts and linguistic forms, to a more complex mapping between syntax and semantics and the interpretation of logical expressions, and finally to the relation between semantics and pragmatics in speaker meaning.

In each section, researchers highlight the basic learning and interpretational problems the young language learner faces – some universal across languages and some specific to a particular language, some universal across lexical and grammatical categories and some specific to particular expressions. These researchers also present experimental investigations, evidence from spoken language corpora, and/or data from previous studies that bear on this issue in order to make a proposal about children's developing linguistic capacity with respect to semantics, and in some cases its relation to syntax and/or pragmatics.

In **Section I (Lexical meaning)**, we begin with lexical semantics and the semantic features and domains present as children build their early lexicons. Eve Clark's chapter "Word meanings and semantic domains in acquisition" starts us off on the initial mapping between sound and meaning, and the gradual building up of semantic domains in the lexicon, which rely upon the link between words and concepts, and similar words to each other. Over the years, Clark has beautifully and skillfully drawn our attention to the semantic and pragmatic force of seemingly mundane daily interactions between children and their caregivers, documenting the development of these interactions over time, and describing in detail the semantic and pragmatic challenges inherent to acquiring words describing concepts such as space, relations, and properties. In this chapter, Clark highlights the conceptual sources for word meaning and the social and linguistic support provided to children about the meaning of words and the structure of a semantic domain by their caregivers in their interactions. Thus, Clark introduces us to the young language learner who is rapidly transforming into a young conversationalist.

One might be tempted to think that young children iron out the details of word meaning in the first two to three years of language acquisition, and from then on, it is semantic composition and more complex sentence structures that pose the challenge. However, this is not the case. Fast forward a number of years into language development, and children as old as five years of age are still struggling to pin down the meaning of certain words, even seemingly simple ones such as those mapping on to periods of time or temporal ordering. (We return to the struggles that five- and six-year-olds experience when trying to pin down the meanings of certain words in our last chapter on semantics and pragmatics.) In her chapter "The influence of linguistic temporal organization on children's understanding of temporal terms and concepts," Laura Wagner presents the results of two experiments in which children age five to eight were asked to order events on a timeline, given descriptions of the events and/or explicit linguistic temporal terminology. In doing so, she tests the hypothesis that if a particular semantic element is typologically prevalent across the world's languages, then it will be comparatively easy for children to learn (because concepts that humans are predisposed to learn will shape learning and will be grammaticized). While Wagner finds limited support for this hypothesis, the results highlight the role of a 'deictic center' in relative temporal ordering and lend potential support to a "thinking for speaking" position (Slobin, 1987, 1996), which holds that our grammar shapes the way that we perceive and think about events in the world and the way in which we speak about them.

Among the words and grammatical categories every child acquires are verbs. Investigations of how the child acquires verb meaning have played a central role in the field of language acquisition, because of the relative delay of the appearance of verbs in the lexicons of many languages relative to nouns (Gentner, 1978, 1982) and the tight relation between syntax and semantics in verb meaning due to the semantic roles linked to argument structure and event representation (Gleitman, 1990; Grimshaw, 1981; Pinker, 1994a). In their chapter "Semantic features of early verb vocabularies," Sabrina Horvath, Leslie Rescorla, and Sudha Arunachalam take this long-standing observation about the acquisition of verbs as a starting point for their comprehensive survey addressing *which* kinds of verbs are difficult to acquire initially, and why. Horvath et al. examined parental reports in a vocabulary checklist for toddlers acquiring Greek, Italian, Korean, and Portuguese, and coded verbs for a subset of key semantic dimensions tied to event representation: manner or result, durative or punctual events, and the number of event participants encoded in the verb's denotation. They find that both durativity and the number of event participants are predictors of order of acquisition.

We dig further into the link between event representation and verb meaning in **Section II (Event Semantics)**, in which the two chapters target one aspect of verbs in particular: telicity (event completion, from the Greek *telos* 'goal'). This aspect of verb meaning and events was most notably brought to light by Dowty (1979) and Vendler (1957). The two chapters in this section outline the learning problem and relevant experimental findings to date on the acquisition of telicity. Angeliek van Hout's chapter "On the acquisition of event culmination" presents the linguistic sources of telicity and perfectivity located in the lexical semantics of the verb, verb phrases, and aspectual morphology. She reviews in exceptional detail previous studies covering the acquisition of verb meaning as it relates to manner and end-states, and children's acceptance of telic sentences in events without culmination, incorporating a range of linguistic features and methodologies across a number of languages, including English, Dutch, German, and Spanish. Van Hout concludes by offering a proposal about why children consistently accept non-culminating interpretations of telic-perfective sentences.

Petra Schulz turns our focus specifically to German in her chapter, "Telicity in typical and impaired acquisition." Beginning with a concise and crisp presentation of the semantics of telicity in language in general, she then transitions to the encoding of telicity in German verb particles and the lexical semantics of verbs, highlighting the contribution of both semantics and pragmatics. How does a child learning to speak German acquire telicity in verb meaning? Schulz proposes an *Event Structural Bootstrapping* account in which the child focuses on a verb's event structure rather than on its core meaning or its argument selection, and demonstrate an *Endstate Orientation* in which they focus on expressions that encode an endstate. She presents empirical support for this position from previous studies of typically developing children in German, but contrasts this with findings from children who exhibit Specific Language Impairment (SLI) (Developmental Language Disorder). Atypically developing children appear to have an unstable *Endstate Orientation* (or lack an endstate orientation) and do not reliably understand that the endstate is entailed by telic verbs. The findings from German may be easily generalized to make predictions about the acquisition of telic verbs by typically and atypically developing children in other languages.

Given the complexity that arises in verb meaning – even with the simplest of verbs and event expressions – a question that naturally arises is what strategies children use to acquire verb meaning. **Section III (Syntactic Structure and Semantic Meaning)** presents four chapters that address the role that the relation between syntax/constructions and semantic meaning plays in this process. In their chapter "Not all subjects are agents: Transitivity and meaning in early language comprehension," Rose Scott, Yael Gertner, and Cynthia Fisher present two preferential-looking experiments with novel verbs to test the hypothesis that young language learners approximately two years of age take transitive syntax to indicate an asymmetric expression of the subject corresponding to the agent and the object corresponding to the patient (or proto-agent and proto-patient in the sense of Dowty (1991)) and not merely a causal interpretation. Their findings strongly suggest that young children at this age link transitive word order with semantic roles that extend beyond prototypical agents and patients, demonstrating a more abstract representation of meaning early in acquisition than was previously thought.

Ben Ambridge, Micah Goldwater, and Elena Lieven in their chapter "Analogical structure mapping and the formation of abstract constructions: A novel construction learning study" also capitalize on the value of incorporating novel language into an experimental design with a forced choice pointing paradigm to test the power of the link between syntax and semantics in verb learning. Instead of novel verbs, however, their experiment tests novel constructions in order to test Tomasello's (2003) proposal that children perform an analogical structure mapping across specific slot-and-frame patterns, and use this information to form abstract constructions. The premise of the experiment is that particular constructions may reliably express relations between event participants (e.g., the transitive expressing a relation between an agent and patient, or the ditransitive expression transfer of an object between possessors). Whether an overlap among constructions is sufficient for children to generalize to an abstract argument structure construction is another question, and one that they probe in their experiments with four- to six-year-olds and non-English OSV and VOS constructions. The null results raise a number of interesting questions about the viability of Tomasello's proposal and the support needed and strategies enlisted for analogical learning.

Aaron White, Valentine Hacquard, and Jeffrey Lidz move us beyond verbs denoting perceivable events of manner, motion, transfer of possession, and enabling or preventing an action to verbs that have more abstract meanings tied to belief and desire (so-called 'propositional attitude verbs'). In their chapter, "The labeling problem in syntactic bootstrapping: Main clause syntax in the acquisition of propositional attitude verbs," White et al. consider how a learner might use cues in a syntactic bootstrapping framework to identify 'clusters' of verbs that share either a belief or a desire component and to 'label' them as such. Like the previous chapters, they depend on a tight connection between syntax and semantics. However, they observe that while the syntactic frame in which verbs like *think*, *believe*, *know*, and *want* appear in English may display regularities with respect to ±tense, this correlation is not robust cross-linguistically. Their solution is to invoke abstract projection rules in which the semantic meaning maps onto not one syntactic feature, but a set of such features (a 'featural anchor'). They then go further to formalize their proposal in a probabilistic model, which depends on an incremental learner that observes pairings of such verbs and syntactic features and makes subsequent inferences about the verbs' semantic representation. This model is then applied to child-directed speech data.

Jill de Villiers' chapter "Perspectives on truth: The case of language and false belief reasoning" is also concerned with the information that an embedded clause conveys, and how children come to understand propositional attitude verbs. De Villiers has carefully highlighted how the language learner must eventually come to appreciate the force of linguistic devices such as pronouns, deixis, and evidentials in expressing another speaker's perspective. But at the same time, the child must understand that another speaker may share a different perspective than their own, and be able to take on that perspective. That is, the child must have a 'Theory of Mind'. Preschoolers' Theory of Mind abilities are often tested in 'false belief' tasks. De Villiers advances the thesis, supported by findings that she reviews in detail, that four-year-olds who pass such a task are able to do so because of their acquisition of sentential complement structures, which enables false belief reasoning. Specifically, they have come to recognize that the tensed complements that verbs such as *believe* and *think* take convey *realis*, or truth, assertion, and point of view.

Of course, the language acquisition process is not limited to the grammatical categories indicated above. Especially vexing are words that function as logical operators, or which undergo covert movement, or which are not present at all in the surface string. In **Section IV (Logical Interpretations)**, we turn to how children age four to six interpret sentences involving such expressions, and evaluate what the experimental findings presented in these chapters says about children's abstract semantic representations. The work presented in this section follows an approach in language acquisition shaped in many ways by one of the section authors, Crain, of demonstrating that preschoolers not only possess abstract representations involving logical operations, but that their interpretations are also often adult-like, presumably reflecting continuity in language development.

In their chapter, "The meaning of question words in statements in child Mandarin," Crain and Zhou present two experiments using the question/statement task, a derivative of the Truth Value Judgment Task (TVJT) originally pioneered by Crain (Crain & Thornton, 1998), and show that Mandarin-speaking preschoolers appreciate the existential force of a question word (*shenme* 'what') appearing in declarative contexts, but treat it as a question word in interrogative, information-seeking contexts (that is in the scope of different expressions).

In the second chapter in this section, "Overt, covert, and clandestine operations: Ambiguity and ellipsis in acquisition," Kristen Syrett reviews experimental evidence from a series of studies, some of which employ Crain's TVJT, testing young children's comprehension of ambiguous sentences involving overt movement (*wh-* questions) or covert movement (quantifier raising), and the interpretation of elided material in verb phrase ellipsis or comparatives. The three main sentence types under investigation involve the interaction of *wh*-phrases and universal quantifiers, antecedent-contained deletion, and comparative constructions involving pronominal reference in the elided material. The experimental findings not only clearly demonstrate that children as young as age four have access to abstract representations and operations at the syntax-semantics interface (even when they do not arrive at the intended adult-like interpretation), but also highlight the path of language development, provide a test of linguistic theory, and help us to better understand the endstate of the adult grammar.

As research on semantics in acquisition has burgeoned, so too has research on pragmatics and the relation between semantics and pragmatics in language development. The three chapters in the final section, **Section V (The Relation between Semantics and Pragmatics)** cover children's knowledge of different linguistic phenomena related to the calculation of semantic and pragmatic meaning. Lyn Tieu, Cory Bill, Jérémy Zehr, Jacopo Romoli, and Florian Schwarz start the section off with their chapter, "Developmental insights into gappy phenomena: Comparing presupposition, implicature, homogeneity, and vagueness." This chapter introduces a range of phenomena that have been debated as being handled either by the semantics or pragmatics, all of which involve so-called 'gappy' interpretations. These phenomena include presupposition failure when a presupposition is triggered under conditions that do not support it, the calculation of scalar implicatures when the literal meaning of the target expression is considered, statements that are not clearly true or false because they violate homogeneity, and instances of vagueness that involve borderline cases in fuzzy regions. Tieu et al. present a clear and concise review of the theoretical background on these topics, paired predictions for the path of acquisition and a review of a series of experimental studies directly comparing them and revealing their differences. The results of the studies not only provide insight into language development, but speak directly to the merits of the theoretical approaches that attempt to capture these phenomena.

In their chapter, "Four-year-old children compute scalar implicatures in absence of epistemic reasoning," David Barner, Lara Hochstein, Miriam Rubenson, and Alan Bale home in on children's calculation of 'ad hoc' scalar implicatures – those that are not based on strength by entailment relations – and compare this capacity to children's ability to make epistemic inferences at the same age. They show that the ability to compute ad hoc implicatures precedes the ability to compute ignorance implicatures. Here, too, in this chapter, Barner et al. offer a simple methodology that gets at a core understanding of the nature of children's ability to compute pragmatic inferences. Children were shown two dolls and provided with a statement with either a central *or* or *and*, and were asked to determine which of the puppets (the silly or straight faced one, or the seeing or blindfolded one) said it. The findings speak directly to neo-Gricean approaches to pragmatic meaning that implicate reasoning about another speaker's knowledge (or lack thereof).

Closing the volume is the chapter "The acquisition path of near-reflexivity" by Valentina Brunetto and Thomas Roeper, which tackles 'near reflexive' interpretations (Jackendoff, 1992; Lidz, 2001) – interpretations illustrated by a scenario in which Ringo sees a statue of himself in a museum, and the situation is described by saying that 'Ringo saw *himself*.' Brunetto and Roeper test children's interpretation of English 'self' and Italian *se* and show that English children allow for near-reflexivity with *himself* while Italian children tend not to do so with *se*. Brunetto and Roeper propose that near-reflexivity involves a mapping between sets of individuals (the real people and the statues), and that there is a pragmatic partition of the sets into subsets, giving rise to reciprocal and near-reflexive readings. They argue that at age four, children have difficulty computing multiple partitions, and therefore overaccept near-reflexives. Thus, an interface between syntax, semantics, and pragmatics explains the acquisition path of this linguistic phenomenon.

It is our hope that the collection of chapters in this volume on *Semantics in Acquisition* will not only showcase the range of linguistic phenomena that can be investigated in this subfield of language acquisition, but also highlight the tight connections between semantics and the other subfields of syntax and pragmatics but also between language, concepts, and cognition. In addition, the clear explication of theoretical underpinnings and the diversity of methodologies represented will help current and future researchers conduct further investigations of these topics. We are confident that the clarity of exposition in the chapters, the expertise and enthusiasm of the authors, and the inherent beauty and intrigue of the semantic topics investigated in the experimental research presented in this volume will draw future students to the study of semantics in language acquisition for years to come.

References

Barner, D., & Snedeker, J. (2008). Compositionality and statistics in adjective acquisition: 4-year-olds interpret *tall* and *short* based on the size distributions of novel noun referents. *Child Development*, 79, 594–608. https://doi.org/10.1111/j.1467-8624.2008.01145.x

Barwise, J., & Cooper, R. (1981). Generalized quantifiers and natural language. *Linguistics and Philosophy*, 4, 159–219. https://doi.org/10.1007/BF00350139

Bresnan, Joan. (1982). Control and complementation. In Joan Bresnan (Ed.), *The mental representation of grammatical relations* (pp. 282–390). Cambridge: MIT Press.

Bloom, L. (1970). *Language development: Form and function in emerging grammars*. Cambridge, MA: The MIT Press.

Bowerman, M. (1978). The acquisition of word meaning: An investigation into some current conflicts. In N. Waterson & C. Snow (Eds.), *The development of communication* (pp. 263–287). New York, NY: Wiley & Sons.

Bowerman, M. (1980). The structure and origin of semantic categories in the language learning child. In M. L. Foster & S. H. Brandes (Eds.), *Symbol as sense* (pp. 277–299). New York, NY: Academic Press.

Brooks, P. J., & Braine, M. D. S. (1996). What do children know about the universal quantifiers *all* and *each*? *Cognition*, 60, 235–268. https://doi.org/10.1016/0010-0277(96)00712-3

Brooks, P. J., Braine, M. D. S., Jia, G., & Dias, M. da Graca. (2001). Early representations for *all*, *each*, and their counterparts in Mandarin Chinese and Portuguese. In S. C. Levinson (Eds.), *Language acquisition and conceptual development* (pp. 316–339). Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511620669.013

Brown, R. (1973). *A first language*. Cambridge, MA: Harvard University Press. https://doi.org/10.4159/harvard.9780674732469

Chierchia, G., & McConnell-Ginet, S. (1993). *Meaning and grammar*. Cambridge, MA: The MIT Press.

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

Chomsky, N. (1981). *Lectures on government and binding*. Dordrecht: Foris.

Chomsky, Noam. (1980). *Rules and representations*. London: Basil Blackwell.

Clark, E. V. (1973). What's in a word? On the child's acquisition of semantics in his first language. In T. E. Moore (Ed.), *Cognitive development and the acquisition of language*. New York, NY: Academic Press.

Clark, E. V., & Amaral, P. (2010). Children build on pragmatic information in language acquisition. *Language and Linguistics Compass*, 4, 445–457. https://doi.org/10.1111/j.1749-818X.2010.00214.x
Crain, S., & McKee, C. (1986). The acquisition of structural restrictions on anaphora. In S. Berman, J-W. Choe, & J. McDonough (Eds.), *Proceedings of 16th annual meeting of the North East Linguistic Society* (pp. 94–111). Amherst, MA: GLSA.
Crain, S., Meroni, L., & Minai, U. (2004). If everybody knows, then every child knows. In J. van Kampen & S. Baauw (Eds.), *Proceedings of the Conference on Generative Approaches to Language Acquisition 3* (pp. 127–138). Utrecht: LOT.
Crain, S., & Thornton, R. (1998). *Investigations in Universal Grammar: A guide to research on the acquisition of syntax and semantics*. Cambridge, MA: The MIT Press.
Crain, S., Thornton, R., Boster, C., Conway, L., Lillo-Martin, D., & Woodams, E. (1996). Quantification without qualification. *Language Acquisition*, 5, 83–153. https://doi.org/10.1207/s15327817la0502_2
Dowty, D. R. (1979). *Word meaning and Montague grammar*. Dordrecht: Reidel. https://doi.org/10.1007/978-94-009-9473-7
Dowty, D. R. (1991). Thematic proto-roles and argument selection. *Language*, 67, 547–619. https://doi.org/10.1353/lan.1991.0021
Drozd, K. F. (2001). Children's weak interpretations of universally quantified questions. In M. Bowerman & S. C. Levinson (Eds.), *Language acquisition and conceptual development* (pp. 340–376). Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511620669.014
Fodor, J. A. (1975). *The language of thought*. Cambridge, MA: Harvard University Press.
Freeman, N. H., Sinha, C. G., & Stedmon, J. A. (1982). All the cars–which cars? From word meaning to discourse meaning to discourse analysis. In M. Beveridge (Ed.), *Children thinking through language* (pp. 52–74). London: Edward Arnold.
Gamut, L. T. F. (1991). *Logic, language, and meaning, Vol. I: Introduction to logic, Vol. II: Intensional logic and logical grammar*. Chicago, IL: University of Chicago Press.
Gentner, D. (1978). On relational meaning: The acquisition of verb meaning. *Child Development*, 49, 988–998. https://doi.org/10.2307/1128738
Gentner, D. (1982). Why nouns are learned before verbs: Linguistic relativity versus natural partitioning. In S. A. Kuczaj (Ed.), *Language development, Vol. 2: Language, thought, and culture* (pp. 301–334). Hillsdale, NJ: Lawrence Erlbaum Associates.
Gleitman, L. R. (1990). The structural sources of verb meanings. *Language Acquisition*, 1, 3–55. https://doi.org/10.1207/s15327817la0101_2
Gleitman, L. R., & Landau, B. (1994). *Acquisition of the lexicon*. Cambridge, MA: The MIT Press.
Goldberg, A. (1995). *Constructions: A construction grammar approach to argument structure. Cognitive theory of language and culture series*. Chicago, IL: University of Chicago Press.
Goro, T., & Akiba, S. (2004). The acquisition of acquisition of disjunction and positive polarity in Japanese. In V. Chand, A. Kelleher, A. J. Rodríguez, & B. Schmeiser (Eds.), *Proceedings of the 23rd West Coast Conference on Formal Linguistics (WCCFL)* (pp. 251–264). Summerville, MA: Cascadilla Press.
Grice, H. P. (1975). Logic and conversation. In P. Cole & J. Morgan (Eds.), *Syntax and semantics*, Vol.3. New York, NY: Academic Press. Reprinted in Grice (1989) (pp. 22–40).
Grice, H. P. (1989). *Studies in the way of words*. Cambridge, MA: Harvard University Press.
Grimshaw, J. (1981). Form, function, and the language acquisition device. In C. L. Baker & J. J. McCarthy (Eds.), *The logical problem of language acquisition* (pp. 183–210). Cambridge, MA: The MIT Press.

Gualmini, A. (2003). The ups and downs of child language. *Doctoral dissertation*, University of Maryland, College Park.

Gualmini, A., & Crain, S. (2005a). Operator conditioning. In A. Brugos, L. Micciulla, & C. E. Smith (Eds.), *BUCLD 28: Proceedings of the 28th annual Boston University Conference on Language Development* (pp. 232–243). Somerville, MA: Cascadilla Press.

Gualmini, A., & Crain, S. (2005b). The structure of children's linguistic knowledge. *Linguistic Inquiry*, 36, 463–474. https://doi.org/10.1162/0024389054396944

Gualmini, A., & Schwarz, B. (2009). Solving learnability problems in the acquisition of semantics. *Journal of Semantics*, 26, 185–215. https://doi.org/10.1093/jos/ffp002

Heim, I. (1982). The semantics of definite and indefinite NPs (Doctoral dissertation). University of Massachusetts. Amherst, MA: GLSA.

Heim, I., & Kratzer, A. (1998). *Semantics in generative grammar*. Malden, MA: Wiley-Blackwell.

Horn, L. R. (1972). On the semantic properties of the logical operators in English (Doctoral dissertation). UCLA, Los Angeles, CA. Distributed by the Indiana University Linguistics Club (1976). Bloomington, IN: Indiana University.

Horn, L. R. (1989). *A natural history of negation*. Chicago, IL: University of Chicago Press.

Horn, L. R. (2004). Implicature. In L. R. Horn & G. Ward (Eds.), *The handbook of pragmatics* (pp. 3–28). Oxford: Blackwell.

Inhelder, B., & Piaget, J. (1958). *The growth of logical thinking from childhood to adolescence*. New York, NY: Basic Books.

Inhelder, B., & Piaget, J. (1964). *The early growth of logic in the child: Classification and seriation*. London: Routledge and Kegan Paul.

Jackendoff, R. (1983). *Semantics and cognition*. Cambridge, MA: The MIT Press.

Jackendoff, R. (1990). *Semantic structures*. Cambridge, MA: The MIT Press.

Jackendoff, R. (1992). Mme. Tussaud meets the binding theory. *Natural Language and Linguistic Theory*, 10, 1–31. https://doi.org/10.1007/BF00135357

Kamp, H., & Partee, B. (1995). Prototype theory and compositionality. *Cognition*, 57, 129–191. https://doi.org/10.1016/0010-0277(94)00659-9

Kartunnen, L. (1974). Presuppositions of compound sentences. *Linguistic Inquiry*, 4, 169–193.

Karttunen, L., & Peters, S. (1979). Conventional implicature. *Syntax and Semantics, Vol. 11: Presupposition* (pp. 1–56). New York, NY: Academic Press.

Katsos, N., & Bishop, D. (2011). Pragmatic tolerance: Implications for the acquisition of informativeness and implicature. *Cognition*, 120, 67–81). https://doi.org/10.1016/j.cognition.2011.02.015

Ladusaw, W. (1979). Polarity sensitivity as inherent scope relations (Doctoral dissertation). University of Texas, Austin. Republished in (1980) in the series Outstanding dissertations in linguistics. New York, NY: Garland.

Landau, B., & Gleitman, L. R. (1985). *Language and experience: Evidence from the blind child*. Cambridge, MA: Harvard University Press.

Lewis, D. (1979). Scorekeeping in a language game. *Journal of Philosophical Logic*, 8, 339–359. https://doi.org/10.1007/BF00258436

Locke, J. (1690/1894). *An essay concerning human understanding*. Ed. by A. Campbell Fraser, 2 Vols. Oxford: Clarendon Press.

Lidz, J. (2001). Condition R. *Linguistic Inquiry*, 32, 123–140. https://doi.org/10.1162/002438901554603

Lidz, J., & Musolino, J. (2002). Children's command of quantification. *Cognition*, 84, 113–154. https://doi.org/10.1016/S0010-0277(02)00013-6

May, Robert. 1977. The grammar of quantification. *Doctoral dissertation*, MIT. Cambridge: MA.

May, R. (1985). *Logical form*. Cambridge, MA: The MIT Press.

McDaniel, D., McKee, C., & Smith Cairns, H. (1996). *Methods for assessing children's syntax*. Cambridge, MA: The MIT Press.

Montague, R. (1973). The proper treatment of quantification in ordinary English. In K. J. J. Hintikka, J. M. E. Moravcsik, & P. Suppes (Eds.), *Approaches to natural language* (pp. 221–242). Dordrecht: Reidel. https://doi.org/10.1007/978-94-010-2506-5_10

Montague, R. (1974). *Formal philosophy: Selected papers of Richard Montague*. Ed. by R. H. Thomason. New Haven, CT: Yale University Press.

Musolino, J. (1998). Universal Grammar and the acquisition of semantic knowledge: An experimental investigation into the acquisition of quantifier-negation interaction in English (Unpublished doctoral dissertation). University of Maryland, College Park, MD.

Noveck, I. (2001). When children are more logical than adults: Experimental investigations of scalar implicature. *Cognition*, 78, 165–188. https://doi.org/10.1016/S0010-0277(00)00114-1

Papafragou, A., & Musolino, J. (2003). Scalar implicatures: Experiments at the semantics-pragmatics interface. *Cognition*, 86, 253–282. https://doi.org/10.1016/S0010-0277(02)00179-8

Partee, B. (1995). Lexical semantics and compositionality. In L. Gleitman, M. Liberman, & D. Osherson (Eds.), *An invitation to cognitive science, Part I: Language* (pp. 311–360). Cambridge, MA: The MIT Press.

Partee, B. (2006). Montague, Richard (1930–1971). In K. Brown (Ed.), *Encyclopedia of language and linguistics* (pp. 255–257). Oxford: Elsevier. https://doi.org/10.1016/B0-08-044854-2/02760-7

Philip, W. (1995). Event quantification in the acquisition of universal quantification (Unpublished doctoral dissertation). University of Massachusetts, Amherst.

Pinker, S. (1989). *Learnability and cognition: The acquisition of argument structure*. Cambridge, MA: The MIT Press.

Pinker, S. (1994a). How could a child use verb syntax to learn verb semantics? *Lingua*, 92, 377–410. https://doi.org/10.1016/0024-3841(94)90347-6

Pinker, S. (1994b). *The language instinct: How the mind creates language*. New York, NY: Harper Perennial Modern Classics.

Reinhart, T. (1983). Coreference and bound anaphora: A restatement of the anaphora questions. *Linguistics and Philosophy*, 6, 47–88. https://doi.org/10.1007/BF00868090

Rizzi, L. (1986). Null objects in Italian and the theory of pro. *Linguistic Inquiry*, 17, 501–557.

Ross, J. R. (1967). Constraints on variables in syntax (Unpublished doctoral dissertation). MIT.

Slobin, D. I. (1987). Thinking for speaking. In *The proceedings of the thirteenth annual Meeting of the Berkeley Linguistics Society* (pp. 435–445). Available at: <https://journals.linguisticsociety.org/proceedings/index.php/BLS/article/view/1826/1598>

Slobin, D. I. (1996). From "thought and language" to "thinking for speaking.". In J. J. Gumperz & S. C. Levinson (Eds.), *Rethinking linguistic relativity* (pp. 70–96). Cambridge: Cambridge University Press.

Smith, C. L. (1980). Quantifiers and question answering in young children. *Journal of Experimental Child Psychology*, 30, 191–205. https://doi.org/10.1016/0022-0965(80)90057-0

Stalnaker, R. (1974). Pragmatic presuppositions. In R. Stalnaker (Ed.), *Context and content* (pp. 47–62). Oxford: Oxford University Press.

Syrett, K. (2007). Learning about the structure of scales: Adverbial modification and the acquisition of the semantics of gradable adjectives (Unpublished doctoral dissertation). Northwestern University.

Syrett, K., Kennedy, C., & Lidz, J. (2010). Meaning and context in children's understanding of gradable adjectives. *Journal of Semantics*, 27, 1–35. https://doi.org/10.1093/jos/ffp011
Tomasello, M. (2003). *Constructing a language: A usage-based theory of language acquisition*. Cambridge, MA: Harvard University Press.
Vendler, Z. (1957). Verbs and times. *The Philosophical Review*, 66, 143–160. https://doi.org/10.2307/2182371
Wilson, D., & Sperber, D. (2004). Relevance theory. In L. R. Horn & G. Ward (Eds.), *The handbook of pragmatics* (pp. 607–632). Oxford: Blackwell.
Williams, Edwin. (1980). Predication. *Linguistic Inquiry*, 11, 203–238.

SECTION I

Lexical meaning

CHAPTER 2

Word meanings and semantic domains in acquisition

Eve V. Clark
Stanford University

As children accumulate words, they build up semantic domains. In doing this, they start to link the meanings of words, depending on how they are related to each other. They rely on conceptual representations of objects and events, and on how adults talk about objects and events. Adults typically provide information along with new-word offers: facts about class membership, parts and properties, motion, sound, and function provide a basis for semantic relations. Semantic domains built up early include many general domains as well as some domains of intense interest (e.g., dinosaurs or cars), also elaborated with parental support. As children learn more words, they structure each domain and link new terms to ones they already know.

Keywords: word meaning, semantic domain, semantic relations, sources, expertise

1. Introduction

Adults talk with young children about the here-and-now, about what is happening. At the table, they talk about eating and drinking, plates and cups, spoons and forks, and specific kinds of food and drink depending on the meal. The terms they use for food, plates, and utensils constitute a semantic domain. Some semantic domains are loosely structured sets of nouns and verbs for talking about a particular activity, such as eating and drinking. Others contain nouns that are hierarchically structured, with from two to seven levels, as in various plant and animal term hierarchies. Semantic domains also contain terms for activities associated with objects and their roles (see Lyons, 1977; Lehrer, 1974; Lehrer & Kittay, 1992). This chapter focusses on how, as children acquire word meanings, they organize them into semantic domains.

https://doi.org/10.1075/tilar.24.02cla

2. Acquiring word meanings

From the end of their first year on, children hear many new words. Some are flagged as new when adults introduce them explicitly in a frame like 'This is a….' or 'That's called a ….'. They encounter many new words in passing, and there they must assign some meaning in context. Parents introduce their children to new words when they and their children share joint attention on some physically present object or activity (Moore & Dunham, 1995).

When adults offer a new word, children attend to it and frequently repeat it in their next turn. This repetition ratifies the adult offer, showing that children have noticed the new word and are acknowledging it, and it allows children to try out the new word themselves (e.g., Clark, 2007, 2010). Notice that ratification places the new word in common ground: both adult and child know that the new word has been applied in that context on that occasion.

When adults make explicit offers, they don't only offer a label. They often add information about the referent of the word too (Clark & Wong, 2002; Clark & Estigarribia, 2011). With words for objects, they offer information about class membership or inclusion, about parts, properties, characteristic noises, movements, and functions.

(1) Parent holding out measuring spoon to Child, 1;6: That's a spoon.
(child touches it)
Do you know what you do with a spoon? What do you do with a spoon?
You eat with a spoon.

(2) Parent showing toy crocodile to Child, 1;9: You wanna see? Oh look, look!
(holds it out to child) Who's that? Who is that? That's an alligator. Say alligator.
Child: /a/
Parent: Yeah. Alligator. What color is the alligator? … What color is he? Is he green?
…. Look at his teeth. You see the teeth? (parent close-points to teeth)

(3) Parent puts toy stingray in front of Child, 2;10: What's this? (child looks)
It goes in the sea.
Child (hesitant): *Yeah…*
Parent: You call this a sting-ray.

With words for actions, adults often supply the kind of object associated with the action, or ask a general question with *do* ('what's the doggie *doing*?'), then repeat the question offering a new verb ('he's *sniffing* the shoe'). With adjectives, they contrast the new adjective with a familiar one ('this ball is orange, not red'). With prepositions, they tend to add heavy contrastive stress on the new preposition ('he fell down *from* a tree'). Overall, adults are quite consistent in how they offer new words, using distinct presentation frames for each part of speech (Clark, 2010).

2.1 The first step: Fast mapping

Adults work to make sure children are attending before they offer a new word. When this happens, children must assign the word some preliminary meaning. What they select varies with context and with any words they already know. For very young children, what is in joint attention may fill their field of vision. As a result, if adults seize the right moment to offer a new word, what the child is currently attending to is the best, and perhaps the only, candidate-referent and so provides a first meaning (Tomasello, 1995; Estigarribia & Clark, 2007; Yurovsky, Smith, & Yu, 2013). Adults and children tend to focus on whole objects here.

This first step in assigning meaning has been called 'fast mapping' (Carey & Bartlett, 1978; Heibeck & Markman, 1987). Once children have assigned a preliminary meaning, they can supplement it, adding to or changing the meaning in light of further adult uses of the same word. Once children try out new words themselves, they can also use adult feedback to supplement their initial meanings (Chouinard & Clark, 2003). For example, when children apply words inappropriately, adults give feedback in which they check up on what children intended, or tacitly accept the child's word choice but offer an alternative in their next utterance:

(4) Checking up on what the child meant:
D (2;3.25) out of the blue, at breakfast, remembering an episode from when living in NL four months earlier
D: *i climb up* a *ladder with 'vrouw groot*.
Mother: You climb<u>ed</u> up a ladder with <u>Me</u>vrouw Groot?
D (explaining): *we go on a slide*.

(5) Following up with an alternative word:
D (2;0.5) told that a friend in the house was getting ready to leave
D: *Christiane packing…/ə./../ə/…/ə/ pack-case*.
Father: yes, she's packing a <u>suitcase</u>.
D: *i go see*.

How much does context influence fast mapping? Most experimental studies of word learning have focussed on object labels, namely nouns, assuming that these are easier to learn because there is a concrete referent visible. But when researchers vary the objects being acted on with the same action, or the actions being performed by the same object, children map novel words differently. With repetitions of the same action on different objects, children treat the novel word as a verb for that action, but when the same object carries out several different actions, they treat the novel word as a noun for that object (Tomasello & Barton, 1994; Tomasello & Kruger, 1992; see also Haryu, Imai, & Okada, 2011). In short, children attend closely to what is going on and what the speaker is attending to as they assign an initial meaning to a new word.

How many exposures do children need in order to add a new word to their vocabulary? Researchers have attempted to answer this question experimentally by looking at novel words in word learning tasks. They typically teach new words for nonsense objects or drawings, then test children's comprehension as soon as the learning trials are completed (e.g., Markman & Wachtel, 1988). At this point, preschool children usually do well. But when researchers follow up such a task 24 hours later, they find little retention of words learnt the day before (Bion et al., 2013; Horst & Samuelson, 2008). Compared to real words, of course, nonsense words have no history in ongoing interactions so children receive no exposure to adult usage. In spontaneous exchanges, children attend to explicit offers of new words and, about half the time, repeat them immediately in their next turn. In doing this, they both ratify the adult offers that have been made, and try out the new words themselves (Clark, 2007, 2010). But just when to count words as 'acquired' is hard to establish with production data alone.

3. Conceptual and social sources of information

In assigning initial meanings, children rely on two major sources: what they know from their perceptual and conceptual knowledge about the object or activity in question (Clark, 2004), and what they know from social interaction, in terms of cooperation, conventionality, and contrast (Grice, 1989; Clark, 1987, 1990).

3.1 Conceptual sources for word meanings

Children rely on conceptual categories and organization acquired earlier when first mapping meanings to words. They draw on information already established about objects, actions on objects, and relations between objects in space, for example, when asked to group like-objects together or carry out simple instructions (e.g., Clark, 1973b, 1980). Many initial meaning assignments depend on conceptual categories children are already familiar with (Clark, 2004).

In identifying members of a category with a label, one-year-olds consistently rely on shape (e.g., Baldwin, 1989; Booth & Waxman, 2002; Clark, 1973a; Landau, Smith & Jones, 1988; Gershkoff-Stowe & Smith, 2004). In responding to instructions to place objects, they rely on their conceptual organization of space. Whether told, 'Put the mouse in the box' or 'Put the mouse on the box', one-and-a-half year olds always place the mouse 'in' the box. But when shown a small table, and asked to put the mouse 'on' or 'under' it, they consistently place the mouse on top of the table. In short, containers hold things, and surfaces support things. But once children begin to map the meanings of *in, on*, and *under*, they replace their former reliance on purely conceptually-based strategies, and make use instead of the meaning of support with *on*, containment with *in*, and placement below with *under* (Clark, 1973b).

3.2 Social sources for word meanings

In assigning word meanings, children also rely heavily on social sources of information. They keep track of who used what word where and when. In this they follow general pragmatic principles that also guide adult uses of language. These principles include cooperation in conversation, conventionality, and contrast.

3.2.1 *The cooperative principle*

This principle of conversation captures the fact that when speakers say something, they communicate in a cooperative fashion, generally by being clear, truthful, relevant, and by providing enough (but not too much) information (Grice, 1989). In assuming that speakers intend to be cooperative, their addressees can – given what they already know, and what they expect to hear – make appropriate inferences about what speakers intend. The Cooperative Principle depends on speakers making use of the conventions of the language and of the contrasting meanings carried by the terms and constructions of the language.

3.2.2 *Conventionality and contrast*

For any communication system to work effectively, users must observe the conventions agreed on by the community. For certain meanings, speakers assume that there is a conventional form that should be used in that language, for example, the English word *dog* denotes dogs. And speakers assume that any difference in form – *dog* versus *cat* – signals a difference in meaning. Notice that the reverse doesn't hold: a single form can carry several meanings. When speakers don't use the expected conventional form, but instead produce another expression or coin a new term, their addressees must compute that meaning, in contrast to what they had expected. Contrast works hand-in-hand with conventionality, and speakers rely on both in making use of the cooperative principle in conversation (Clark, 1987, 1990).

In accumulating and organizing vocabulary, therefore, children must attend to (a) how a new word is used by the adult (the category it refers to); (b) any familiar words it contrasts with, for example, words in the same domain; and (c) the relation between the new word and other words already known.

4. Accumulating words

As children add more words, they accumulate clusters of words that belong in the same semantic domain. In Box A, I have listed all the terms produced for animals – domestic, farm, zoo, wild – as well as for birds and reptiles, for a total of 61 different terms, from one child, D, between the ages of 1;0.20 and 3;0. These terms were mostly acquired and then produced as singletons, in relevant contexts. In some contexts, the child would produce two or more terms from this domain.

The terms shown emerged in the order given, presumably reflecting the child's exposure to each category, in real life and in books. From 1;0.20 to 1;11.30, most of the terms he added were for mammals (12 terms), plus two terms for birds (*chicken* and *duck*), one for an insect (*ladybug*), and five for reptiles and sea creatures. In addition, he produced three potentially superordinate terms: *animal, bird*, and *fish*.

Box A. First productions of animal terms added between age 1;0 and 3;0 by one child

(a) 1;0.20–1;6.0:
1;0.29 doggie, 1;1.15 dog, 1;1.26 bear, 1;2.0 bird, 1;3.0 duck, 1;3.3 mouse, 1;3.23 cat, 1;4.0 chicken, 1;4.28 horse, 1;5.6 turtle

(b) 1;6.1–1;11.30
1;6.4 fish, 1;6.13 cow, 1;6.28 rabbit, 1;7.2 goose, 1;7.19 lion, 1;7.20 *animal*, 1;8.7 frog, 1;8.8 alligator, snake, 1;8.22 crab, 1;10.19 ladybug, 1;10.24 seal, 1;11.16 gorilla, 1;11.30 goat

(c) 2;0–3;0
2;0.8 puppy-dog <toy>, 2;0.11 hippo, 2;1.7 camel, 2;2.24 baby-rabbit, 2;3.21 stork, ostrich, 2;4.6 robin, 2;4.10 baboon, 2;4.16 tiger, 2;4.17 sparrow, 2;4.19 flamingo, 2;4.20 dove, 2;5.1 trout, 2;5.4 flounder, 2;5.6 spider, grass-hopper, 2;5.10 grouse, 2;5.14 woodpecker, 2;5.21 mummy-bunny <toy>, monkey, 2;7.2 fly, bee, 2;7.15 wolf, raccoon, 2;7.27 sheep, 2;8.6 owl, 2;8.8 armadillo, fox, 2;8.14 butterfly, 2;8.17 beaver, 2;9.21 lizard-animal <live>, frog-animal <live>, 2;11.1 cattle, baby-cattle, daddy-cattle, 2;11.3 bronco-horse, 2;11.10 bucking-horse
[Clark, diary data]

D's uses of early bird terms initially made a three-way distinction based on air (*bird*: generally flying), water (*duck*: swimming) and ground (*chicken*: walking), but uses of these terms moved towards the conventional adult meanings as he added more terms for birds from age 2;3 on, e.g., *stork, ostrich, robin, sparrow, dove, flamingo*, then *grouse* and *woodpecker*.

In some parts of the domain shown in Box A, the child spontaneously set up subcategories, using novel compounds as in *bronco-horse* and *bucking-horse*. He also began to pay attention to family relations (e.g., *baby-rabbit, mummy-bunny; cattle, baby-cattle*, *daddy-cattle*). It took considerably longer, though, before he added any terms for habitat, sound, or locomotion specific to particular kinds (Clark, 1995).

From age 3 years on, he organized the words shown here into farm versus zoo animals, birds, reptiles, and so on, as he learned still more words in this domain.

Organization and re-organization within domains is consistent with findings for older children for the domain of mammals. Johnson and her colleagues (1992) looked at how children aged 7 and 10 judged similarity compared to adults, for 25 mammals. The child and adult groups were much the same in their understanding of similarities and differences, with 10-year-olds making the same taxonomic judgements as adults. But 7-year-olds still had some way to go. Although they produced appropriate names over 80% of the time, they had yet to distinguish primates as a sub-category of mammals.

The domain of animals is just one domain where children start acquiring terms early on. Other domains set up quite early include food, mealtimes, and associated activities; clothing; household objects; activities linked to different rooms; toys; cars; birds and flowers, plants, and trees. In each domain, children necessarily start small, with just one or two words. Building up a domain depends on experience with the 'contents' and exposure to adult talk about the domain (Roy, Frank & Roy, 2012; Hills 2013).

5. Setting up semantic domains

On what basis do children set up a semantic domain? As they take up new words, adults provide them with information about the referents of those words (e.g., Callanan, 1990; Clark & Wong, 2002; Clark & Estigarribia, 2011). They give information about class-membership (*A seal is a mammal; An owl is a bird*); about related subkinds (*This is a Siamese cat: see his dark ears and paws? And this one is a Persian with all that long fur*). They draw children's attention to parts and properties (*This is the handle; Look at the rabbit's tail*). They talk about habitats (*Dolphins live in the sea; Rabbits make burrows to live in*). They talk about the young of different species (*A baby horse is called a foal; Cats have kittens*), about the food each kind eats, the noise it makes, the way it moves. With artefacts, they talk about functions (*This ball's for throwing, The sieve is for letting water through and catching the rice*). And adults often highlight the relevant details with gestures (Clark & Estigarribia, 2011).

Such added information informs children about what the members of a domain have in common (e.g., animals eat, breathe, and produce young) and since the information is usually presented in generic form, children infer that it applies to all members of a domain (Cimpian & Markman, 2008; Gelman, Goetz, & Sarnecka, & Flukes, 2008). The information adults provide also focusses on how categories and category members differ, for example, on parts that differentiate subcategories (see Tversky & Hemenway, 1984; Mervis & Rosch, 1981). Finally, much of this information provides children with the relations that hold among the meanings of familiar and new words.

5.1 Some semantic domains

Researchers have studied children's acquisition and organization of a number of semantic domains, including spatial terms like prepositions in English (e.g., Clark, 1973b; Casasola, Bhagwat, & Burke, 2009; Choi et al., 1999; Grimm, 1975; Murphy & Jones, 2008; Rice, 2003; Tomasello, 1987); terms for spatial dimensions (e.g., Clark, 1972; Donaldson & Wales, 1970; Ebeling & Gelman, 1988; Murphy & Jones, 2008); terms for plants (Dougherty, 1979); terms for kinship (e.g., Haviland & Clark, 1974); terms for containers (Ameel, Malt, & Storms, 2008), and terms for color (e.g., Au & LaFramboise, 1990; Mervis, Bertrand, & Pani, 1995; O'Hanlon & Roberson, 2006; Sandhofer & Smith, 1999; Shatz, et al., 1996). They have also studied the acquisition of verbs like *give/take* and *buy/sell* (e.g., Gentner, 1975); *cutting, hitting*, and *breaking* (e.g., Majid, Boster, & Bowerman, 2008; Majid et al., 2007; Pye, Loeb, & Pao, 1996); *carrying* (e.g., Saji et al., 2011); verbs for motion in space (e.g., Allen et al., 2007; Malt et al., 2016; Naigles & Terrazas, 1998; Özçalışkan & Slobin, 1999; Papafragou, 2010); and categorization of manner and paths in events related to verb learning (e.g., Pruden et al., 2013; Pulverman et al., 2008).

In acquiring terms in these domains, children draw on conceptual information about the domain, and on social information – the language the people around them generally use for talking about each domain. These two sources, nonlinguistic (conceptual) and linguistic (social), inform both children's construction of categories and their acquisition of terms for talking about them.

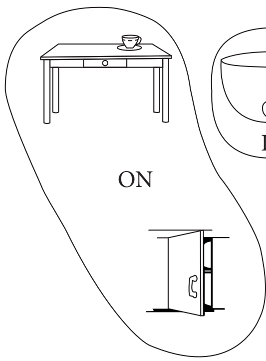
5.1.1 *Placement in space*

Children attend closely to where objects are placed and how they move in space. They attend to topological relations like containment and support, and, in many languages, learn how to talk about these early on. But at first, they attend only to certain perceptual or conceptual dimensions such that they always place small objects in containers and on supporting surfaces (Clark, 1973b). Where a language encodes such factors as tight versus loose fit for containment (compare a tape in a cassette case with an apple in a bowl) as in Korean, children also attend to how an object is contained (Choi, 1991; Choi et al., 1999; see also Casasola et al., 2009). And having learnt how adults talk about containment and support, they take into account both properties of objects and of where they are to go, along with the relevant linguistic expressions for more complex talk about space and spatial relations (Clark, 1980; Choi & Bowerman, 1991; Landau et al., 2017; Rice, 2003).

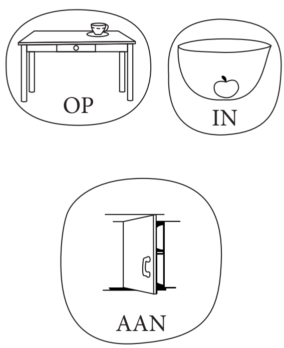
Just how each language distinguishes containment from support varies considerably, even for such closely related languages as English and Dutch. While English relies on IN for containment and ON for support, Dutch divides the same domain three ways, with IN for containment, OP for support on a horizontal surface, and AAN for support through attachment, as shown in Box B.

Box B. Containment and Support in English and Dutch

(a) English IN and ON



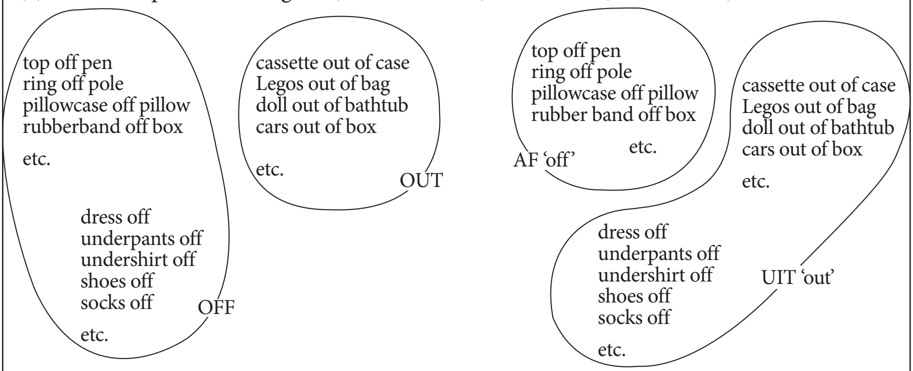
(b) Dutch IN, OP, and AAN



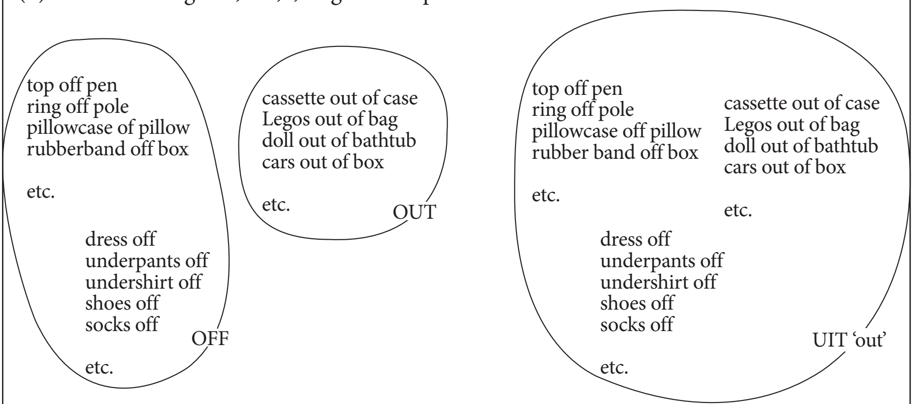
When talking about removal from a container or a supporting surface, English and Dutch simply divide the domain differently, as shown in the top half of Box C. But children acquiring these two languages appear to find it easier to grasp the dividing lines in English than in Dutch where they initially opt for just one term, UIT, as shown in the lower half of Box C (Bowerman, 1996; Gullberg, 2011).

Box C. Removal or separation in English and Dutch

(a) in adult speakers of English (OFF and OUT) and Dutch (AF and UIT)



(b) in children aged 2;0–2;5, English compared to Dutch



5.1.2 *Joining and separating*

Containment sometimes includes various kinds of combining or joining (think of pop-beads, nested bowls) as well as the reverse: separation of two pop-beads, removal of one bowl from the stack, and so on. In English, such relations in space are generally encoded with *put in, put together, join* versus *take out, take apart, separate*. In Korean, speakers use specific verbs for joining or inserting depending on tight or loose fit, and each verb has a counterpart for reversing that action (Choi & Bowerman, 1991; Narasimhan et al., 2012). Children typically begin with just one or two expressions: English-speaking children start out with *in* or *put in* and reverse *in*-placements with *back* or *out*, later replaced by *take out*. (*Put* and *take* are flexible locative verbs that occur with a variety of locative particles in English, e.g., *in, on, beside, above, up, along*, *inside*.) This pattern of starting out with just one of the available expressions for joining (e.g., *in, shut*) and separating (e.g., *out, open*) appears typical as children start to talk about such events (see Bowerman, 1996; Choi et al., 1999).

5.1.3 *Holding* and *carrying*

One can hold or carry objects in one hand or in both, on a shoulder or on one's back, on one's head, in one's arms or under an arm. One can support the object being held, or let it dangle from a strap or handle. In Chinese, speakers have to learn which verb to use for each type of holding or carrying, from among a number of verbs for these everyday activities. Saji and colleagues (2011) looked at how children acquired some 13 distinct verbs by looking at how they described video clips of people either holding or carrying a familiar object in some way. Saji et al. (2011) also collected comprehension data, asking children and adults whether an action shown in a video clip was an instance of X, where X was one of the target verbs.

Children added verbs to this domain very slowly: they produced the same number of verb types at age 3, 5, and 7 years, but the Pearson correlation between child and adult usage increased with age, from a low of .17 for 3's to .53 for 5's, and .58 for 7 year olds. The changes in this correlation, though, were not attributable to shifts in adult usage: adults maintained the same patterns of use as in talking to other adults. One obstacle for children with these verbs is identifying the dimensions on which the verb meanings differ, the most prominent being *manner* of holding or carrying. Children instead focussed first on the objects being carried. As in other domains, they started out with a single frequent verb, generally used with too-broad an extension, and only later narrowed its meaning down. Acquisition of this domain takes time because children can't converge on adult usage until they work out how each verb contrasts in meaning with its neighbours. In short, children have to find the boundaries between the words for holding or carrying events before they can structure this domain in an adult fashion.

Two factors seemed to be important here. First, the frequency with which children hear a verb contributes to its early uptake (but not necessarily to adult-like usage). If the verb-boundary is clear in adult usage, with no overlap with other verbs in the domain, even if it covers a range of actions, children converge with adult usage earlier. In English, children need only work out the meanings of two verbs, *hold* and *carry*, where both cover a variety of actions, but differ on whether the focus is on motion (*carry*) or not (*hold*). In Chinese, children must attend to the manner of holding and carrying, and to the type of object involved. Consequently, acquisition is more protracted.

5.1.4 *Motion, manner, and path*

When people talk about motion in space, they are constrained by the patterns available in their language. In some, verbs of motion include information about manner-of-motion, as in English *stroll, jog*, and *sprint* (where manner information encodes increasing speed). Particles provide direction of motion. But in many languages, motion verbs contain information about direction rather than manner, as in Spanish *entrar* 'go in' or *bajar* 'go down', so speakers have to add information about manner in some other way (Talmy, 1985; Slobin, 2004, 2006; Beavers et al., 2010). Children start out with simple verbs of motion like *walk* and *run* (Malt et al., 2016), with parents providing the relevant verbs in talk about motion events (Hohenstein, 2013). Only later do children come to understand and add verbs for motion + direction or motion + manner (Allen et al., 2007; Naigles & Terrazas, 1998; Özçalışkan & Slobin, 1999; Papafragou, 2010; Pruden et al., 2013; Pulverman et al., 2008).

Motion and direction are closely linked to notions of source and goal. Children treat goals of motion as primary, and more salient than sources: when moving, one looks to the goal (not the source; Clark & Carpenter, 1989a, 1989b; Lakusta et al., 2007; Papafragou, 2010). The salience of goals is also reflected in the attention paid to resultant states, the end-points or goals of changes in state (Cheung & Clark, 2006). Children combine talk about motion with goals but rarely mention sources. This asymmetry also holds for adults (Lakusta & Landau, 2012).

6. Structure within semantic domains

Semantic domains contain nouns, verbs, prepositions, and adjectives, and while some domains may consist predominantly of a single word class (e.g., dimensional adjectives in English), many contain a mix of word classes that capture information relevant to the semantic relations in that domain. Among the commonest relations are **inclusion** or **membership** in a superordinate category, as in 'A seal is a mammal', where *seal* is a hyponym or subordinate of the superordinate term *mammal*. Another relation is that of **opposite**, as in *alive* vs. *dead*, or *up* versus *down*. **Parts** are related to wholes, as in '<u>The wheel</u> is part of the truck'. 'Rabbits have <u>ears</u>', or '<u>The pump</u> belongs to that bicycle'. Other terms in a domain pick out **properties** like size or texture (*big*, *furry*, *smooth*, etc.), characteristic actions and noises ('Horses can gallop', 'Sheep go baa'), and **functions** ('Sieves are for straining out water'), or information about **habitat** and **ontogenesis**.

By way of illustration, take the domain of farm animals. Within this domain, while children often acquire the term *horse* early on, there are many related terms they may eventually acquire. These might include *mare* and *stallion* (also *gelding*), *foal, colt*, and *filly*; for equine motion, the noun *gait* and the verbs *walk, trot, canter,* *gallop*; also *shy, rear, buck*; for characteristic horse sounds, *neigh, whinny, whicker*; for characteristic colors, *black, white, brown, palomino, roan, piebald, skewbald*; and for feed, *oats, hay, grass*. Terms for harness might include *bridle, bit, reins, saddle, crupper*, and *nightingale*; and terms for habitat, *farm, stable, stall, paddock*, and *range* or *prairie*. The same accretion of related words applies to other animal terms as well. But in a domain like this, children start out with just one or two terms (as shown in Box A), and then add more over time.

As children build up different semantic domains, they come to realize that certain terms overlap in meaning, so the word *animal*, for example, includes a number of more specific terms ranging from *lion* to *goat* to *squirrel*. And while they may at first treat *animal* as a collection-term (Macnamara, 1982), children soon grasp the inclusion relation involved and begin to treat *animal* as a superordinate term instead. Once they realize that a lion is a kind of animal, and that donkeys, horses, cows, and sheep are all animals too, they can generalize about each new animal term: A *lemur* is also an animal, as is an *okapi*. Inclusion appears to be one of the earliest semantic relations children grasp. Acceptance of various kinds of overlap in meaning is also basic to the realization that a single instance of a category may be talked about with more than one term: the family dog can be referred to as *the dog* or *the animal*. But it can also be referred to as *Toby*, *that dog*, *our retriever*, *the pest*, *the scavenger*, and more (see Ravn, 1988; Clark, 1997).

Children demonstrate clear awareness of the hierarchical relation between terms like *animal* and *dog*, or *whale* and *beluga*, for instance, from as young as 2 years of age (Callanan & Sabbagh, 2004; Clark & Svaib, 1997). Moreover, when taught two novel words, *ruk* and *dib*, for two unfamiliar categories of objects (whisks and honey-sticks, say), and given just one utterance linking the two, 'a *dib* is a kind of *ruk*', during learning trials, young 2-year-olds reliably pick out both ruks and dibs when asked for 'all the *ruks*', but only dibs when asked for 'all the *dibs*' (Clark & Grossman, 1998; see also Haryu & Imai 2002; Mervis, Golinkoff, & Bertrand, 1994; Waxman & Senghas, 1992).

Children also pay attention to the relation part-of (or meronymy) that links a part to the relevant whole. They attend to offers of part-terms in talk about unfamiliar objects (Clark & Wong, 2002; Clark & Estigarribia, 2011; Masur, 1997), and learn some body-part terms quite early (Andersen, 1978; Sinha & Jensen de López, 2000; also Bergelson & Swingley, 2012). Parents typically label whole objects first, and only then point to and label parts (e.g., Masur, 1997). This whole-part juxtaposition appears to help children infer that the unfamiliar term is for a part where the part in question is also highlighted with a gesture (Clark & Estigarribia, 2011; Kobayashi, 1998; Saylor & Sabbagh, 2002; also Gentner, Loewenstein, & Hung, 2007). In short, children start in early on two semantic relations that link terms within a domain, **inclusion** and **meronymy**.

Children also attend to opposites. Consider their acquisition of dimensional pairs, where one is unmarked and positive, encoding the *presence* of size, height, width, etc., and its opposite is marked and negative, encoding the relative *absence* of extent. Certain opposites emerge early: children start with *big* versus *little* or *small* for size (e.g., Donaldson & Wales, 1970; Ebeling & Gelman, 1994), and only later add terms for height, length, width, thickness, and depth. Adults often refer to several dimensions when talking to small children about whether a particular truck, for instance, will or won't fit under a bridge, or whether a doll's chair is the right height for a table in the doll's house (e.g., Rogers, 1978; Murphy & Jones, 2008). These uses expose children to different pairs of dimensional terms and how they are related.

After *big* and *small*, children add further dimensional pairs in rough order of complexity, with terms for a single dimension – *tall* and *short*, then *long* and *short* (with *high* and *low*), before pairs like *thick* and *thin*, *wide* and *narrow*, before *deep* and *shallow*. When children are unsure of the negative member in a pair, they typically replace it with *little* or *small*, or negate the positive term, as in *not-deep* (Clark, 1972; Murphy & Jones, 2008).

Children also take in other information about semantic relations: which terms are at the same level in a taxonomy (**co-hyponyms**), for example, *zebra, giraffe, tiger*, and *monkey*, among zoo or jungle animals; or *sheep, cow*, *goat, donkey* among farm animals; *duck, chicken, blackbird, robin*, among common birds, and so on. In each domain they gradually add more terms and relate them to such superordinates as *bird*, *mammal*, and *animal*. Adults also talk about parts and properties. For animals, this may include terms like *ear, eye, mouth, hoof, mane, tail, claw*; *collar*, *saddle, reins*, and so on. Adults also supply information about **ontogenesis** with terms for the young for each kind – e.g., *calf, lamb, cub, chick*; for the usual **habitat**: *burrows* or *hutches* for rabbits; *kennels* or *baskets* for dogs; *paddocks* or *stables* for horses; *pastures* or *barns* for cows; for the sounds different kinds make (*roar, neigh, bark*; *chirp, twitter, sing*); and for what they eat.

This information is offered in installments along with new words: different pieces of information relate familiar words to a new word just offered, and in doing this, adults enable children to link new words to terms already known in each domain. The more adults talk about a domain, whether with a puzzle containing zoo-animal shapes or with a toy-farmyard, the more links they offer for integrating new words into existing domains. In short, the semantic relations that children learn, that link word meanings, come from the added information adults provide. This ancillary information comes in the form of 'pragmatic directions' about the relevant relation (Clark, 1998; Clark & Wong, 2002; Clark & Grossman, 1998; Clark & Estigarribia, 2011). While adults appear to focus on inclusion and meronymy early on, they also provide information about opposites, properties, and many other details, depending on what their children already know, and what they ask questions about (Chouinard, 2007).

7. Elaborated domains and islands of expertise

As children add more words to semantic domains, they may develop small islands of expertise. As Crowley and Jacobs (2002) point out, the development of such islands of expertise depends critically on family interactions and encouragement. Some children as young as 2 years of age focus on balls, cars, planes, or trains. They fixate on instances they encounter, collect toy versions of cars, for instance, and play with them obsessively (DeLoache, Simcock, & Macari, 2007).

Early enthusiasm for specific domains is exhibited by some 20% of 4-year-olds. Some develop an intense interest in dinosaurs, for example: they recognize and learn terms for different kinds; they read about them, and may go to museum exhibits, learn about dinosaur fossils, and about modern descendants of dinosaurs (Chi & Koeske, 1983; Gobbo & Chi, 1986). By age 4 and 5, they can develop a rich and quite deep knowledge of their particular interest. Such interests accrue over weeks and months, with support from family members and extensive family discussions. Their commitment sometimes lasts for several years, and generally draws in the rest of the family as well (Crowley & Jacobs, 2002; Johnson et al., 2004).

The children generally master an extensive vocabulary for their domain, with as many as 40 distinct terms (see Chi & Koeske, 1983). In follow-up research on 7-year-olds, Gobbo and Chi (1986) showed children pictures of dinosaurs and asked them to come up with any information they had, including names. They also asked them to sort the pictures that went together. Expert children consistently talked about such factors as diet, and used these in their sorts. Novices sorted the pictures on the basis of perceptual similarity alone. Experts and novices appeared to organize what they knew differently, depending on their degree of expertise (see also Carey, 1985).

In another study of expertise, Johnson and Mervis (1994) looked at how children acquired knowledge about shore birds. They studied 5-year-olds' acquisition of terms for some 14 shore birds and information about each type of bird. They followed what children absorbed and remembered and how they organized what they were learning over four sessions within 17 days, in order to track the transition as children went from novice to (more) expert in this domain. As the children played board games with pictures, they heard names for the birds they saw, and were given information about physical and behavioral attributes for each type.

They were assessed with tests of comprehension and production during the board games, as well as with follow-up tasks such as sorting. Children showed increased comprehension over time. Their knowledge of attributes increased more slowly, and they did least well on name production. The authors also reported a case study of a boy who at age 2 could identify 20 bird types, and, by 4;5, produce over 100 bird names. Since some shore birds were included among the birds he knew, he was asked to play the same board games. His knowledge about birds clearly helped: his scores for all 14 birds in the game were at or close to the maximum. This strongly suggests that children (and adults) are more successful when building on relevant prior knowledge, because they have already organized named types they know, along with their properties.

Alongside any such domains of intense interest, children acquire more terms in everyday domains that form part of what might be considered core knowledge of the language: words for colors and shapes; for foods, utensils, and meal-time routines; for clothing: clothes worn to bed versus clothes for daytime activities, for indoors and outdoors, summer and winter; for toys: for blocks, balls, books, dolls and action figures, animal toys and puppets, pencils and crayons, coloring books and drawing blocks, paints and paint brushes; kitchen utensils and saucepans; certain plants and garden tools like trowels; beach toys including spades, molds, and buckets, as well as words for seaweed, crabs, limpets, and various sea-shells. They also learn words for actions involving all these objects: filling and emptying, opening and closing; gathering, sorting, and separating; placing and moving in space, as they manipulate objects they are playing with.

8. Conclusion

Speakers of any age need to acquire the relevant vocabulary for talking about any new domain they learn about, as Bross (1973) argued for the learning of medical terms: knowing the relevant terminology is a prerequisite for learning about a domain such as anatomy. Without words, speakers find it hard to organize and remember the information they need. As children master a larger vocabulary with more precise labels for the available distinctions among objects, actions, and relations that they attend to every day, they can call on more vocabulary and on more complex ways of talking about what is going on (Crowley & Jacobs, 2002; Rigney & Callanan, 2011).

Gaining any degree of knowledge for children depends heavily on interaction. It is in interaction that children are offered new words and all kinds of supporting information. And it is in interaction that children acquire knowledge about each domain (Chouinard, 2007; Gentner et al., 2016). This interaction occurs in everyday conversational exchanges, in exchanges that accompany activities, whether reading books together or building elaborate Lego constructions, sorting pebbles or growing crystals in a jam-jar, floating ducks in the bath or looking at family photos. It is in joint activities like these that children take up new words. And when they do this, they can also draw on any accompanying information being offered. The more words they acquire in a domain, the better that domain becomes organized, and the easier they find it to add further words linked to the domain. But these interactions are characteristic of middle-class Western families, and may not be represented in all cultures or all social classes. Research is still needed on how speakers build up and organize semantic domains where different patterns of adult-child interaction are the norm.

References

Allen, S., Özyürek, A., Kita, S., Brown, A., Furman, R., Ishizuka, T., & Fujii, M. (2007). Language-specific and universal influences in children's syntactic packaging of Manner and Path: A comparison of English, Japanese, and Turkish. *Cognition*, 102, 16–48. https://doi.org/10.1016/j.cognition.2005.12.006

Ameel, E., Malt, B., & Storms, G. (2008). Object naming and later lexical development: From baby bottle to beer bottle. *Journal of Memory and Language*, 58, 262–285. https://doi.org/10.1016/j.jml.2007.01.006

Andersen, E. S. (1978). Lexical universals of body-part terminology. In J. H. Greenberg, C. A. Ferguson, & E. A. Moravcsik (Eds.), *Universals of human language* (Vol. 3, pp. 335–368). Stanford, CA: Stanford University Press.

Au, T. K. F., & LaFramboise, D. E. (1990). Acquiring color names via linguistic contrast: The influence of contrasting terms. *Child Development*, 61, 1808–1823. https://doi.org/10.2307/1130839

Baldwin, D. A. (1989). Priorities in children's expectations about object-label reference: Form over color. *Child Development*, 60, 1291–1306. https://doi.org/10.2307/1130921

Beavers, J., Levin, B., & Tham, S-W. (2010). The typology of motion expressions revisited. *Journal of Linguistics*, 46, 331–377. https://doi.org/10.1017/S0022226709990272

Bergelson, E., & Swingley, D. (2012). At 6-9 months, human infants know the meanings of many common nouns. *Proceedings of the National Academy of Sciences of the United States of America*, 109(9), 3253–3258. https://doi.org/10.1073/pnas.1113380109

Bion, R. A. H., Borovsky, A., & Fernald, A. (2013). Fast mapping, slow learning: Disambiguation of novel word-object mappings in relation to vocabulary learning at 18, 24, and 30 months. *Cognition*, 126, 39–53. https://doi.org/10.1016/j.cognition.2012.08.008

Booth, A. E., & Waxman, S. R. (2002). Word learning is 'smart': Evidence that conceptual information affects preschoolers' extension of novel words. *Cognition*, 84, B11–B22. https://doi.org/10.1016/S0010-0277(02)00015-X

Bowerman, M. (1996). Learning how to structure space for language: A crosslinguistic perspective. In P. Bloom, M. A. Peterson, L. Nadel, & M. F. Garrett (Eds.), *Language and space* (pp. 385–436). Cambridge, MA: The MIT press.

Bross, I. D. J. (1973). Languages in cancer research. In G. P. Murphy, D. Pressman, & E. A. Mirand (Eds.), *Perspectives in cancer research and treatment* (pp. 213–221). New York, NY: Alan R. Liss.

Callanan, M. A. (1990). Parents' descriptions of objects: Potential data for children's inferences about category principles. *Cognitive Development*, 5, 101–122. https://doi.org/10.1016/0885-2014(90)90015-L

Callanan, M. A., & Sabbagh, M. A. (2004). Multiple labels for objects in conversations with young children: parents' language and children's developing expectations about word meanings. *Developmental Psychology*, 40, 746–763. https://doi.org/10.1037/0012-1649.40.5.746

Carey, S. (1985). *Conceptual change in childhood*. Cambridge, MA: The MIT Press.

Carey, S., & Bartlett, E. (1978). Acquiring a single new word. *Papers and Reports on Child Language Development* [Stanford University] 15, 17–29.

Casasola, M., Bhagwat, J., Burke, A. S. (2009). Learning to form a spatial category of tight-fit relations: How experience with a label can give a boost. *Developmental Psychology*, 45, 711–723. https://doi.org/10.1037/a0015475

Cheung, S-L., & Clark, E. V. (2006). Actions and results in the acquisition of Cantonese verbs. In P. Li, L. H. Tan, E. Bates, & O. Tzeng (Eds.), *Handbook of East Asian psycholinguistics, Vol. 1: Chinese* (pp. 13–22). Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511550751.003

Chi, M. T. H., & Koeske, R. D. (1983). Network representation of a child's dinosaur knowledge. *Developmental Psychology*, 19, 29–39. https://doi.org/10.1037/0012-1649.19.1.29

Choi, S. (1991). Influence of language-specific input on spatial cognition: Categories of containment. *First Language*, 26, 207–232. https://doi.org/10.1177/0142723706060748

Choi, S., & Bowerman, M. (1991). Learning to express motion events in English and Korean: The influence of language-specific lexicalization patterns. *Cognition*, 41, 83–121. https://doi.org/10.1016/0010-0277(91)90033-Z

Choi, S., McDonough, L., Bowerman, M., & Mandler, J. M. (1999). Early sensitivity to language-specific spatial categories in English and Korean. *Cognitive Development*, 14, 241–268. https://doi.org/10.1016/S0885-2014(99)00004-0

Chouinard, M. M. (2007). Children's questions: A mechanism for cognitive development. *Monographs of the Society for Research in Child Development* 72 (Serial No. 286).

Chouinard, M. M., & Clark, E. V. (2003). Adult reformulations of child errors as negative evidence. *Journal of Child Language*, 30, 637–669. https://doi.org/10.1017/S0305000903005701

Cimpian, A., & Markman, E. M. (2008). Preschool children's use of cues to generic meaning. *Cognition*, 107, 19–53. https://doi.org/10.1016/j.cognition.2007.07.008

Clark, E. V. (1972). On the child's acquisition of antonyms in two semantic fields. *Journal of Verbal Learning and Verbal Behavior*, 11, 750–758. https://doi.org/10.1016/S0022-5371(72)80009-4

Clark, E. V. (1973a). What's in a word? On the child's acquisition of semantics in his first language. In T. E. Moore (Ed.), *Cognitive development and the acquisition of language* (pp. 65–110). New York, NY: Academic Press.

Clark, E. V. (1973b). Non-linguistic strategies and the acquisition of word meanings. *Cognition*, 2, 161–182. https://doi.org/10.1016/0010-0277(72)90010-8

Clark, E. V. (1980). 'Here's the Top': Nonlinguistic strategies in the acquisition of orientational terms. *Child Development*, 51, 329–338.

Clark, E. V. (1987). The principle of contrast: A constraint on language acquisition. In B. MacWhinney (Ed.), *Mechanisms of language acquisition* (pp. 1–33). Hillsdale, NJ: Lawrence Erlbaum Associates.

Clark, E. V. (1990). On the pragmatics of contrast. *Journal of Child Language*, 17, 417–431. https://doi.org/10.1017/S0305000900013842
Clark, E. V. (1995). Language acquisition: The lexicon and syntax. In J. L. Miller & P. D. Eimas (Eds.), *Handbook of perception and cognition, Vol. 11: Speech, language, and communication* (2nd ed., pp. 303–337). New York, NY: Academic Press.
Clark, E. V. (1997). Conceptual perspective and lexical choice in acquisition. *Cognition*, 64, 1–37. https://doi.org/10.1016/S0010-0277(97)00010-3
Clark, E. V. (1998). Lexical structure and pragmatic directions in acquisition. In M. C. Gruber, D. Higgins, K. S. Olson, & T. Wysocki (Eds.), *Papers from the 34th Meeting of the Chicago Linguistic Society 1998: The Panels* (pp. 437–446). Chicago, IL: Chicago Linguistic Society.
Clark, E. V. (2004). How language acquisition builds on cognitive development. *Trends in Cognitive Sciences*, 8, 472–478. https://doi.org/10.1016/j.tics.2004.08.012
Clark, E. V. (2007). Young children's uptake of new words in conversation. *Language in Society*, 36, 157–182. https://doi.org/10.1017/S0047404507070091
Clark, E. V. (2010). Adult offer, word-class, and child uptake in early lexical acquisition. *First Language*, 30, 250–269. https://doi.org/10.1177/0142723710370537
Clark, E. V. (2016). *First language acquisition*. Cambridge: Cambridge University Press.
Clark, E. V., & Carpenter, K. L. (1989a). The notion of source in language acquisition. *Language*, 65, 1–30. https://doi.org/10.2307/414840
Clark, E. V., & Carpenter, K. L. (1989b). On children's uses of from, by, and with in oblique noun phrases. *Journal of Child Language*, 16, 349–364. https://doi.org/10.1017/S030500090001045X
Clark, E. V., & Estigarribia, B. (2011). Using speech and gesture to introduce new objects to young children. *Gesture*, 11, 1–23. https://doi.org/10.1075/gest.11.1.01cla
Clark, E. V., & Grossman, J. B. (1998). Pragmatic directions and children's word learning. *Journal of Child Language*, 25, 1–18. https://doi.org/10.1017/S0305000997003309
Clark, E. V., & Svaib, T. A. (1997). Speaker perspective and reference in young children. *First Language*, 17, 57–74. https://doi.org/10.1177/014272379701704903
Clark, E. V., & Wong, A. D-W. (2002). Pragmatic directions about language use: Offers of words and relations. *Language in Society*, 31, 181–212.
Crowley, K., & Jacobs, M. (2002). Building islands of expertise in everyday family activity. In G. Leinhardt, K. Crowley, & K. Knutson (Eds.), *Learning conversations in museums* (pp. 333–358). Mahwah, NJ: Lawrence Erlbaum Associates.
DeLoache, J. S., Simcock, G., & Macari, S. (2007). Planes, trains, automobiles – and tea sets: Extremely intense interests in young children. *Developmental Psychology*, 43, 1579–1586. https://doi.org/10.1037/0012-1649.43.6.1579
Donaldson, M., & Wales, R. J. (1970). On the acquisition of some relational terms. In J. R. Hayes (Ed.), *Cognition and the development of language* (pp. 235–268). New York, NY: Wiley.
Dougherty, J. W. D. (1979). Names for plants and plants for names. *Anthropological Linguistics*, 21, 298–315.
Ebeling, K. S., & Gelman, S. A. (1988). Coordination of size standards by young children. *Child Development*, 59, 888–896. https://doi.org/10.2307/1130256
Ebeling, K. S., & Gelman, S. A. (1994). Children's use of context in interpreting 'big' and 'little'. *Child Development*, 65, 1178–1192. https://doi.org/10.2307/1131313
Estigarribia, B., & Clark, E. V. 2007. Getting and maintaining attention in talk to young children. *Journal of Child Language*, 34(4), 799–814. https://doi.org/10.1017/S0305000907008161
Gelman, S. A., Goetz, P. J., Sarnecka, B. W., & Flukes, J. (2008). Generic language in parent-child conversations. *Language Learning and Development*, 4, 1–31. https://doi.org/10.1080/15475440701542625

Gentner, D. (1975). Evidence for the psychological reality of semantic components: The verbs of possession. In D. A. Norman, D. E. Rumelhart, & the LNR Research Group, *Explorations in cognition* (pp. 211–246). San Francisco, CA: W. H. Freeman.

Gentner, D., Levine, S. C., Ping, R., Isaia, A., Dhillon, S. Bradley, C., & Honke, G. (2016). Rapid learning in a children's museum via analogical comparison. *Cognitive Science*, 40(1), 224–240. https://doi.org/10.1111/cogs.12248

Gentner, D., Loewenstein, J., & Hung, B. (2007). Comparison facilitates children's learning of names for parts. *Journal of Cognition and Development*, 8, 285–307. https://doi.org/10.1080/15248370701446434

Gershkoff-Stowe, L., & Smith, L. B. (2004). Shape and the first hundred nouns. *Child Development*, 75, 1098–1114. https://doi.org/10.1111/j.1467-8624.2004.00728.x

Gobbo, C., & Chi, M. (1986). How knowledge is structure and used by expert and novice children. *Cognitive Development*, 1, 221–237. https://doi.org/10.1016/S0885-2014(86)80002-8

Grice, H. P. (1989). *Studies in the ways of words*. Cambridge, MA: Harvard University Press.

Grimm, H. (1975). On the child's acquisition of semantic structure underlying the word field of prepositions. *Language and Speech*, 18, 97–119.

Gullberg, M. (2011). Language-specific encoding of placement events in gestures. In J. Bohnemeyer & E. Pederson (Eds.), *Event representation in language and cognition* (pp. 166–188). Cambridge: Cambridge University Press.

Haryu, E., & Imai, M. (2002). Reorganizing the lexicon by learning a new word: Japanese children's interpretation of the meaning of a new word for a familiar artifact. *Child Development*, 73, 1378–1391. https://doi.org/10.1111/1467-8624.00478

Haryu, E., Imai, M., & Okada, H. (2011). Object similarity bootstraps young children to action-based verb extension. *Child Development*, 82, 674–686. https://doi.org/10.1111/j.1467-8624.2010.01567.x

Haviland, S. E., & Clark, E. V. (1974). This man's father is my father's son: A study of the acquisition of English kin terms. *Journal of Child Language*, 1, 23–47. https://doi.org/10.1017/S0305000900000064

Heibeck, T., & Markman, E. M. (1987). Word learning in children: An examination of fast mapping. *Child Development*, 58, 1021–1034. https://doi.org/10.2307/1130543

Hills, Thomas. (2013). The company that words keep: Comparing the statistical structure of child- versus adult-directed language. *Journal of Child Language*, 40, 586–604. https://doi.org/10.1017/S0305000912000165

Hohenstein, Jill. (2013). Parent-child talk about motion: Links to children's development of motion event language. *First Language*, 33, 411–425. https://doi.org/10.1177/0142723713490599

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

Johnson, K. E., Alexander, J. M., Spencer, S., Leibham, M. E., & Neitzel, C. (2004). Factors associated with the early emergence of intense interests within conceptual domains. *Cognitive Development*, 19, 325–343. https://doi.org/10.1016/j.cogdev.2004.03.001

Johnson, K. E., & Mervis, C. B. (1994). Microgenetic analysis of first steps in children's acquisition of expertise on shorebirds. *Developmental Psychology*, 30, 418–435. https://doi.org/10.1037/0012-1649.30.3.418

Johnson, K. E., Mervis, C. B., & Boster, J. S. (1992). Developmental changes within the structure of the mammal domain. *Developmental Psychology*, 28, 74–83. https://doi.org/10.1037/0012-1649.28.1.74

Kobayashi, H. (1998). How 2-year-old children learn novel part names of unfamiliar objects. *Cognition*, 68, B41–B51. https://doi.org/10.1016/S0010-0277(98)00044-4
Lakusta, L., & Landau, B. (2012). Language and memory for motion events: Origins of the asymmetry between source and goal paths. *Cognitive Science*, 36, 517–544. https://doi.org/10.1111/j.1551-6709.2011.01220.x
Lakusta, L., Wagner, L., O'Hearn, K., & Landau, B. (2007). Conceptual foundations of spatial language: Evidence for a goal bias in infants. *Language Learning and Development*, 3, 179–197. https://doi.org/10.1080/15475440701360168
Landau, B., Johannes, K., Skordos, D., & Papafragou, A. (2017). Containment and support: Core and complexity in spatial language learning. *Cognitive Science*, 41(S4), 748–779. https://doi: 10.1111/cogs.12389
Landau, B., Smith, L., & Jones, S. (1998). Object shape, object function, and object name. *Journal of Memory and Language*, 38(1), 1–27. https://doi.org/10.1006/jmla.1997.2533
Lehrer, A. J. (1974). *Semantic fields and lexical structure*. Amsterdam: North-Holland.
Lehrer, A., & Kittay, E. F. (Eds.). (1992). *Frames, fields, and contrasts: New essays in semantic and lexical organization*. Hillsdale, NJ: Lawrence Erlbaum Associates.
Lyons, J. (1977). *Semantics*, Vol. 1. Cambridge: Cambridge University Press.
Macnamara, J. (1982). *Names for things: A study of human learning*. Cambridge, MA: The MIT Press.
Majid, A., Boster, J., & Bowerman, M. (2008). The cross-linguistic categorization of everyday events: A study of cutting and breaking. *Cognition*, 100, 235–250. https://doi.org/10.1016/j.cognition.2008.08.009
Majid, A., Gullberg, M., van Staden, M., & Bowerman, M. (2007). How similar are semantic categories in closely related languages? A comparison of cutting and breaking in four Germanic languages. *Cognitive Linguistics*, 18, 179–194. https://doi.org/10.1515/COG.2007.007
Malt, B., White, A., Ameel, E., & Storms, G. (2016). Learning the language of locomotion: Do children use biomechanical structure to constrain hypotheses about word meaning? *Language Learning and Development*, 12, 357–379.
Markman, E. M., & Wachtel, G. (1988). Children's use of mutual exclusivity to constrain the meanings of words. *Cognitive Psychology*, 20, 121–157. https://doi.org/10.1016/0010-0285(88)90017-5
Masur, E. F. (1997). Maternal labelling of novel and familiar objects: Implications for children's development of lexical constraints. *Journal of Child Language*, 24, 427–439. https://doi.org/10.1017/S0305000997003115
Mervis, C. B., Bertrand, J., & Pani, J. R. (1995). Transaction of cognitive-linguistic abilities and adult input: A case study of the acquisition of colour terms and colour-based subordinate object categories. *British Journal of Developmental Psychology*, 13, 285–302. https://doi.org/10.1111/j.2044-835X.1995.tb00680.x
Mervis, C. B., Golinkoff, R. M., & Bertrand, J. (1994). Two-year-olds readily learn multiple labels for the same basic-level category. *Child Development*, 65, 1163–1177. https://doi.org/10.2307/1131312
Mervis, C. B., & Rosch, E. (1981). Categorization of natural objects. *Annual Review of Psychology*, 32, 89–115. https://doi.org/10.1146/annurev.ps.32.020181.000513
Moore, C., & Dunham, P. J. (Eds.). (1995). *Joint attention: Its origins and role in development*. Hillsdale, NJ: Lawrence Erlbaum Associates.
Murphy, M. L., & Jones, S. (2008). Antonyms in children's and child-directed speech. *First Language*, 28, 403–430. https://doi.org/10.1177/0142723708091047

Naigles, L. R., & Terrazas, P. (1998). Motion-verb generalization in English and Spanish. *Psychological Science*, 9, 363–369. https://doi.org/10.1111/1467-9280.00069

Narasimhan, B., Kopecka, A., Bowerman, M., Gullberg, M., & Majid, A. (2012). Putting and taking events: A crosslinguistic perspective. In A. Kopecka & B. Narasimhan (Eds.), *Events of putting and taking: A crosslinguistic perspective* (pp. 1–18). Amsterdam: John Benjamins. https://doi.org/10.1075/tsl.100.03nar

O'Hanlon, C. G., & Roberson, D. (2006). Learning in context: Linguistic and attentional constraints on children's color term learning. *Journal of Experimental Child Psychology*, 94, 275–300. https://doi.org/10.1016/j.jecp.2005.11.007

Özçalışkan, Ş., & Slobin, D. I. (1999). Learning 'how to search for the frog': Expression of manner of motion in English, Spanish and Turkish. In A. Greenhill, H. Littlefield, & C. Tano (Eds.), *BUCLD 23: Proceedings of the 23rd annual Boston University Conference on Language Development* (pp. 541–552). Somerville, MA: Cascadilla Press.

Papafragou, A. (2010). Source-goal asymmetries in motion representation: Implications for language production and comprehension. *Cognitive Science*, 34, 1064–1092. https://doi.org/10.1111/j.1551-6709.2010.01107.x

Pruden, S. M., Roseberry, S., Göksun, T., Hirsh-Pasek, K., & Golinkoff, R. M. (2013). Infant categorization of path relations during dynamic events. *Child Development*, 84, 331–345. https://doi.org/10.1111/j.1467-8624.2012.01843.x

Pulverman, R., Golinkoff, R. M., Hirsh-Pasek, K., & Buresh, J. S. (2008). Infants discriminate manners and paths in non-linguistic dynamic events. *Cognition*, 108, 825–830. https://doi.org/10.1016/j.cognition.2008.04.009

Pye, C., Loeb, D. F., & Pao, Y. (1996). The acquisition of breaking and cutting. In E. V. Clark (Ed.), *Proceedings of the Twenty-Seventh Annual Child Language Research Forum* (pp. 227–236). Stanford, CA: CSLI.

Ravn, K. (1988). On calling things names (Unpublished doctoral dissertation). Stanford University.

Rice, S. (2003). Growth of a lexical network: Nine English prepositions in acquisition. In H. Cuyckens, R. Dirven, & J. R. Taylor (Eds.), *Cognitive approaches to lexical semantics* (pp. 243–280). Berlin: De Gruyter. https://doi.org/10.1515/9783110219074.243

Rigney, J., & Callanan, M. (2011). Patterns in parent-child conversations about animals at a marine science center. *Cognitive Development*, 26, 155–171. https://doi.org/10.1016/j.cogdev.2010.12.002

Rogers, D. (1978). Information about word-meaning in the speech of parents to young children. In R. N. Campbell & P. T. Smith (Eds.), *Recent advances in the psychology of language* (pp. 187–198). London: Plenum.

Roy, B. C., Frank, M. C., & Roy, D. (2012). Relating activity contexts to early word learning in dense longitudinal data. *Proceedings of the 34th Annual Meeting of the Cognitive Science Society* (UC Merced), vol. 34, pp. 935–940. https://escholarship.org/uc/item/9rq595qm

Saji, N., Imai, M., Saalbach, H., Zhang, Y., Shu, H., & Okada, H. (2011). Word learning does not end at fast-mapping: Evolution of verb meanings through reorganization of an entire semantic domain. *Cognition*, 118(1), 45–61. https://doi.org/10.1016/j.cognition.2010.09.007

Sandhofer, C. M., & Smith, L. B. (1999). Learning color words involves learning a system of mappings. *Developmental Psychology*, 35, 668–679. https://doi.org/10.1037/0012-1649.35.3.668

Saylor, M. M., & Sabbagh, M. A. (2002). Children use whole-part juxtaposition as a pragmatic cue to word meaning. *Developmental Psychology*, 38, 993–1003. https://doi.org/10.1037/0012-1649.38.6.993

Shatz, M., Behrend, D., Gelman, S. A., & Ebeling, K. S. (1996). Colour term knowledge in two-year-olds: Evidence for early competence. *Journal of Child Language*, 23, 177–200. https://doi.org/10.1017/S030500090001014X
Sinha, C., & Jensen de López, K. (2000). Language. culture, and the embodiment of spatial cognition. *Cognitive Linguistics*, 11, 17–41.
Slobin, D. I. (2004). The many ways to search for a frog: Linguistic typology and the expression of motion events. In S. Strömqvist & L. Verhoeven (Eds.), *Relating events in narratives, Vol. 2: Typological and contextual perspectives* (pp. 219–257). Mahwah, NJ: Lawrence Erlbaum Associates.
Slobin, D. I. (2006). What makes manner of motion salient? Explorations in linguistic typology, discourse, and cognition. In M. Hickmann & S. Robert (Eds.), *Space in languages: Linguistic systems and cognitive categories* (pp. 59–81). Amsterdam: John Benjamins. https://doi.org/10.1075/tsl.66.05slo
Talmy, L. (1985). Lexicalization patterns: Semantic structure in lexical forms. In T. Shopen (Ed.), *Language typology and syntactic description, Vol. 3: Grammatical categories and the lexicon* (pp. 57–149). Cambridge: Cambridge University Press.
Tomasello, M. (1987). Learning to use prepositions: A case study. *Journal of Child Language*, 14, 79–98. https://doi.org/10.1017/S0305000900012745
Tomasello, M. (1995). Joint attention as social cognition. In C. Moore & P. J. Dunham (Eds.), *Joint attention: Its origins and role in development* (pp. 103–130). Hillsdale, NJ: Lawrence Erlbaum Associates.
Tomasello, M., & Barton, M. (1994). Learning words in non-ostensive contexts. *Developmental Psychology*, 30, 639–650. https://doi.org/10.1037/0012-1649.30.5.639
Tomasello, M., & Kruger, A. (1992). Joint attention on actions: Acquiring verbs in ostensive and non-ostensive contexts. *Journal of Child Language*, 19, 311–333. https://doi.org/10.1017/S0305000900011430
Tversky, B., & Hemenway, K. (1984). Objects, parts, and categories. *Journal of Experimental Psychology: General*, 113, 169–193. https://doi.org/10.1037/0096-3445.113.2.169
Waxman, S. R., & Senghas, A. (1992). Relations among word meanings in early lexical development. *Developmental Psychology*, 28, 862–873. https://doi.org/10.1037/0012-1649.28.5.862
Yurovsky, D., Smith, L. B., & Yu. C. (2013). Statistical word learning at scale: The baby's view is better. *Developmental Science*, 16, 959–966.

Additional resources

These references are the primary resource for research. Other general resources include the CHILDES Archive <http://childes.psy.cmu.edu/>, the bibliography 'Acquisition of language' in Oxford Bibliographies Online: Linguistics <http://www.oxfordbibliographiesonline.com/>, and the text and reference source: Clark, E. V. (2016). *First language acquisition* (3rd ed.). Cambridge: Cambridge University Press.

CHAPTER 3

The influence of linguistic temporal organization on children's understanding of temporal terms and concepts

Laura Wagner
Ohio State University

The Typological Prevalence Hypothesis argues that the cognitive biases that lead to common cross-linguistic patterns also guide language development. Two studies investigated English learning children's understanding of temporal terms and concepts and how they might be influenced by three typologically common patterns: the use of a deictic center, the asymmetry between past and future times, and the use of temporal remoteness. Using a timeline task, five- and eight-year-old children were asked to locate events with a range of temporal expressions as well as with general event descriptions. The results provided only limited support for the TPH for the temporal domain but do demonstrate the potential for the hypothesis as a research approach.

Keywords: tense, time, timeline, temporal semantics, typology

1. Introduction

This paper investigates children's developing understanding of temporal terms in English and the concepts that organize linguistic temporal systems. The approach adopted is inspired by the Typological Prevalence Hypothesis (TPH) of Gentner & Bowerman (2009, see also Slobin, 1985). This hypothesis states that, all things being equal, if a particular semantic element is commonly grammaticized across a diverse set of the world's languages (i.e., it is typologically prevalent) then it will be comparatively easy for children to learn. The intuition is that those concepts which humans are cognitively pre-disposed to learn will not only shape the local learning processes of individual children, but will also shape larger linguistic patterns around the globe. Of course, in the real world, not all things are equal, and so the hypothesis is not couched in terms of absolute universals or necessities. Nevertheless, the TPH serves as a useful tool for researchers by suggesting what the learning biases may be

https://doi.org/10.1075/tilar.24.03wag

for children acquiring a language and therefore, the kinds of meaning components that might influence children's interpretations – even components that are not directly encoded in the child's target grammar.

Temporality is one of the handful of conceptual domains that is reliably grammaticized by languages, although the specific encodings vary dramatically from language to language (see Dahl, 1985). Moreover, temporality is a multi-layered domain involving event structure, perspective, aspectual information, and linear ordering among other things (see Binnick, 2012). The breadth of this domain is too large for the scope of this paper, and the current experiments will focus exclusively on tense, that is, on the specification of the past, present, and future. More specifically, they will focus on three dimensions of tense meaning that are commonly grammaticized within a range of languages and therefore good candidates for the TPH. These dimensions are the DEICTIC CENTER, an ASYMMETRY BETWEEN THE PAST AND FUTURE, and the ENCODING OF TEMPORAL REMOTENESS (or distance) from the deictic center.

Most formal accounts organize tenses around a DEICTIC CENTER (cf. Reichenbach, 1947; Binnick, 1991; Klein, 1994, inter alia); usually that center is the speaker's time of utterance, which is also often linked to the present moment. Past tenses refer to times occurring before the time of utterance and future tenses refer to times occurring after the time of utterance. This type of organization is not the only way one can define time. For example, many philosophers (and some physicists) have argued that time is better understood without the deictic reference point, as a simple linear sequence of moments with no special status for the speaker's role within it (see, for example, Ludlow, 1999). Thus, while using a deictic center is not a necessary way of encoding time, it is a common one across languages. In addition, given that the current work will focus on children acquiring English, it is worth noting that the deictic center is part of how English grammaticizes tense.

An ASYMMETRY BETWEEN THE PAST AND FUTURE manifests itself across the world's languages in a variety of ways. Most notably future markers in many languages, including English, are homophonous with markers of present tense and modality (Comrie, 1985; Deo, 2012). In English, there are no verb auxiliaries or morphological markers that exclusively express future meaning: all markers used for the future are also, synchronically, used for other functions such as present tense (the progressive *go*) or modality (*will, shall*). By contrast the English past tense has distinctive forms (*was, -ed)*. And, while it is true that English's simple past tense form conveys perfective aspectual meaning in addition to tense meaning, not all past tense forms do so (e.g., *was* does not).

TEMPORAL REMOTENESS, where remoteness is calculated relative to the deictic center, is also a common feature of many linguistic tense systems (Dahl, 1985; Botne, 2012). For example, Washo (a native American language in the Hokan family) grammatically marks not only whether an event happened in the past, but whether it happened within the past day (hodiernal -*leg*), previous to yesterday but not too far away (-*ayʔ*), within the speaker's memory (-*gul*), or beyond the speaker's memory (-*lul*) (Botne, 2012). Remoteness distinctions are found in both the past and future tenses, although they are more common in the past tenses (Deo, 2012). Note that for languages with remoteness marking, the specification is obligatory. While a language like English certainly allows one to provide a precise temporal distance (e.g., I wrote this sentence at 4:23 on October 18th), those distances are not grammaticized in the language and are wholly optional; English does not directly code remoteness within the grammatical system.

These three temporal elements, therefore, pass the first hurdle for the TPH: they are all well described as being part of the grammatical systems in many of the world's languages. The next step is to consider whether any corroborating evidence exists within the cognitive domain. By hypothesis, these elements are prevalent because they reflect something "natural" about how we build these concepts and so they should matter for children's learning trajectories (and potentially also for adults' representations).

The strongest support for cognitive naturalness for any of these three temporal concepts can be found for the idea that the past and future are asymmetric and further, that the future is less well specified. For example, McCormack and Hanley (2011) found that pre-school aged children were better at reasoning about the past than the future, and Busby Grant and Suddendorf (2009) similarly found that three- and four-year-olds are more accurate at locating events on a timeline for past times than future times. There is more limited support for the cognitive underpinnings of the other typological elements considered here. Suddendorf (2010) reviews ERP research that found that adult brains show more activation when reasoning about distal (i.e., remote) events than close ones. He argues that it takes more effort to think about things that are not close to the immediate moment. Moreover, there is also a long tradition going back to Piaget (1969) suggesting that children initially reason better about the here-and-now (the present) than they do about more distant time periods. With respect to the deictic center, there is some evidence that children as old as nine years may have difficulties using it when reasoning about times. In various studies by Friedman (2002; Friedman, Garder & Zubin, 1995), children located events along a road that corresponded to a timeline. A subset of the children – including some 9-year-olds – interpreted the road in a non-deictic fashion even after being explicitly told not to. These data, however, are difficult to interpret as the critical cases involved cyclical daily events such as having breakfast. The non-deictic responses arose when children put breakfast at the beginning of the road (since it starts the day) instead of much further on (since children were tested mid-day and their next breakfast was in the future). Cyclic reasoning may be especially difficult to do, but it is something that children must connect to their deictically organized tense system; the fact that this integration takes years to accomplish suggests some dimension of this integration is quite difficult.

Overall, the existing work on temporal cognition does not provide particularly strong evidence for the TPH and raises the possibility that the causal direction between cognitive organization and language development goes the other way. An alternative to the TPH is Slobin's "thinking for speaking" hypothesis (e.g., Slobin, 1996). From this point of view, children are initially more open-minded about how to conceptualize domains such as time, but through the process of learning their native language they develop habits for how to think about the domain. This position predicts that influences should be restricted to elements children are exposed to (i.e., that are grammaticized in their native language). Moreover, it predicts that these deep organizational influences will take time to be felt: it could take years before a child's target language organization has a strong influence on their habits of cognitive organization.

One final literature to consider is work looking explicitly at children's acquisition of tense meanings. Temporal morphology is among the earliest produced by children (e.g., Brown, 1973) and its use appears to be coordinated with systematic grammatical rules, as has been discussed in the Optional Infinitive debates (see Poeppel & Wexler, 1993). Studies examining children's ability to assign meaning to tense markers suggest that children begin to make language specific mappings before the age of 3 years (Valian, 2006) and that they have largely completed their acquisition of grammatical tense by the age of 5 years across a variety of different languages (Wagner, 2012; Maastricht & Hollebrandse, 2011; Weist, Atanassova, Wysocka, & Pawlak, 1999). All of the languages examined in detail have been organized around a deictic center, and so these data support the idea that children can learn such tense systems in this time period. Interestingly, a recurring theme of the literature has been that children's performance is often facilitated by the inclusion of a future tense form in the task (Weist, Wysocka & Lyytenin, 1991; Weist et al., 1999; Wagner, 2001; Valian, 2006) suggesting not only that the future may have a distinctive status in children's early grammatical systems, but that it may be easier to acquire; early facility with future forms is not obviously predicted by the typological or cognitive results noted above but it could be a consequence of the less differentiated nature of the future – if there's less specificity to learn, it may be easier to accomplish. No studies have explicitly considered how temporal remoteness might influence children's early tense understanding but this is hardly surprising given that none of the languages closely investigated have grammaticized remoteness.

The current pair of studies investigated the TPH by asking whether children's temporal understanding was influenced by typologically common patterns even when those patterns are not actually grammaticized in their own language. A timeline task was used in which children were asked to locate events along a line with a pre-established deictic center. This task allows children to demonstrate their understanding of three typologically prevalent temporal properties. The use of a deictic center to organize temporal interpretations (a property that is grammaticized in the participants' native language of English) would be shown by correctly placing events relative to the established center. An asymmetry between the past and future (a property that is debatably grammaticized in English) would be shown by more correct event placements, either in isolation or in ordered pairs, in one timezone (past or future) than the other. An influence of remoteness (a property that is not grammaticized in English) would be shown by a differentiation in performance as a function of how distant from the present time the event was/will be.

The timeline task is not without limitations. Children in the younger age group (5-year-olds) have relatively little experience with timelines and this necessitated a brief training session. Children in the older age group (8-year-olds) by contrast have likely received some amount of explicit training with timelines in a school setting. As shall be shown, these issues are of little practical importance: the younger children did perform coherently on the task in general, making performance differences between ages meaningful. One additional issue concerns the radically different size of the temporal intervals being mapped onto lines of identical length: for example, across items, children located events from the recent past (e.g., eating breakfast), the distant past (e.g., being a baby), and the mythic past (e.g., when dinosaurs were on earth). Children were not provided with any instructions about how to handle the range problem across items, however, a very conservative coding strategy that simply considered regions of the line (rather than absolute distances) was adopted to help counteract differences in range interpretations across items.

To assess children's temporal interpretations with elements outside of the grammaticized tense system of English, children were tested with event descriptions that required the child to make knowledge-based inferences about when the events occurred, and they were also tested on a range of temporal adverbials. The adverbial forms were all checked in the CHILDES database and all appeared in the speech of multiple children. All the events used were chosen to reflect events children would be reasonably familiar with (e.g., driving, being a baby) or had stereotypical associations with (e.g., science fiction scenarios, famous historical periods). Informal piloting confirmed that children had relevant associations to the events used.

A possible concern is that the distal future elements all had a science-fiction aspect to them and suggested it might invoke a fantasy element for these items. And indeed, all the pictures were chosen to encourage children to see the events as very unlike the current day (e.g., the future robot cleaning one's home was not depicted as a Roomba, but more like Rosie of the Jetsons). The future is inherently irrealis – by definition, it does involve possible worlds and an imaginative sense the past does not. To that extent, this problem is inherent in what it means to be part of the future. However, in the analyses, the more mythic times were collapsed with more "realistic" future/past times (being a baby, being a mom/dad) and inspection showed that they did not pattern substantially differently from each other.

The dimension of language that the TPH is generally thought to depend on is grammatical: grammars functionally organize a language to a much greater extent than open class words within a lexicon. Grammatical tense systems clearly fall within this dimension, and general descriptions of events clearly do not. Temporal adverbs, however, potentially occupy an intermediate space. Investigations of children's understanding of temporal adverbs suggests that preschoolers understand that they convey some kind of temporal information – but that it takes until children are around 7 to 8 years old before most individual items are correctly mapped onto specific temporal concepts (Shatz, Tare, Nguyen, & Young, 2010; Tillman & Barner, 2015; but see Busby Grant & Suddendorf, 2011). This shift from a more general temporal grouping to specific temporal understandings may explain why temporal adverbs have inconsistently influenced how pre-school aged children interpret grammaticized tense marking (Wagner, 2001; Valian, 2006). Within the current context, the shift may also reflect development in children's understanding about exactly which temporal concepts are in fact being grammaticized in their language. Adverbial expressions do permit a much wider range of temporal points to be targeted (including points not grammaticized by the children's target language); however, given their intermediate status, adverbs will be analyzed separately from the event descriptions whose temporal status depends on knowledge-based inferences.

If the TPH is grounded in cognition itself, then any effects of common patterns should arise regardless of whether children are tested with explicit linguistic expressions about time (the adverbials) or with general event descriptions. However, it is also possible that the effects of linguistic typology are restricted to linguistic items; if so, then the effects of the more common patterns may be seen more strongly (or perhaps only) when children are tested with linguistic encodings of time. By contrast, a "thinking for speaking" approach would more easily predict that the presence of explicit linguistic elements would better foster habits that are grounded in speech. More generally, the "thinking for speaking" position predicts that children should not be influenced by features not grammaticized in their target language (like remoteness) and in general, any effects of language structure should be stronger as children get older and have more developed habits.

2. Experiment 1

2.1 Participants

Participants were tested in a local science center. The Younger group consisted of 24 children (*M* = 5;7, 9 girls) and the Older group consisted of 10 children (*M* = 8;5, 8 girls). All children had English as their sole or dominant language.

2.2 Stimuli

The task was performed using a timeline consisting of a line approximately 7 inches long with a large star in the middle (approximately 1 inch wide). Each trial used a different line, so children always started with a fresh, blank timeline. Children marked times by placing pictures (~1 inch wide) on the line. The pictures were described in different ways depending on the phase of the task – in the Explicit Linguistic phase, the picture depicted a potentially recurring event (e.g., having a birthday party, playing with Lego blocks) combined with a specific temporal descriptor (such as *last year*). In the General Event Description phase, the pictures depicted events that were naturally anchored in time. The events were described in terms of their content (e.g., "imagine this is a picture of you as a baby") but no specific temporal descriptors were added. That said, it frequently happened that the experimenter added appropriate tense forms on the verb ("when you **were** a baby") so there was a limited amount of linguistic tense that occurred as well. The full set of items and their descriptors are shown in Figure 1.

The target items ranged over the past, present and future times, reflecting the three possible relations to a deictic center. Moreover, for the past and future times, the items also ranged over levels of remoteness. Three remoteness divisions were used. Very Close times (within a matter of minutes or hours) were used to target a potential "vast now" period that might include quite loose boundaries of the present time; Close times (within a matter of days or the week), and Far times (including the span of a person's lifetime as well as the mythic past and future) were chosen to link to distances that are grammaticized in other languages (but not in English). Note that typologically, many languages that grammaticize remoteness also combine those levels of remoteness into a single marker. However, because the span of a lifetime and mythic times are conceptually different, the data was inspected to insure that the effects were not being driven primarily by one or the other of these distances. In addition, one Unspecified level of remoteness was used for a subset of items in the Explicit Linguistic phase that were described with markers that signaled a deictic relationship but no specific temporal remoteness (e.g., plain tenses).

Explicit Linguistic Adverbs/Tenses					
Past Items		**Present Items**		**Future Items**	
Did (unspecified)	E1	Is doing	E1	Is going to do (unspecified)	E1
Before (unspecified)	E2	Today	E1	After (unspecified)	E2
Just (very close)	E1	This week	E2	Soon (very close)	E1
A little bit ago (very close)	E2	This year	E1	In a little bit (very close)	E2
Yesterday (close)	E1			Tomorrow (close)	E1
Last week (close)	E2			Next week (close)	E2
Last year (far)	E1			Next year (far)	E1

Naturally anchored general event descriptions
Note that only a single Present Time item was used in both experiments: *When you are sitting with me.*

Past Items		**Future Items**	
When you came to <museum space> (very close)	E1 (single) E2 (pair)	When you leave <museum space> (very close)	E1 (single) E2 (pair)
When you drove here (close)	E1 (pair)	When you drive home (close)	E1 (pair)
When you ate breakfast (close)	E2 (single) E2 (pair)	When you go to bed (close)	E2 (single) E1 (pair)
When you wore diapers (far)	E1 (single) E2 (pair)	When you are a mom or dad yourself (far)	E1 (single) E2 (pair)
When people wore togas (far)	E1 & 2 (pair)	When people drive spaceships to work (far)	E2 (single) E1 & 2 (pair)
When dinosaurs were on earth (far)	E2 (single) E1 & 2 (pair)	When robots clean our homes (far)	E2 (pair) E1 & 2 (pair)

Figure 1. Stimuli used across both experiments. The temporal remoteness classification for past and future items is indicated in parentheses; all present items were treated as equivalently remote from the current time. The experiment in which the item was used is indicated after each item. For the naturally anchored event descriptions, some items were placed singly (single) and others were placed as parts of pairs (pair).

2.3 Procedures

All sessions began with a brief practice session about how to use the timeline. Children were told that the star represented *right now*, the region to the left of the star was where things have happened *already* and the region after the star was *later* than now (note, specific adverbs were chosen based on Wagner, 2001, as well as informal piloting). Children placed three pictures and received corrective feedback. No additional feedback was given after the practice phase.

Testing began with the Explicit Linguistic phase, in which children were provided with explicit temporal specifications to describe recurring events. Children were told explicitly that the events recurred (e.g., "This is something the boy does a lot") but the event itself was not described. Children also received one recurring item with plain tense marking, but the tense bearing element received strong prosodic stress (e.g., "Show me where it [the picture] goes for when the boy IS doing it") to encourage attention to the tense per se. Children located the recurring events multiple times (e.g., "Show me where it goes for last year/this year/next year"), but they received a new timeline for each picture placement. Children placed pictures for 12 items in this phase.

Testing continued with the General Event Description phase which had two parts. First, children were provided with a single picture and a description of it that implied a particular time (e.g., people don't currently wear togas, nor can the children being tested currently drive their own cars). Children were asked to place the picture on the timeline where it was supposed to go. Next, children were given pairs of pictures and were provided with descriptions for each. They were asked to place both pictures on the timeline. Children placed 5 single pictures and 4 pairs of pictures.

2.4 Coding

For all experimental phases, the experimenter marked on the line where children located the picture – the center of the picture was used as the chosen location regardless of orientation. To code the time markings, a clear plastic template was made that broke the line down into one-inch regions; these regions were given numerical ranks with −3 being the left-most (or most distantly past) region, 0 being the center star (or present time) and + 3 being the right-most (or most distantly future) region. The template was laid over the child's line and the location of the child's picture placement was given a numerical score corresponding to its region. Placements on the border between two regions were given intermediate credit (e.g., 1.5 indicated that the picture's midpoint was on the line between regions 1 and 2).

In effect, this coding scheme gives a basic temporal rank ordering to children's temporal placements, where the size of the number signals remoteness from the present, and the valence signals general time (negative numbers are past, positive numbers are future, zero is present).

2.5 Results

2.5.1 *Ability to use the timeline appropriately*

The first analysis consisted of a validity check: do children understand how to use a timeline in an appropriate fashion at all? The essential feature of a timeline is that spatial distance should correspond to temporal distance. To check whether children were sensitive to this overall fact, all the pictures used in the task were ordered temporally from past to future; events happening at the same time (e.g., at the present time) were given the same rank. Temporally unspecified items (e.g., plain tenses) were ranked as very close to the present time. The ordering was established by discussion among the author and the two coders, but there was little disagreement. This ordering was treated as the "objectively true" times for each event. Children's temporal choices for each picture were determined by where they were located on the timeline. Significant positive correlations were found between the objectively true ordering and the ordering on the timeline for all the items placed by the Younger age group (Pearson's $r = .26$, $N = 572$, $p < .001$) and by the Older age group (Pearson's $r = .71$, $N = 237$, $p < .001$). Figure 2 shows the correlations for the children in both Experiments 1 and 2. On the whole, these data show that children are indeed capable of translating their temporal understanding into this spatial representation.

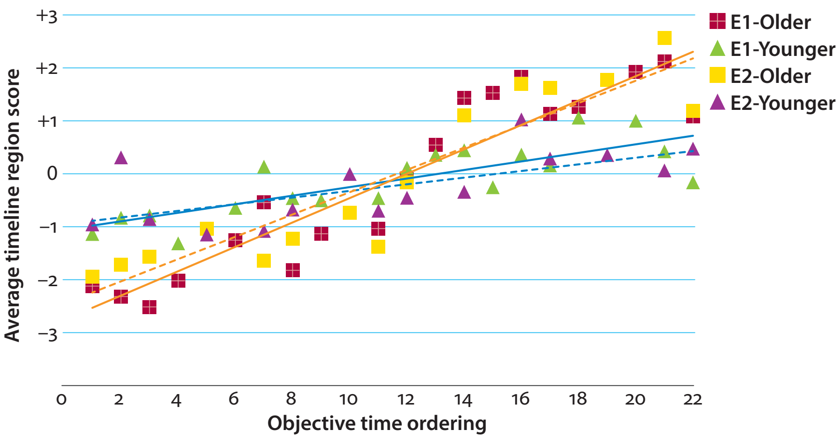


Figure 2. Children's use of the timeline

The Objective Time Ordering is the true temporal ordering of the items used with lower scores being further in the past. The Average Timeline Region Score is the location on the timeline where children placed that item with lower scores being further in the past (zero represents the present timezone). All groups of children showed significant positive correlations, reflecting veridical use of the timeline. Filled-in symbols/solid lines represent children from Experiment 1 and Open symbols/dashed lines represent children from Experiment 2; Squares/grey lines represent Older children (8-year-olds) and Triangles/black lines represent Younger children (5-year-olds).

2.5.2 *Analysis of effects*

The next set of analyses examined whether children understood time as organized around a deictic center. Each item was categorized as picking out a time in the past, the present, or the future (for the paired event descriptions, each half of the pair was treated as a separate item in this analysis); children's answers were coded as being on the central star (associated with the present), to the left of the star (associated with the past) or to the right of the star (associated with the future). Overall rates of correctness for all three timezones were very high for both age groups, and most scores were above chance levels (see Table 1). A repeated measures ANOVA was conducted using age group as a between-subjects variable, cue phase (Explicit Linguistic vs. General Event Descriptions) and target timezone (Past, Present, and Future) as the independent variables, and proportion correct as the dependent variable. A main effect was found for age group ($F(1, 32) = 27.12$, $p < .001$, partial eta^2 = .46) as the Older children out-performed Younger children in general, but there were no significant interactions with age group as the overall patterns of ease and difficulty were the same across ages. No main effects were found between the Explicit Linguistic items and General Event Descriptions ($F(1, 32) = 1.56$, $p = 0.22$) or for the three timezones ($F(2, 31) = 0.84$, $p = 0.49$), but there was a significant interaction between those two variables ($F(2, 31) = 4.11$, $p = .026$, partial eta^2 = .21). The interaction stemmed from the fact that while there was a slight (non-significant) advantage favoring the General Event Descriptions over the Explicit Linguistic cues for items in the past and present, but for the future times, there was a significant advantage favoring the items with Explicit Linguistic cues ($M = .69$) over those with General Event Descriptions ($M = .56$) ($t(33) = 2.35$, $p < .025$). Thus, overall children are highly accurate in this task, and their ability to reason temporally about the future was facilitated by having explicit linguistic direction.

The paired Event Descriptions permitted two further analyses: First, did children do better when the two items within the pair were in different timezones? To the extent that the deictic center organizes temporal organization, it was predicted that children would do better with pairs containing one item from the past and one from the future relative to two past items or two future items. A repeated measures ANOVA was conducted with the independent variables of age group and time-pairings (Both Past, Both Future, One Past & One Future). Results showed that there was improvement with age ($F(1, 32) = 12.69$ $p < .001$, partial eta^2 = .28) but no interaction between age group and time pairings. However, there was a significant main effect of type of time pairing ($F(2, 32) = 9.98$, $p < .001$, partial eta^2 = .24). Post-hoc paired t-tests showed that children were more likely to be correct when the two items crossed the deictic center with one in the past and one in the future than when both items were in the past ($t(33) = 3.4$, $p = .002$) or both items were in the future ($t(33) = 5.2$, $p < .001$). There was no difference in overall

Table 1. Summary of accuracy scores for both experiments
The overall accuracy for correct timezone placement for the different conditions across both age groups and both experiments (standard errors in parentheses). Asterisks mark values that are significantly different from chance ($p < .05$), which was set at .33 reflecting the three major timezones of past, present and future. All non-significant results are highlighted. For paired event descriptions, each picture was coded separately.

			Mean Timezone Accuracy			
			5-year-olds		8-year-olds	
			Exp 1	Exp 2	Exp 1	Exp 2
Linguistic Items	Past	Far	.74 (.09)*	n/a	.9 (.1)*	n/a
		Close	.58 (.1)*	.68 (.1)*	.9 (.11)*	.77 (.12)*
		Very Close	.29 (.09)	.36 (.1)	.6 (.16)	.77 (.12)*
		Unspecified	.67 (.09)*	.55 (.1)	.9 (.1)*	.85 (.1)*
	Present		.51 (.07)*	.45 (.1)	.77 (.09)*	.46 (.14)
	Future	Very Close	.57 (.1)*	.32 (.1)	.7 (.15)*	.69 (.13)*
		Close	.61 (.1)*	.45 (.1)	.9 (.11)*	.92 (.08)*
		Far	.78 (.09)*	n/a	.9 (.1)*	n/a
		Unspecified	.5 (.1)	.45 (.1)	.9 (.1)*	.77 (.12)*
Event Descriptions	Past	Far	.60 (.06)*	.61 (.07)*	.9 (.05)*	.87 (.08)*
		Close	.71 (.09)*	.72 (.1)*	.9 (.1)*	.84 (.1)*
		Very Close	.52 (.08)*	.72 (.1)*	.85 (.07)*	.77 (.12)*
	Present		.63 (.1)*	.55 (.1)	1.0 (.0)*	1.0 (.0)*
	Future	Very Close	.52 (.08)*	.95 (.05)*	.95 (.05)*	.92 (.08)*
		Close	.25 (.09)	.55 (.1)	.6 (.2)	.92 (.08)*
		Far	.56 (.08)*	.57 (.06)*	.93 (.04)*	.83 (.07)*

deictic accuracy (that is, getting the pictures on the correct side of the timeline) between the all-past and all future items ($t(33) = 1.7$, $p = .10$). Also note that the paired items crossing the deictic center did not all create a larger objective temporal distance between the two items and children performed virtually identically with items that were maximally distal from each other ("people dressed in togas" vs. "when people have robots clean their homes") and minimally distal from each other ("when you drove here" vs. "when you drove home"). What facilitates children's performance is having the deictic center temporally located in between the events.

The paired event descriptions also allowed for the examination of children's ability to correctly order the two events in a linear sequence. For the pairs where one item was in the past and the other was in the future, the timezone accuracy score is equivalent to linear ordering success: if children correctly placed the past item in the past timezone and the future item in the future timezone, then they also by definition provided the correct linear order of the items. This analysis asks whether children differentiated pairs with respect to their linear ordering within a timezone: were children better (or worse) able to sequence two past items compared to two future items? The linear sequence was coded as being correct based on whether the linear order of the items was appropriate (e.g., "When dinosaurs were on earth" should precede "When you wore diapers"), but did not consider the actual distance between the items. A repeated measures ANOVA was conducted with the independent variables of age group and time-pairings (Both Past vs Both Future) and the dependent measure of linear sequencing accuracy. Results showed no effect or interaction with age group, but a significant effect of the time-pairing ($F(1,32) = 7.72$, $p < .01$, partial eta^2 = .194). Children were more accurate in ordering two past items ($M = .76$) than two future items (.41). Children were better able to make precise judgments about the past than the future.

The final analysis asked how the temporal remoteness of the items influenced children's performance. A modified version of the initial omnibus ANOVA was conducted. This version also used the independent variables of age group and type of cue (Explicit Linguistic vs. General Event Description) and the dependent variable of timezone accuracy. However, the three remoteness levels were also used as independent variables (Very Close, Close, and Far); moreover, as these levels were only defined for the past and future times, only those times were included as levels in the analysis. In addition, plain tenses (which lack any remoteness value), were excluded from the analysis[1]. Not surprisingly, this analysis reproduced the previously noted main effects of age group (Older children out-performed Younger children) and the interaction between cue type and timezone resulting from the difference between the cue types in the future timezone. In addition, this analysis found a main effect for temporal remoteness ($F(2, 31) = 7.67$, $p = .002$ partial eta^2 = .33) as well as an interaction between remoteness and timezone ($F(2, 31) = 5.0$, $p = .013$ partial eta^2 = .24) and a three-way interaction between remoteness, timezone, and cue type ($F(2, 31) = 4.07$ $p = .027$ partial eta^2 = .21). Overall, the most distal (Far) items were the most accurate, and children performed significantly better with them than either the Close items ($t(33) = 2.3$, $p = .026$) or the Very Close items ($t(33) = 4.7$, $p < .001$). However, as the interactions suggest, this effect was driven in large part by the fact that two specific cases were significantly lower than the others (paired t-test p-values ranged from .001 to 032): children were significantly

1. Paired t-tests were conducted to see how the plain tense items compared to the linguistic items with more precise temporal specification. The unspecified items received an accuracy score of .68 which was not statistically different for either age group from any level of specified linguistic remoteness.

less accurate with the General Event Descriptions for Close Future times (e.g., when you go to bed) and with Explicit Linguistic items for Very Close Past times (e.g., *She just did it*) although those two items were not significantly different from each other ($t(33) = 0.3$, $p = .7$). Although these conditions were the primary drivers of significance, it is worth noting that the overall pattern within both timezones and with both cue types is for increasing accuracy with increasing distance from the present time.

2.6 Discussion

These children showed significant improvement with age: Relative to the 5-year-olds, the 8-year-olds showed more precise use of the timeline (as seen in the higher correlation between their picture placements and the objectively true times), and they were more likely to place items in the correct overall timezone (past, present, or future). This improvement, however, should be seen in the context of 5-year-olds' overall high rates of success in using the correct overall timezone and significant positive correlation with the objective times of the events. More importantly, it should be seen in the context of the complete absence of interactions of age with the other factors: the influences of the deictic center, event remoteness, and differential ordering in the past and future were qualitatively the same across the age groups. Consistent with the TPH, children may get better at the task, but the influences of the temporal components are already in place for the younger age group.

The data also support several other predictions of the TPH. The children were influenced by the deictic center, as seen by their ability to correctly locate events with respect to it, and by the fact that they were better able to locate pairs of events when the items crossed the deictic center compared to when both were on the same side of it. Children also demonstrated an asymmetry in their past and future understanding, as seen by the fact that they were more likely to get the linear order correct for two paired items in the past than in the future timezone. Most strikingly, even though the children were acquiring a language (English) that does not grammaticize remoteness, the children nevertheless were influenced by how remote the event was from the deictic center. Specifically, the children were more accurate in locating items in the correct overall timezone when the item was linked to a time far from the deictic center than when it was linked to times close or very close to the center. Thus, even though the English grammatical tense system treats near and far past events as equivalent (Mary wore a hat yesterday/1000 years ago) these children – like many other languages of the world – did not.

One piece of data that suggests that temporal organization might be influenced by something more along the lines of the "thinking for speaking" approach is the fact that children's accuracy in locating future events in the correct general timezone was better with explicit linguistic temporal adverbials than without them. It is interesting to note, though, that language appears to help in organizing a dimension of time which the TPH suggests would be less well defined through cognitive biases.

3. Experiment 2

One potential concern with Experiment 1 is that a limited number of items were used and the effects may depend on those specific words and events rather than the more general features they embodied. Experiment 2 is a conceptual replication of Experiment 1. It uses the same task and procedures, and even a few of the same items. However, it also uses all new Explicit Linguistic items, which will help determine if children's abilities are linked to the more abstract features of the items. Moreover, in the General Event Description phase, in addition to using a few new items, it also re-arranges how the paired event descriptions were matched up. Experiment 1 found an advantage for pairs in which the items were on different sides of the deictic center, but it is also possible that those pairs simply involved particularly easy items to locate. Every paired item that was in a highly successful configuration from Experiment 1 has been switched to be in a configuration that was found to be more difficult in that experiment. Again, the intention is to insure that it is the abstract temporal configurations that matter and not the specific items.

3.1 Participants

Participants were tested in a local science center. The Younger group consisted of 22 children (*M* = 5;6, 13 girls) and the Older group consisted of 13 children (*M* = 8;5, 9 girls). All children had English as their sole or dominant language.

3.2 Stimuli and procedures

The procedures used in this study were identical to those used in Experiment 1. The stimuli consisted of the same types of items, but as noted, new specific linguistic terms were used (see Figure 1) and new arrangements of the events pairs were used.

3.3 Results

The same analyses were conducted with these data. Figure 2 shows the correlation between children's placements and the objectively true temporal ordering of the events. As with Experiment 1, significant positive correlations were found between these two measures for both the Younger children (Pearson's $r = .23$, $N = 440$, $p < .001$) and the Older children (Pearson's $r = .65$, $N = 260$, $p < .001$).

A repeated measures ANOVA was conducted using age group as a between-subjects variable, cue phase (Explicit Linguistic vs. General Event Descriptions) and timezone (Past, Present, and Future) as the independent variables, and proportion correct as the dependent variable. A main effect was found for age group ($F(1, 33) = 13.60$, $p < .001$, partial eta$^2 = .29$) as the Older children continued to out-perform the Younger children in general; there was also a main effect for cue type ($F(1, 33) = 18.27$, $p < .001$, partial eta$^2 = .36$) with children doing better overall with the General Event Descriptions than the Explicit Linguistic cues. The only significant interaction was the full 3-way interaction of all the variables ($F(1, 33) = 3.88$, $p = .031$, partial eta$^2 = .20$) which was driven by the fact that the Younger children did unaccountably badly with the General Event Description item involving the Present time while the Older children achieved a perfect score with this item. This effect does not have a principled explanation (the identical item was used in Experiment 1 and did not cause problems) and appears to be a local anomaly. It is worth noting that when the present time items are omitted from this analysis, the main effect for cue phase remains (children overall did better with the General Event Descriptions) but the interaction does not.

The analysis of the paired event descriptions yielded somewhat different effects from Experiment 1. A repeated measures ANOVA with time pairing (Both Past, Both Future, and One Past & One Future) found a main effect for age group ($F(1, 33) = 13.00$ $p < .001$, partial eta$^2 = .28$) and time pairing ($F(2, 32) = 5.21$, $p < .011$, partial eta$^2 = .25$) and an interaction between the two ($F(2, 32) = 6.17$, $p < .005$, partial eta$^2 = .28$). Follow-up analyses showed that Older children succeeded with all time pairings (means ranged from .85 to .88 correct) but that Younger children – similar to Experiment 1 – were more accurate with pairs that crossed the deictic center relative to pairs where both items were in the past ($t(21) = 5.1$, $p < .001$) or in the future ($t(21) = 4.1$, $p < .001$). However, unlike in the last study, there was no difference in how accurately the children linearly ordered the items in the past and future times for either the Younger children ($t(21) = 0.9$, $p = .38$.) or the Older children ($t(12) = 0.81$, $p = .44$.). Comparing across the two studies, it appears that children did equivalently well (or rather, badly) at ordering the all-future items ($t(67) = .14$, $p = .71$.) but their performance dropped in Experiment 2 for ordering the all-past items ($t(67) = 8.89$, $p = .004$).

The final analysis incorporated remoteness information. As with Experiment 1, present time items were omitted as were linguistic items with no specified temporal distance (e.g., *before, after*)[2]. Because only two remoteness levels (Very Close

2. As with the previous study, these unspecified linguistic items received intermediate accuracy scores which did not statistically differentiate them from any of the specified remoteness levels.

and Close) were used in both the cue conditions the omnibus repeated measures ANOVA used just those two levels of remoteness, along with the independent variables of cue type (Explicit Linguistic vs. General Event Descriptions, timezone (Past vs. Future) and age group; the dependent variable was timezone accuracy. This analysis found a main effect for age group ($F(1, 33) = 8.25$, $p = .007$, partial eta$^2 = .20$) and several interactions with age group all of which stemmed from the fact that Older children performed almost at ceiling and all the condition differences depended on Younger children. Therefore, a second ANOVA looking just at the Younger age group found these children were more accurate with the General Event Descriptions than the Explicit Linguistic items ($F(1, 21) = 16.30$, $p = .017$, partial eta$^2 = .43$). No other main effects were significant but there was a significant interaction between cue type and remoteness ($F(1, 21) = 12.05$ $p = .002$, partial eta$^2 = .37$) and between time and remoteness ($F(1, 21) = 5.50$, $p = .03$, partial eta$^2 = .21$). Inspection of the data suggested that these interactions arose because the Younger children did particularly well with the Very Close Future items in the General Event Descriptions, and this intuition was confirmed by post-hoc t-tests (p's ranged from .001 to .06). In addition, a further ANOVA conducted just within the General Event Descriptions found that the most remote (Far) items in this condition patterned consistently with the intermediately remote (Close) items across the board.

3.4 Discussion

With the change of specific items, Experiment 2 was able to replicate some – but not all – of the results from Experiment 1. One overall general difference between the studies was the fact that the 5-year-olds failed to perform above chance on many of the linguistic items and thus performed worse overall compared to Experiment 1; similar declines were not seen with the general event descriptions, however. With respect to the TPH, one of the new results provides extra support. In Experiment 2, 5-year-olds again located pairs of events better when the deictic center was between them, but the 8-year-olds no longer showed that effect (they were effectively at ceiling). The fact that the influence of a potential cognitive bias (the desire to organize time around a deictic center) can decline with age supports the notion that the bias guides the learning process and is not the outcome of the process. By contrast, unlike the previous experiment, Experiment 2 found no asymmetry between the past and future and no effects of remoteness. Given that remoteness is the best test of the TPH because it is the only temporal feature that is not grammaticized in English, the lack of effect speaks against the hypothesis. Indeed, these results suggest that children's ability to interpret remoteness is more ad hoc and dependent on the specific arrangement of events and linguistic items used.

4. General discussion

The TPH states that common grammatical patterns across languages arise fundamentally from the same biases in cognition that guide early language development. Cross-linguistic studies of linguistic typology can thus provide working hypotheses about what concepts children may use, regardless of the specific elements in their native language. Three typologically prevalent temporal elements were investigated using a timeline task: the use of the deictic center to organize time into a past, present, and future; an asymmetry between the past and future times; and the use of temporal remoteness from the present moment as an organizing metric. The first of these elements is clearly used to organize the English temporal system, the second is potentially instantiated in English, and the third is not grammaticized in English.

The results showed that children as young as five years old can appropriately use a timeline and understand a range of temporal adverbs and general event descriptions. However, their performance on the task provided only limited support for the TPH. The children showed strong understanding that time was deictically organized – they correctly placed items in the past, present, and future times, and their ability to linearly order pairs of events was improved when the deictic center could be used as an intervening anchor point. However, such results are weak evidence for the TPH as the use of a deictic center is not only common typologically, but it is also a part of the grammatical system of the language the children are learning. For the typological patterns that are not as fully grammaticized in English (the asymmetry between past and present) or not grammaticized at all (remoteness), the data was far more ambiguous: Experiment 1 found support for the idea that children's past and future representations are asymmetric (with past being better understood) and also that their ability to correctly place events into the past, present or future timezones was influenced by temporal remoteness (with more distal items being more accurately located); however, Experiment 2 failed to replicate these latter two findings. Thus, it is still possible that children's understanding of English is influencing their performance, but less likely that non-English patterns of organization – however typologically prevalent – are guiding their understanding.

This failure to replicate for the non-English elements lends support to the alternative "thinking for speaking" position because it shows that children are being influenced by their own native language's organization (and not by other common means of organizing time in language). Moreover, this point of view is also supported by another core dimension of these results, namely that specifics matter. The differences between the two experiments were all in the specific items – slightly different adverbials and different combinations of specific items within the task. While these small changes did not influence the importance of the English instantiated dimension of using a deictic center, they did change the influence of the non-English related temporal dimensions on locating the events. The acquisition of specific adverbs is a protracted process (see Tillman & Barner, 2015) and as noted previously, these items have an intermediate status with respect to the grammar. For these items to develop the habits necessary to guide "thinking for speaking" would take time and even so, might develop on a word-by-word basis as the different adverbials differ in terms of complexity and frequency. Thus, the "thinking for speaking" alternative predicts that the specific adverbials should matter in this task.

However, the "thinking for speaking" position does not predict the pattern of results found in Experiment 1 nor does it explain the typological prevalence of an asymmetry between the past and future tenses nor the use of remoteness as a temporal organizer more generally. Perhaps a third alternative is in order. Cognitive biases can be conceptualized in many different ways and it is possible that a more computationally oriented perspective would be helpful. Temporal elements which are typologically common may not arise because of some a priori conceptual bias to organize time around those dimensions but may instead reflect local maxima in any system designed to express time. As such, these local maxima might influence children's temporal organization but their effects would be expected to be fleeting and fragile, as they would need to compete with a strong external signal, namely the actual temporal organization of the child's language. These ideas are speculative, of course, but they suggest fruitful lines of ongoing inquiry.

There is an old story about a drunk looking for her keys on a dark street underneath a streetlamp. When asked why she is looking there (of all places) she replies "Because that's where the light is." One of the hardest aspects of studying the relationship between cognitive biases and language development is identifying elements that are not in the direct spotlight of the child's target language. The TPH is a way of adding a streetlamp that can widen our hypothesis space. The current data provided equivocal evidence that the keys to understanding the cognition that undergirds temporal semantics are actually lying in this particular patch of light, but the general approach has expanded the possible places we could look for them.

Acknowledgements

These data were collected at the Center of Science and Industry (COSI) in Columbus OH in the little kidspace® area and in the Buckeye Language Network's Language Sciences Research Lab. My thanks to COSI for their support and enthusiasm for allowing me to conduct research on-site. Thanks go to the research assistants who created the stimuli, ran the participants, and coded the data – especially Kerrianne Morrison and Sonam Jindal.

References

Binnick, R. I. (1991). *Time and the verb: A guide to tense and aspect*. Oxford: Oxford University Press.

Binnick, R. I. (Ed.). (2012). *The Oxford handbook of tense and aspect*. Oxford: Oxford University Press. https://doi.org/10.1093/oxfordhb/9780195381979.001.0001

Botne, R. (2012). Remoteness distinctions. In R. I. Binnick (Ed.), 536–562.

Brown, R. (1973). *First language*. Cambridge, MA: Harvard University Press. https://doi.org/10.4159/harvard.9780674732469

Busby Grant, J., & Suddendorf, T. (2009). Preschoolers begin to differentiate the times of events from throughout the lifespan. *European Journal of Developmental Psychology*, 6, 746–762. https://doi.org/10.1080/17405620802102947

Busby Grant, J., & Suddendorf, T. (2011). Production of temporal terms by 3-, 4-, and 5-year-old children. *Early Childhood Research Quarterly*, 26, 87–95. https://doi.org/10.1016/j.ecresq.2010.05.002

Comrie, B. (1985). *Tense*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9781139165815

Dahl, Ö. (1985). *Tense and aspect systems*. Oxford: Basil Blackwell.

Deo, A. (2012). Morphology. In R. I. Binnick (Ed.), 155–183. https://doi.org/10.1093/oxfordhb/9780195381979.013.0005

Friedman W. J. (2002). Children's knowledge of the future distances of daily activities and annual events. *Cognition and Development*, 3, 333–356. https://doi.org/10.1207/S15327647JCD0303_4

Friedman W. J., Gardner A. G., & Zubin N. R. E. (1995). Children's comparisons of the recency of two events from the past year. *Child Development*, 66, 970–983 https://doi.org/10.2307/1131792

Gentner, D., & Bowerman, M. (2009). Why some spatial semantic categories are harder to learn than others: The typological prevalence hypothesis. In J. Guo, E. Lieven, N. Budwig, S. Ervin-Tripp, K. Nakamura, & S. Ozcaliskan (Eds.), *Crosslinguistic approaches to the psychology of language* (pp. 465–480). London: Taylor & Francis.

Klein, W. (1994). *Time in language*. New York, NY: Routledge.

Ludlow, P. (1999). *Semantics, tense, and time*. Cambridge, MA: The MIT Press.

Maastricht, L., & Hollebrandse, B. (2011). Tense acquisition in French 5-year-olds. *Groninger Abeiten zur Germanistischen Linguistik*, 51, 75–86.

McCormack, T., & Hanley, M. (2011). Children's reasoning about the temporal order of past and future events. *Cognitive Development* 26, 299–314. https://doi.org/10.1016/j.cogdev.2011.10.001

Poeppel, D., & Wexler, K. (1993). The full competence hypothesis of clause structure in early German. *Language*, 69, 1–33. https://doi.org/10.2307/416414

Piaget, J. (1969). *The child's conception of time*. English translation by A. J. Pomerans. New York, NY: Basic Books.

Reichenbach, H. (1947). *Elements of symbolic logic*. New York, NY: McMillan.

Shatz, M., Tare M., Nguyen S. P., & Young, T. (2010). Acquiring non-object terms: The case for time words. *Journal of Cognitive Development*, 11, 16–36. https://doi.org/10.1080/15248370903453568

Slobin, D. I. (1985). Crosslinguistic evidence for the language-making capacity. In D. I. Slobin (Ed.), *The crosslinguistic study of language acquisition, Vol 2: Theoretical issues* (pp. 265–322). Mahwah NJ: Lawrence Erlbaum Associates.

Slobin, D. I. (1996). From "thought and language" to "thinking for speaking.". In J. J. Gumperz & S. C. Levinson (Eds.), *Rethinking linguistic relativity* (pp. 70–96). Cambridge: Cambridge University Press.
Suddendorf, T. (2010). Episodic memory versus episodic foresight: Similarities and differences. *Wiley Interdisciplinary Reviews in Cognitive Science*, 1, 99–107. https://doi.org/10.1002/wcs.23
Tillman, K. A., & Barner, D. (2015). Learning the language of time: Children's acquisition of duration words. *Cognitive Psychology*, 78, 57–77. https://doi.org/10.1016/j.cogpsych.2015.03.001
Valian, V. (2006). Young children's understanding of present and past tense. *Language Learning and Development*, 2, 251–78. https://doi.org/10.1207/s15473341lld0204_2
Wagner, L. (2001). Aspectual influences on early tense interpretation. *Journal of Child Language*, 28, 661–81. https://doi.org/10.1017/S0305000901004792
Wagner, L. (2012). Primary language acquisition. In R. I. Binnick (Ed.), 458–480.
Weist, R. M., Wysocka, H., & Lyytinen, P. (1991). A cross-linguistic perspective on the development of temporal systems. *Journal of Child Language*, 18, 67–92. https://doi.org/10.1017/S0305000900013301
Weist, R. M., Atanassova, M., Wysocka, H., & Pawlak, A. (1999). Spatial and temporal systems in child language and thought: A cross-linguistic study. *First Language*, 19, 267–311. https://doi.org/10.1177/014272379901905701

Appendix of supplementary resources

There are many resources available to learn more about the typological variety of the world's languages – for temporal systems as well as other dimensions of language. The links below provide a good starting place:

Typological Tools for Field Linguists: The Tense, Moods and Aspect Questionnaire
From the Department of Linguistics at the Max Planck Institute for Evolutionary Anthropology:
<https://www.eva.mpg.de/lingua/tools-at-lingboard/questionnaire/tma_description.php>

The World Atlas of Language Structure
(Search for "tense" as a name under the Features tab to see temporal language information)
<http://wals.info/>

Ethnologue Website
(Search under the Languages tab to see the languages described)
<https://www.ethnologue.com>

CHAPTER 4

Semantic features of early verb vocabularies

Sabrina Horvath[i], Leslie Rescorla[ii] and Sudha Arunachalam[i]
[i]Boston University / [ii]Bryn Mawr College

Much research has asked why verbs are difficult to acquire, and how toddlers nevertheless acquire them. Still, we know little about what *kinds* of verb meanings are easy or difficult to acquire. We revisit Rescorla and colleagues' data on vocabulary knowledge in toddlers acquiring English, Italian, Greek, Korean, and Portuguese measured using the Language Development Survey. We coded the survey's verbs for several semantic features to determine which features predict appearance in toddlers' vocabularies. For English, manner and result verbs were equally well known across samples, but verbs labeling durative events and events with fewer event participants were more likely to be known than those labeling punctual events and events involving more participants. Similar trends held in the other languages.

Keywords: vocabulary composition, verbs, lexical semantics, verb learning, cross-linguistic, language development survey

1. Introduction

Acquiring the meanings of verbs is a notoriously difficult challenge for young learners for several reasons, including the complexity of the meanings they encode (e.g., Gentner, 1978) and the difficulty of identifying the intended referent simply by observing the world when a new verb is uttered (e.g., Gleitman, 1990). However, the trajectory of verb acquisition appears far from arbitrary; young children typically acquire similar verbs early on (e.g., Naigles et al., 2009). Several factors are likely to affect whether a verb will be easy or difficult to acquire, such as its frequency in the input (Goodman, Dale, & Li, 2008), the frequency and diversity of contexts in which it appears (e.g., Naigles & Hoff-Ginsberg, 1998), and the imageability of the concept it labels (e.g., Gillette, Gleitman, Gleitman & Lederer, 1999; Ma, Golinkoff, Hirsh-Pasek, McDonough & Tardif, 2009).

Here, we focus in more detail on the semantic properties of verbs and the concepts they denote as potential indicators of how likely a verb is to occur in the

https://doi.org/10.1075/tilar.24.04hor

early lexicon. For example, we might predict that verbs denoting fleeting events are more difficult to acquire than verbs denoting long-lasting events, as there is simply less time to observe them. As yet there are no systematic studies of which kinds of meaning are easiest to acquire, nor do we know how universal such patterns might be across typologically different languages. Here, we take a first pass at these questions by investigating 2-year-old toddlers' verb vocabularies in a large sample of English-acquiring toddlers as well as toddlers acquiring Greek, Italian, Korean, and Portuguese to determine whether certain semantic features are more widely represented than others. Our data set comes from Rescorla's (1989) Language Development Survey, a parent report vocabulary checklist adapted for each of these languages.

In addition to analysis of parent reports of vocabulary, methods used to understand the composition of toddlers' early lexicons include experimental novel verb learning studies and diary studies. For the youngest toddlers, many studies have asked about the difficulty of acquiring verbs as compared to nouns, and whether such difficulty is universal or language-specific (e.g., Bornstein et al., 2004; Imai et al., 2008; Tardif, 1996; Tardif, Gelman, & Xu, 1999; Waxman et al., 2013). One factor invoked in such discussions is the different kinds of concepts to which (early-acquired) nouns and verbs typically refer; for example, Gentner (1982) argued that nouns are easier to learn because they often refer to concrete entities, whereas verbs often refer to relational concepts. But even within a grammatical category, different types of concepts are represented. For example, verbs can label states as well as actions, and even action verbs can differ in a variety of meaning-related features. For this reason we think it necessary to look at different types of verb meaning rather than only grammatical distinctions between nouns and verbs (e.g., Maratsos, 1990).

Some semantic properties of early verbs have been studied in prior work, albeit in a scattered way. For example, imageability of the concept being labeled relates to ease of verb acquisition (e.g., Gillette et al., 1999; Ma et al., 2009), and verbs describing mental states may be particularly difficult to acquire for this reason (e.g., Papafragou, Cassidy, & Gleitman, 2007). Similarly, concrete action verbs may be more easily acquired than verbs labeling abstract concepts (e.g., Bassano, 2000). Novel verb learning studies, too, have used a variety of verb and event types, but different types are rarely contrasted within a single study (Naigles & Kako, 1993 is an important exception, as is Scott, Gertner, & Fisher, this volume). Perhaps the best studied semantic notion with respect to ease of acquisition in English-acquiring toddlers is whether a verb encodes a manner or a result of action. To date, however, this literature is conflicted, as we will see below.

Central to our approach is the inclusion of data from a variety of languages. Despite linguistic and cultural differences, studies examining the composition of the lexicon, often using a vocabulary checklist, have found similarities in the words and word types that are early acquired across languages (e.g., Bornstein et al., 2004; Dale & Goodman, 2005). However, when it comes to specific verb types, we know little about the extent to which there are cross-linguistic similarities or differences. A notable exception to this is in the domain of manner and path verbs; this distinction has garnered much interest in the literature and has been the focus of much cross-linguistic research (e.g., Allen et al., 2007; Berman & Slobin, 1994). In the case of manner and path, because languages have a strong tendency to lexicalize one or the other, the expected outcome (and indeed the consistent finding) of such studies is that children have a bias corresponding to the bias evident in their language. In the current study, we examine dimensions that are *not* manifested in wholly or systematically different ways in the typologically different languages under investigation to ask whether within each language, some verb types appear in greater numbers than others, and whether such patterns are consistent across languages.

The specific semantic dimensions we investigate are: whether the verb encodes a manner or a result (e.g., *clap* vs. *close*); whether the verb describes events that are typically durative or punctual (e.g., *read* vs. *cough*); and the number of (semantic) event participants involved in the event the verb denotes (e.g., 1: *nap*, 2: *eat*, 3: *bring*). These distinctions are expected to provide insight into what kinds of semantic representations are easy to form.

1.1 Our dataset: Data from the language development survey (Rescorla, 1989)

Rescorla's (1989) Language Development Survey (LDS) is a parent-report questionnaire of expressive vocabulary development designed to screen for language delay. Parents complete a 310-word checklist of expressive vocabulary, in addition to information about the toddler's family and medical history. The LDS has been shown to be a reliable and valid instrument (Rescorla, 1989). Normed on a U.S. national probability sample of 274 toddlers (Rescorla & Achenbach, 2000), the checklist has been adapted for use in several other countries.

For this study, we revisited the data collected by Rescorla and her colleagues, focusing on 45 verbs in the checklist. In addition to the English data, we analyzed data from five other language samples, reported originally for Greek in Papaeliou and Rescorla (2011), for Italian from two different regions – North Italy and Rome – in Rescorla, Frigerio, Sali, Spataro, & Longobardi (2014), for Korean in Rescorla, Lee, Oh, & Kim (2013), and for Portuguese in Rescorla, Nyame, and Dias (2016). In all cases, Rescorla and her team carefully translated and adapted the LDS to be culturally and linguistically appropriate. Overall, they noted considerable similarity in their U.S. and other samples with respect to number and types of words known by toddlers, setting the stage to ask more fine-grained questions about whether the same is true for verbs specifically and for specific semantic categories of verbs.

We chose to use data from the LDS for two reasons. First, the survey methods allowed for a much larger data sample than can be typically achieved through experimental paradigms or diary studies. Second, the survey questionnaire allowed us to study what children are doing naturally in the process of language development, as opposed to what they do in laboratory contexts.

2. Methods

2.1 Participants

The U.S. sample included 274 toddlers (133 males, 141 females) ranging in age from 18 to 35 months (mean age 26 months, SD 5 months). The toddlers were further grouped by age. The youngest group ranged from 18 to 23 months, the next from 24 to 29 months, and the oldest from 30 to 35 months. For detailed sample information, see Rescorla and Achenbach (2000). On average, toddlers knew 53% of the verbs on the checklist, but this varied widely across the age range, with many of the youngest toddlers knowing none and many of the oldest knowing all of the verbs.

The sample sizes and mean age for all language groups are listed in Table 1. All toddlers were ages 18 to 35 months of age, with the exception of two from the Korean sample, one who was 16 months and the other 17 months. We excluded a small number of toddlers from the Greek sample who were over 35 months of age.

Table 1. Sample size and mean age

	English	Greek	Italian (North Italy)	Italian (Rome)	Korean	Portuguese
Total	274	260	324	175	2191	198
Male	131	127	168	93	1162	114
Female	141	133	156	82	1029	84
Youngest age group (18 to 23 months)	101	40	86	48	807	49
Middle group (24 to 29 months)	90	78	126	60	831	70
Oldest group (30 to 35 months)	83	142	112	87	553	79
Mean age (months)	26.3	29.1	27.0	27.3	26.0	27.2

2.2 Verbs

Though we used the LDS checklist's own designation of "action" words to determine which words to include in our list of verbs, we excluded "dinner," "down," "lunch," "outside," "pattycake," "peekaboo," and "up" because they are not realized as verbs in adult speech. Note that the "action" words category of the checklist includes stative as well as eventive verbs. There were words in other sections of the checklist that in English have identical phonological forms for both nouns and verbs (e.g., snow, swing). We excluded these from our analyses because the forms for these in some of the other languages in our sample contained unambiguous morphological cues signaling their grammatical category. We only studied verbs that appeared on all languages' checklists.

2.3 Semantic dimensions

We coded each verb on three dimensions: whether it encodes a manner or result, whether the event it labels is punctual or durative, and the number of event participants typically associated with its referent.

2.3.1 *Manner vs. result*

The distinction between manner and result verbs has garnered significant attention in the study of verb acquisition. However, the results of these studies do not clearly indicate a single trajectory. Some evidence suggests that toddlers' early productions demonstrate a bias for result meanings. Even at the one-word stage, the result components of events appear to be salient for toddlers. Across languages, early words express meanings like "all gone" or "all done" (e.g., Gopnik & Meltzoff, 1986; Behrens, 1993; de Lemos, 1981) and particles like "up" are used to express changes of state (e.g., Clark, 1995; Clark, Carpenter, & Deutsch, 1995; Greenfield & Smith, 1976; Penner, Schulz, & Wymann, 2003; Tomasello, 1992). Once toddlers begin to use verbs to express state changes, result verbs are commonly attested. Huttenlocher, Smiley, and Charney (1983), for example, found that a majority of the verbs used by 2-year-olds were result verbs.

On the other hand, Gentner (1978) and Gropen, Pinker, Hollander, & Goldberg (1991) reported a manner bias; they found that children showed better command of the meanings of manner verbs than result verbs, often mistakenly incorporating a manner meaning into their representations for result verbs (e.g., construing *fill* as entailing *pour*). Similarly, Bowerman (1982) found more argument structure errors with result verbs than manner verbs. One explanation for this is that children had more difficulty encoding appropriate meanings for the result verbs. Forbes and

Poulin-Dubois (1997) found that toddlers under 2 years of age view manner as crucial to the meanings of familiar verbs; they were reluctant to extend familiar verbs – even those with resultive meanings, like "pick up" – to situations involving a different manner than their first training exemplar.

Novel verb learning studies testing extension of novel verb meanings to new situations have had mixed results. Behrend (1990) presented 3-year-olds with novel verbs in the context of a scene depicting both a manner and a change of state. When asked to extend the verbs to new events, children were more likely to extend them to actions with a different manner of motion from the original than a different result state, suggesting that they construed the result state as critical to verb meaning. However, when asked to name actions, the 3-year-olds were more likely to choose a verb encoding the manner of the action rather than result when both were appropriate. In contrast to Behrend's (1990) findings, Forbes and Farrar (1995) found that both manner and result meanings could be extended in a novel verb learning task, dependent on the variability (or lack thereof) among multiple learning opportunities. One important consideration for novel verb studies is whether children find particular kinds of meaning to be better candidates for what is being labeled by an unfamiliar verb. Choi and Arunachalam (2013) argued that, despite the fact that Korean more often labels path in lexical verbs and manner in satellite phrases (e.g., Choi & Bowerman, 1991), young Korean learners nevertheless preferentially assign manner meanings to novel verbs because it is easier to imagine a new manner of action than a new path of movement. A similar situation may hold for manner vs. result meanings, because although there may be many ways of doing something (manners), the number of possible outcomes (results) is limited.

Given the conflicting literature, it is unclear whether children have manner or result biases in early vocabulary development. Our approach in the current study offers a new perspective to this issue. We analyze data from a large data set – a large number of participants, and a large number of (familiar) words in each of several languages – and we ask about toddlers' existing productive vocabularies rather than examining the vagaries of a novel word learning context, in which it can be extremely difficult to equate the salience or naturalness of the manner and result components being depicted.

To pursue this, we coded the verbs on the LDS for whether they lexicalize manner or result. Fortunately for the purposes of this investigation, the manner vs. result distinction is well studied, particularly by Levin and Rappaport Hovav (e.g., 1991, 1995, 2013), who have noted a number of reflexes of manner vs. result encoding in argument realization patterns such as the appearance of manner, but not result verbs, in the conative alternation (e.g. *John hit at the wall* / **John broke at the wall*) (see also Fillmore, 1970). They have also argued that manner and result are in strict complementarity; that is, a verb can encode only one of these two components (but see Beavers & Koontz-Garboden, 2012; Husband, 2011; Rissman, 2015). For our coding, we used diagnostics in Beavers and Koontz-Garboden (2012) (e.g., if it is possible to deny the result of an action, the verb must be a manner verb, as with, *I* walked, *but nothing happened*, but not with *#I* closed *the door, but nothing happened*). We excluded stative verbs and verbs whose semantics we considered unclear on this basis, or whose semantics have been explicitly discussed in the literature and argued to be polysemous – notably, "cut" (Levin & Rappaport Hovav, 2010, 2013).

Although languages may differ in their specific representations of apparently translation-equivalent verbs, for present purposes we extend the coding for English to these other languages as an approximation. We await the development of language-specific diagnostics for each of the languages in our sample before refining our categories.

2.3.2 *Durative vs. punctual*

If learning the meaning of a verb requires that a child match the linguistic referent to the action it denotes, one complication of verb learning particularly is the ephemerality of many actions. For example, if a toddler hears, "Look! The boy's gonna kick the ball!" but fails to quickly orient to the soccer player, she may not witness the kicking, and may thus miss an opportunity to acquire the verb's meaning. We predict, then, that verbs describing typically durative events will be more easily acquired than verbs describing punctual events. This is, of course, a hypothesis about averages; any given verb may not fall into this pattern. For example, although *breaking* is a punctual event, the resultant change of state can be quite salient. (The verb "break" is not on the LDS, although norms for the MacArthur-Bates Communicative Development Inventory suggest that it is relatively early acquired (Dale & Fenson, 1996).) To date, we are aware of only one study that has examined this distinction between punctual and durative actions (Abbot-Smith, Imai, Durrant & Nurmsoo, 2017), which found that children have difficulty learning verbs that describe punctual events.

We coded the verbs on the LDS as "durative" or "punctual," again omitting verbs that were stative or difficult to code along this dimension. Punctual verbs are either incompatible with or receive iterative interpretations when they occur with temporal expressions that denote a protracted duration (e.g., *#The glass broke for two hours*). Note that the durative vs. punctual distinction is orthogonal to the manner vs. result distinction; for example, semelfactive verbs like *clap*, of which the LDS has four, encode manner but reference an event that is temporally punctual (Comrie, 1976; Smith, 1991).

2.3.3 *Number of obligatory event participants*

The complexity of an event may be related to the number of event participants it necessarily involves. Events like *sleeping* require only one participant, while events like *giving* require a giver, a recipient, and a thing given. Fisher and colleagues have argued that toddlers initially map the number of entities they hear named in an utterance to the number of event participants in an event (e.g., Fisher, 1996; Gertner & Fisher, 2012; Yuan, Fisher, & Snedeker, 2012), indicating that the number of event participants is an important cue to which toddlers attend. We hypothesized that the more event participants required, the more difficult the verb would be to acquire. This is because it may be difficult for learners to identify which particular event among all those each event participant is involved in, or which participant's perspective on the event, is being encoded (Gleitman, 1990).

For this dimension, we coded the smallest number of event participants that could occur in the event described by the verb. For example, the verb *hit* obligatorily has an agent and a patient/theme, as in "The girl *hit* her teddy bear," but may also have an optional instrument, as in "The girl *hit* the ball with her racket." Here, we code *hit* as having two obligatory event participants. The distinction is similar to the number of syntactic arguments, except that we included instruments (e.g., for *cut*, which may be realized in a transitive sentence such as "I *cut* the paper," but typically requires an agent, patient/theme, and instrument) and implicit objects (e.g., for *eat*, which may be realized in an intransitive sentence but necessarily requires an agent and patient/theme).

2.4 Covariates

2.4.1 *Frequency*

We expect toddlers to more easily acquire verbs that are highly frequent in their input. To study the importance of the coded semantic dimensions while acknowledging the importance of input frequency, we included it as a factor in our analyses for English, but we lacked rich coded corpus data for the other languages. For English, we used Li's (2001) frequency counts of the speech uttered by parents, caregivers, and experimenters in the corpora in the CHILDES database (Li & Shirai, 2000; MacWhinney, 2000). We included all forms of the verb (e.g., infinitive, perfective, progressive); noun forms (e.g., *a kiss*) were also included because the tokens were not categorized by grammatical category. This yielded a large range of frequencies, from 64 (for *clap*) to 36,581 (for *go*). We grouped these as follows: low frequency (< 800 occurrences, e.g., *knock*), mid frequency (800–2,000 occurrences, e.g., *close*), and high frequency (> 2,000 occurrences, e.g., *give*). Table 2 lists the verbs in each group.

Table 2. Frequency groupings, based on the CHILDES parental corpus

Low frequency (fewer than 800)	Medium frequency (800 to 2000)	High frequency (more than 2000)
catch	close	bring
clap	cut	come
cough	fix	eat
dance	hit	finish
feed	love	get
hug	open	give
jump	push	go
kick	read	have
kiss	ride	help
knock	run	look
nap	sing	make
shut	sleep	see
tickle	stop	show
	throw	sit
	walk	take
	wash	want

2.4.2 *Imageability*

Given the role attributed to imageability in prior work on verb acquisition (e.g., Gillette et al., 1999; Ma et al., 2009) we initially planned to include imageability, using Cortese and Fugett's (2004) imageability ratings; almost all of the LDS words were included in their list. However, imageability according to these ratings and frequency – coded as reported above – were inversely correlated for the verbs on the LDS ($r = -0.74$), and preliminary analyses indicated that imageability played far less of a role than frequency (but see Snedeker, Zeitlin, & Crawford, 2013). Thus, we ultimately did not include imageability in the analyses reported below. The inverse correlation is not surprising given that many of the highest frequency verbs on the LDS are light verbs that are not very imageable, such as "get" and "make."

3. Results

First, we coded each of the verbs from the LDS for the semantic dimensions of manner vs. result, punctual vs. durative, and minimum number of event participants, using the criteria outlined in Section 2.3. Those verbs that could not be coded reliably in one of these dimensions were excluded from analysis for that particular dimension. The results of this coding are listed in Table 3.

Table 3. Semantic coding for all of the verbs that were included in the mixed-effects analyses. Dots indicate that a verb did not receive a code for a particular distinction, due to difficulty establishing the appropriate value (e.g., because of conflict in the literature)

Verb	Manner vs. result	Durative vs. punctual	Number of event participants
bring	result	punctual	3
catch	result	punctual	2
clap	manner	punctual	1
close	result	punctual	2
come	result	punctual	1
cough	manner	punctual	1
cut	.	.	3
dance	manner	durative	1
eat	result	durative	2
feed	.	durative	3
finish	result	punctual	.
fix	result	punctual	2
get	result	punctual	3
give	result	punctual	3
go	result	.	1
have	.	.	2
help	result	.	.
hit	manner	punctual	2
hug	manner	durative	2
jump	manner	punctual	1
kick	manner	punctual	1
kiss	manner	.	2
knock	manner	punctual	2
look	.	durative	2
love	.	.	2
make	result	.	2
nap	manner	durative	1
open	result	punctual	2
push	manner	.	2
read	manner	durative	2
ride	manner	durative	2
run	manner	durative	1
see	.	durative	2
show	manner	durative	3
shut	result	punctual	2
sing	manner	durative	1
sit	.	.	1
sleep	manner	durative	1
stop	result	punctual	.
take	result	punctual	3

Table 3. (*continued*)

Verb	Manner vs. result	Durative vs. punctual	Number of event participants
throw	manner	punctual	2
tickle	manner	durative	2
walk	manner	durative	1
want	.	.	.
wash	manner	durative	2

Following this, for each verb in each language, we calculated the percentage of toddlers who were reported to use it (called a *percentage use score* following Rescorla & Safyer, 2013). Note that this figure is not the same as the age at which the verb is acquired, but we would expect the two to be correlated; the earlier a verb is acquired, the more likely it is to be known across the age range represented in the sample. Table 4 lists the ten verbs with the highest percentage use score in each language. Notably, the lists are similar across languages, with verbs labeling bodily activities (*eat, sleep*) on all lists. Unsurprisingly, the two samples from Italy are the most similar to each other.

Table 4. Ten verbs known by the largest percentage of toddlers in each sample (alphabetical order)

English	Greek	Italian (North Italy)	Italian (Rome)	Korean	Portuguese
eat	eat	close	clap	eat	close
go	have	dance	close	get	come
hug	kiss	eat	dance	go	dance
kiss	love	go	eat	jump	eat
love	make	kiss	kiss	love	give
open	open	open	open	ride	open
see	read	ride	ride	see	sing
sit	run	run	run	sleep	sit
sleep	sleep	sleep	sleep	sit	sleep
stop	want	wash	wash	want	wash

Next, we asked whether toddlers' knowledge of the verbs was related to the semantic dimensions of manner vs. result, punctual vs. durative, and minimum number of event participants. In each analysis, we fit the raw data to a mixed-effects regression model (binomial family); the outcome measure was a binary measure of whether the toddler knew the verb. We included participant and verb as random factors and age group, gender, and the semantic dimension as fixed factors. Analyses were conducted using the glmer() function in R (v. 2.14.2) (Bates, Maechler & Bolker, 2012; R Development Core Team, 2012). To test significance, we used the z-test and p-values output by glmer(). For the English analyses, we also included each verb's input frequency group (as listed in Table 2), and the interaction of frequency with the semantic dimension, as fixed factors.

3.1 Manner vs. result

3.1.1 *English*

Our coding using the diagnostics in Beavers and Koontz-Garboden (2012) yielded 21 manner verbs and 16 result verbs; coding was done by the first and third authors. Interestingly, these were equally represented in the lexicons of young English learners. Across toddlers and verbs, mean knowledge of manner verbs was 52%, and of result verbs, 53%. Note that these percentages are not percentage use scores, which reveal the percentage of *toddlers* who know a particular word, but simply the average scores for the binary variable indicating whether each verb is known or not, collapsing across *all* toddlers and the 37 verbs. For example, given 274 toddlers and 37 verbs for a total of 10,138 data points consisting of 0s and 1s, we determined the average across all of these data points. We call this measure *mean percent knowledge*. Figure 1 depicts scores by frequency group. A binomial mixed-effects model with participant and verb as random factor, and gender, age group, frequency group, and whether the verb encoded manner or result as fixed factors yielded only main effects of gender ($z = 2.93$, $p < 0.005$) and age group ($z = 9.12$, $p < .001$). Here, as with all analyses in which gender and age group were significant, females knew more verbs than males and older toddlers knew more verbs than younger toddlers. However, whether a verb encoded manner or result did not contribute significantly to this

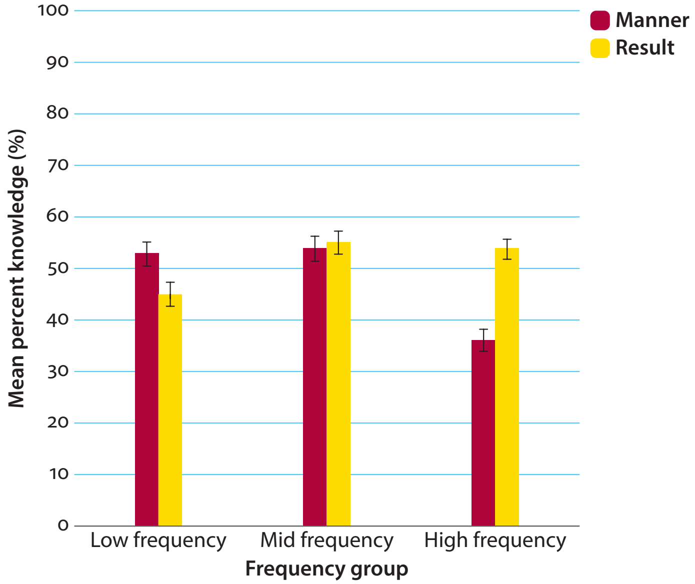


Figure 1. Mean percent knowledge of manner and result verbs by frequency group for English-acquiring toddlers. Error bars indicate standard errors of subject means

model ($z = 0.035$, $p = 0.97$), and thus we infer that the manner-result distinction does not predict verb knowledge. Model parameter estimates are in Table 5 in the online supplementary material.

This finding, despite the fact that it does not confirm most of the prior literature, some of which finds a manner bias and others of which finds a result bias, is perhaps satisfying nevertheless. The similarity in knowledge of manner and result verbs suggests that across the lexicon – as opposed to within a small group of events and verbs tested in laboratory experiments – neither component or interpretation is salient to the exclusion of the other. Of course, it may be that within any given learning situation, a bias in one direction or the other exists, but toddlers apparently manage to overcome such difficulties to acquire both kinds of verbs within the first three years of life.

3.1.2 *Other languages*

Recall that we acknowledge that the classification of English verbs as manner or result verbs carries over only imperfectly to other languages, but we used it as a rough guide in the present analysis. The balanced representation of manner and result verbs held up across languages, as evident in Figure 2. In mixed-effects models for each language as described for English, we found no main effects of manner vs. result (all $zs < 1.4$, all $ps > .15$). For the Greek sample, we found a main effect of age

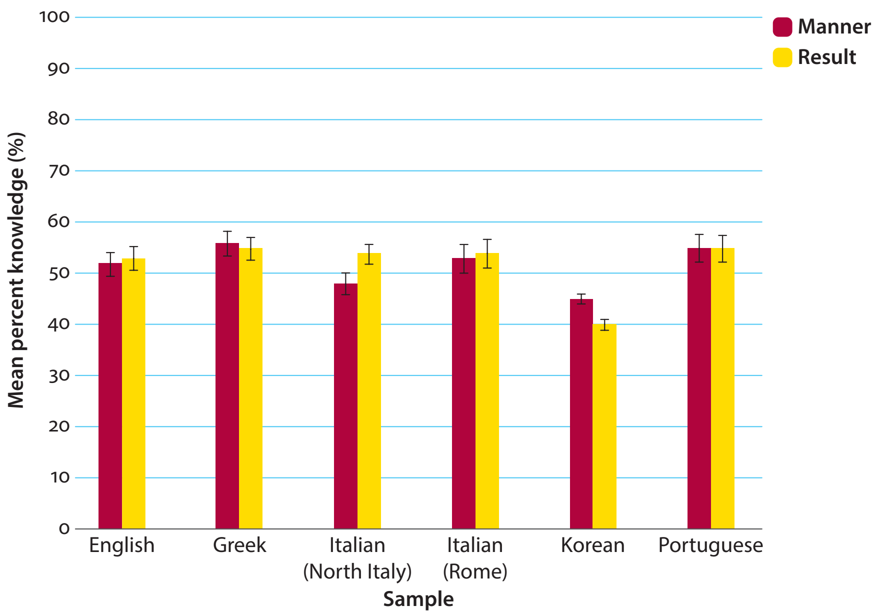


Figure 2. Mean percent knowledge of manner and result verbs across languages. Error bars indicate standard errors of subject means

group ($z = 10.84$, $p < .001$) (and no significant effect of gender, $z = 1.96$, $p = 0.05$). For the two Italian samples and the Korean sample, we found main effects of age group (Italian-North Italy, $z = 12.35$, $p < .001$; Italian-Rome, $z = 9.89$, $p < .001$; Korean, $z = 30.85$, $p < .001$) and gender (Italian-North Italy, $z = 3.14$, $p < .002$; Italian-Rome, $z = 2.039$, $p < .05$; Korean, $z = 11.52$, $p < .001$). For the Portuguese sample, we found only a main effect of age group ($z = 9.63$, $p < .001$). Model parameter estimates are in Table 6 in the online supplementary material.

3.2 Durative vs. punctual

3.2.1 *English*

Our coding, done by the first and third authors, yielded 15 durative verbs and 16 punctual verbs. For English, mean percent knowledge of durative verbs was 55%, and for punctual verbs, 49%; our mixed-effects model yielded a main effect of this factor ($z = -2.060$, $p < 0.05$). We also found main effects of gender ($z = 2.97$, $p < .005$) and age group ($z = 9.045$, $p < .001$), but no for frequency group, no main effect ($z = 0.89$, $p = 0.37$) or interaction with the semantic feature of punctual versus durative ($z = -0.831$, $p = 0.42$). Model parameter estimates are in Table 7 in the online supplementary material. Figure 3 depicts the English scores by frequency group.

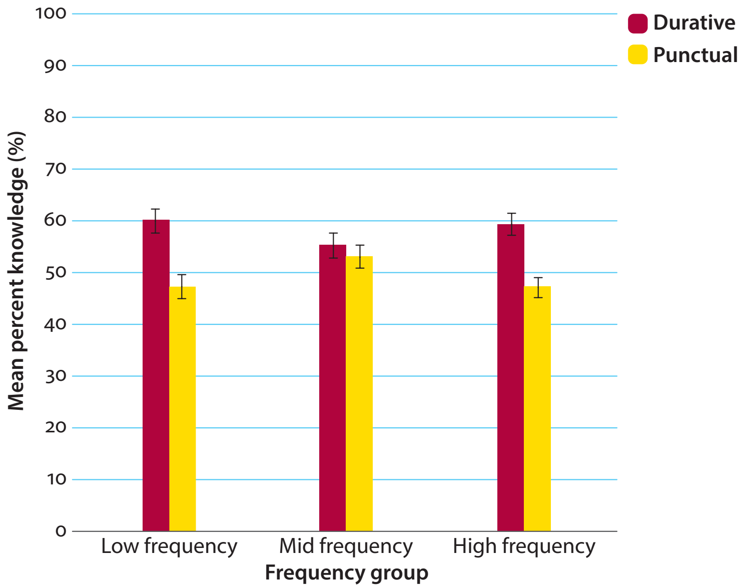


Figure 3. Mean percent knowledge for durative and punctual verbs by frequency group for English-acquiring toddlers. Error bars indicate standard errors of subject means

3.2.2 *Other languages*

See Figure 4. By contrast with English, most of the languages except for Korean did not show an effect of whether a verb labeled a durative or punctual event (all *z*s < 1.5, all *p*s > 0.13). Only in Korean was this effect significant ($z = -2.45$, $p < 0.02$). In all languages we found main effects of age group (Greek, $z = 10.88$, $p < .001$; Italian-North Italy, $z = 12.48$, $p < .001$; Italian-Rome, $z = 9.92$, $p < .001$; Korean, $z = 29.99$, $p < .001$; Portuguese, $z = 9.45$, $p < .001$) and for Greek, Italian, and Korean we also found main effects of gender (Greek, $z = 2.16$, $p < .05$; Italian-North Italy, $z = 3.23$, $p < .002$; Italian-Rome, $z = 2.11$, $p < .05$; Korean, $z = 11.36$, $p < .001$). For Portuguese, we did not find a gender effect (not surprisingly, given the lack of gender effect for manner vs. result). In all languages the trend, at least, is in the same direction as for Korean and English: durative verbs are better known than punctual verbs. Model parameter estimates are in Table 8 in the online supplementary material.

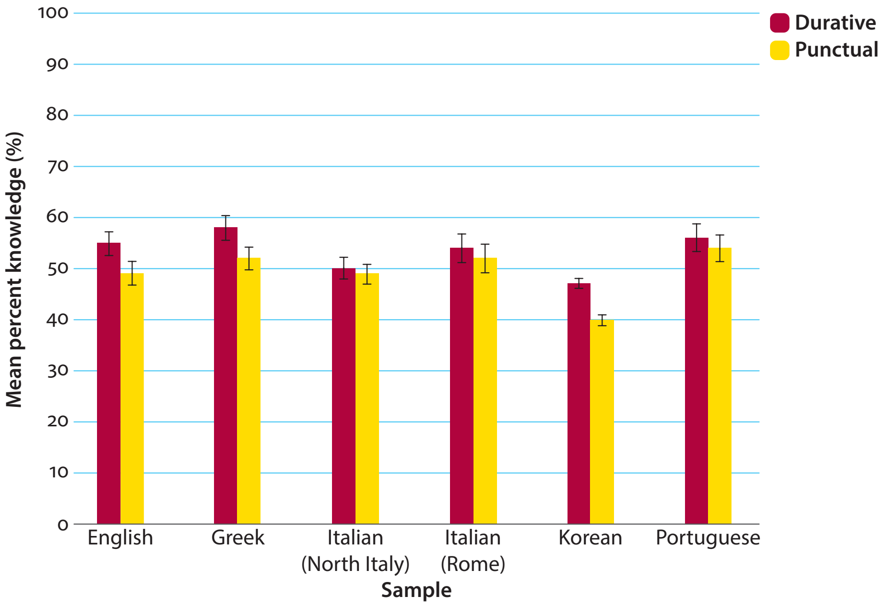


Figure 4. Mean percent knowledge for durative and punctual verbs across languages. Error bars indicate standard errors of subject means

3.3 Number of obligatory event participants

3.3.1 *English*

Our coding yielded 13 1-participant verbs, 21 2-participant verbs, and seven 3-participant verbs. English learners showed a clear effect of the number of event participants; mean percent knowledge scores were 56% for 1-participant verbs, 54% for 2-participant verbs, and only 42% for 3-participant verbs. Inspection of Figure 5, which depicts these means by frequency group, reveals that the number of event participants appeared most important for high-frequency verbs. A mixed-effects model yielded the anticipated main effect of event participant number ($z = -3.23$, $p < 0.005$), as well as main effects of age group ($z = 9.12$, $p < .001$), gender ($z = 2.91$, $p < 0.005$), and frequency ($z = 2.40$, $p < .02$), but no interaction between event participant number and frequency ($z = -1.78$, $p = 0.07$). These findings indicate that the number of event participants mattered over and above differences in input frequency. Model parameter estimates are in Table 9 in the online supplementary material.

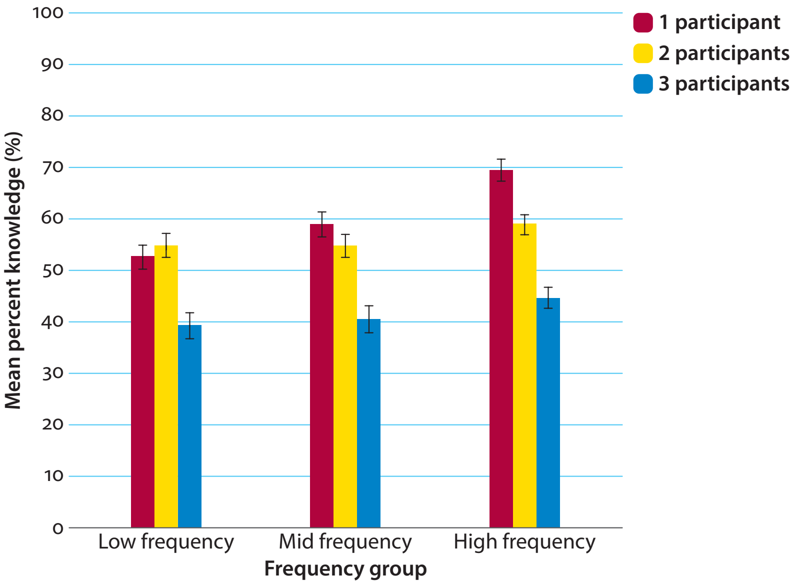


Figure 5. Mean percent knowledge for 1-, 2-, and 3-participant verbs by frequency group for English-acquiring toddlers. Error bars indicate standard errors of subject means

3.3.2 *Other languages*

Findings from the other languages were mixed. See Figure 6. Significant effects of event participant number obtained for toddlers from Rome ($z = -2.16$, $p < .05$) and Portugal ($z = -2.07$, $p < .05$). For Greek, Northern Italian, and Korean, there was no such effect (all *z*s < 1.8, all *p*s > 0.08). It was not the case in any language, however, that toddlers knew 2-participant verbs more than 1-participant verbs, or 3-participant more than 2-participant verbs, though in some cases percent knowledge was the same for two of the categories. All languages showed significant effects of age group (Greek, $z = 11.52$, $p < .001$; Italian-North Italy, $z = 12.52$, $p < .001$; Italian-Rome, $z = 9.77$, $p < .001$; Korean, $z = 30.87$, $p < .001$; Portuguese, $z = 9.61$, $p < .001$), and all except Portuguese showed significant effects of gender (Greek, $z = 2.25$, $p < .02$; Italian-North Italy, $z = 3.18$, $p < .002$; Italian-Rome, $z = 2.22$, $p < .03$; Korean, $z = 11.77$, $p < .001$). Model parameter estimates are in Table 10 of the online supplementary material.

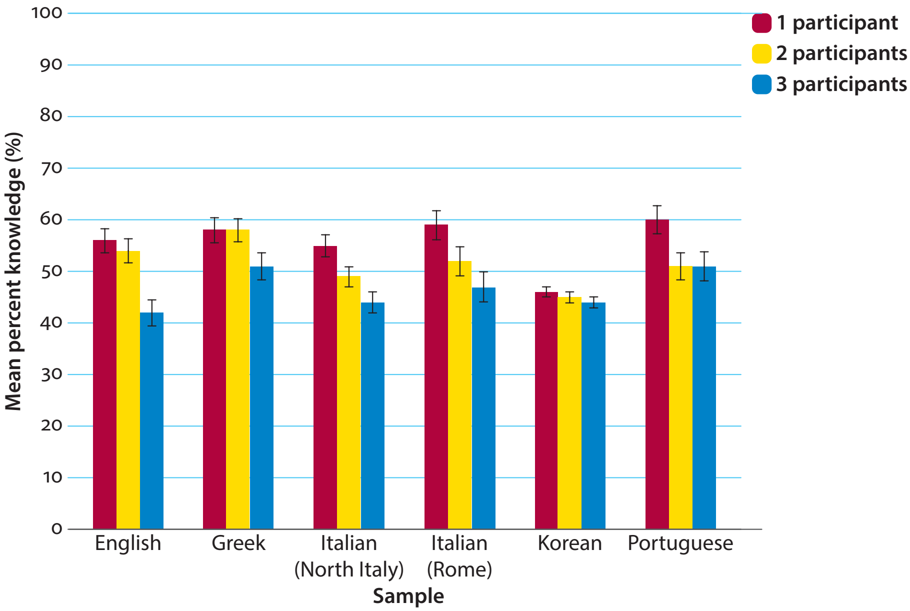


Figure 6. Mean percent knowledge for 1-, 2-, and 3-participant verbs across languages. Error bars indicate standard errors of subject means

4. Discussion

Our goal in the current study was to use vocabulary checklist data from a large number of toddlers and several languages to ascertain what kinds of verbs toddlers have in their productive vocabularies. We chose three semantic properties of the verbs: whether they describe a manner or result, whether they describe durative or punctual events, and whether these events involve one, two, or three event participants. We are most confident about our analyses of English, given that (a) the diagnostics we used are specific to English and translation equivalents in other languages may have quite different grammatical properties, and (b) we were only able to include input frequency for the English data. Nevertheless, it is intriguing that there appear to be similar trends across languages.

4.1 Summary of results

With respect to manner vs. result, we found no difference in any language group. For English, this was true even after accounting for the frequencies of the verbs in child-directed speech. Given that we aimed for a large-scale systematic analysis of manner and result verbs across the lexicon and across languages, it is interesting but perhaps not surprising that our findings point to a middle ground between those focused studies on novel verb learning or those focused on a small subset of verbs that have found either a manner preference or a result preference (see also van Hout, this volume, for related discussion). Our results thus contribute a new perspective on this manner/result distinction, indicating that once a large group of verbs, and a large group of toddlers, is investigated, biases for manner or result components wash out in toddlers' productive vocabularies.

Numerically, across language groups, verbs describing durative events were better known than verbs describing punctual events, although this difference was statistically reliable only for English (which had input frequency included in the analysis) and Korean. The directionality for these two languages, and the non-reliable trends for the other languages, are as predicted: we suspected that verbs describing events that are easily observed over a period of time would be easier to acquire than verbs describing events that occur quickly. The importance of durativity is interesting in light of the null result we found for the manner-result distinction; after all, manner verbs often describe durative events. Four of the punctual verbs on the LDS are semelfactives (manner verbs describing punctual events, such as *cough*), which might be expected to be more difficult to acquire given that they cross-cut other semantic categories, but in our data set the mean percent knowledge for semelfactives (47%) was similar to that for non-semelfactive punctual verbs (49%). Nevertheless, semelfactives may be a useful place for future work to look to understand the ease or difficulty of acquiring verbs with particular semantic features.

There may also be differences in the frequency with which these verbs are used in different languages (recall that we only included frequency information for English) or the ways in which they are used. One relevant note for Italian is that the verb "clap" as listed on the LDS was not a single word, but rather "battere le mani," and this punctual verb was relatively well known compared to the other punctual verbs on the list; it could be that the semantic transparency of this phrase with the high-frequency word "mani" for "hands" increases ease of acquisition for Italian-acquiring toddlers.

For the number of event participants, too, where we found differences, they aligned with our predictions: across languages we saw numerically better knowledge of verbs with fewer event participants than more event participants, although this difference was only reliable in English, the Rome sample of Italian, and Portuguese. But importantly, in no language sample did the trends go in the opposite direction from English. For English, the reliability of event participant number over and above input frequency suggests that the complexity of an event plays a role in its acquisition. This does not necessarily mean that toddlers have difficulty representing complex events. It could instead be due to the presumably greater difficulty of identifying a verb's referent when more event participants are involved. For example, an event in which one person tosses an apple to another can be labeled as *throwing*, *catching*, or *giving* – the first two possibilities require only two event participants, while the third requires all three. A *sleeping* event, by contrast, is more likely to only have one salient individual in the visual scene, making it easier to identify this event participant as the only relevant one.

Although we had data from several languages, with some typological diversity, this study represents a convenience sample of languages for which we had LDS data. Nevertheless, we believe the overall similarity across language groups indicates some cross-linguistic trends, in line with other research examining larger-scale properties of vocabulary such as proportion of nouns and verbs (e.g., Bornstein et al., 2004). Even when we zoom in on verbs only, and semantic categories within verbs, our findings suggest that some of the same features are relevant across languages. Of course, there are limitations of our approach that particularly hinder our interpretation of these cross-linguistic patterns; we turn to these limitations below.

Methodologically, we claim that large-scale surveys can be useful for asking questions about what kinds of words children know. Drawing on semantic and syntactic notions from linguistic and language acquisition theory can provide a framework for categorizing the words along dimensions that will be relevant for learners' future linguistic development. Our approach also permits focus on how verb meanings affect ease of acquisition, while yet controlling for other factors like frequency that are already known to be important, to ascertain the independent influence of semantic features.

4.2 Limitations and future directions

Parent report checklists come with limitations. First, the LDS asks parents to report words their child *produces*. Production is likely to be easier for parents to accurately report than comprehension, but importantly, toddlers' comprehension far outstrips their production; in fact, according to recent evidence even 10-month-olds know the meanings of some verbs (Bergelson & Swingley, 2013). Further, parents may not elicit verb production in their toddlers to the same degree they elicit noun production, and may thus not have heard the full range of verbs their toddler knows (Goldfield, 2000).

Second, checklists do not reveal the linguistic or extralinguistic contexts in which toddlers use verbs. There is ample evidence that toddlers' use of tense and aspect morphemes is initially strongly correlated with verb type; for example, past/perfect marking tends to appear first on verbs like "break" across languages (e.g., Slobin, 1985; Antinucci & Miller, 1976). The range of other words with which verbs are produced (e.g., the number and type of overt arguments) is also likely to be revealing. It may also be that children produce verbs but have incorrect semantic features as part of their representations of them. For this reason, it is important to integrate large-scale studies like ours with experimental and corpus work that can look more closely at children's comprehension and production of verbs. For example, related chapters by van Hout (this volume) and Schulz (this volume) discuss whether children's early verb representations include correct encoding of telicity.

Third, the particular words chosen for inclusion on the checklist will determine the results of an analysis like ours. Although we included a large number of verbs – 48 – in our analyses, in principle it could be that of the verbs that do not appear on the list, punctual verbs are better known than durative verbs, thus evening out the distributions. We think this is unlikely; although in developing the LDS Rescorla did not specifically attend to these semantic dimensions as we have coded them here, her choice of words to include on the checklist was based on studies of early lexical development (Nelson, 1973; Rescorla, 1980), and includes high-frequency words that most toddlers are expected to know as well as less common words.

Fourth, a checklist provides insight into which words a child knows at a particular time point, rather than what happened when the child first encountered that verb. Novel verb learning studies, by contrast, aim to address how and whether, on initial encounters, toddlers posit a correct representation of the novel verb's semantics. It is not surprising that these two approaches may provide different results, although a strength of having cross-sectional checklist data from a wide age range is that we can make inferences about which verbs were easier or more difficult to acquire even without observing the acquisition process.

There are admittedly several limitations associated with our English-centric approach; we categorized the verbs based on diagnostics for English, and these categories may not carry over perfectly to other languages. Manner and result classes do exist cross-linguistically (Levin, 2011), though their syntactic realization as well as the particular translation equivalents may differ. For example, it could be that "cut," argued to be polysemous between manner and result in English (Levin & Rappaport Hovav, 2013; Rappaport Hovav & Levin, 2010), necessarily entails a result in another language. We thus must proceed with caution in interpreting our findings, and we consider this important for future study; careful, language-specific, lexical-semantic analysis must precede thorough investigation of the acquisition question. We offer here a direction for further research, recognizing that ours is only a first attempt at what we consider a critical question of semantic development.

Another cautionary note with interpreting the results, and particularly null effects, in the non-English languages in our sample is that we were unable to include input frequency as a fixed effect; thus we cannot account for the fact that more frequently heard verbs are expected to be early and widely acquired, independently of their semantic properties. For English, we found that semantic features can affect acquisition over and above frequency. For the moment, then, we assume that the same will be true for other languages, although we hope that the coming years will bring larger and richer corpora of child-directed speech for a variety of languages.

Finally, we note that the LDS was developed as a screening tool to easily identify language delay in toddlers. In the current study we did not distinguish between toddlers based on vocabulary. Interestingly, toddlers with low vocabularies who are otherwise developing normally may shed light on whether the ease or difficulty of acquiring different verb types is semantic or conceptual. We might hypothesize that if toddlers with language delay are not delayed in conceptual understanding, they should have similar patterns to typically developing toddlers, even if attenuated (as with, for example, verbs denoting durative vs. punctual events). However, we might expect to see difference in knowledge based on linguistic features of verbs (as with manner vs. result verbs). Horvath, Rescorla, and Arunachalam (2015; in revision) specifically compare toddlers with low vocabularies to toddlers in the normal range, finding that those at risk for language delay do show slightly different patterns with respect to the semantic features studied here. This is true despite that overall, across the full vocabulary, there are strong correlations between the words known by low-vocabulary and typical-vocabulary children (Rescorla, Alley, & Christine, 2001).

We believe that this line of work offers a deeper understanding of early vocabulary growth. We also believe that our findings will be important for understanding toddlers who are growing up multilingually, whose different languages may present different learning challenges for particular lexical items, as well as toddlers with or at risk for language delays and disorders. For these toddlers, some verb types may be particularly unlikely to be acquired without intervention, and different verbs may be optimally presented in different linguistic and extralinguistic contexts.

Acknowledgement

Funding for this study was provided by NIH (K01 DC013306) to Sudha Arunachalam.

References

Abbot-Smith, K., Imai, M., Durrant, S., & Nurmsoo, E. (2017). The role of timing and prototypical causality on how preschoolers fast-map novel verb meanings. *First Language*, 37(2), 186–204.

Allen, S., Ozyürek, A., Kita, S., Brown, A., Furman, R., Ishizuka, T., & Fujii, M. (2007). Language-specific and universal influences in children's syntactic packaging of manner and path: A comparison of English, Japanese, and Turkish. *Cognition*, 102, 16–48. https://doi.org/10.1016/j.cognition.2005.12.006

Antinucci, F., & Miller, R. (1976). How children talk about what happened. *Journal of Child Language*, 3, 167–89. https://doi.org/10.1017/S0305000900001434

Bassano, D. (2000). Early development of nouns and verbs in French: Exploring the interface between lexicon and grammar. *Journal of Child Language*, 27, 521–59. https://doi.org/10.1017/S0305000900004396

Bates, D., Maechler, M., & Bolker, B. (2012). lme4: Linear mixed-effects models using S4 classes. R package version 0.999375–42; 2011.

Beavers, J., & Koontz-Garboden, A. (2012). Manner and result in the roots of verbal meaning. *Linguistic Inquiry*, 43, 331–369.

Behrend, D. A. (1990). The development of verb concepts: Children's use of verbs to label familiar and novel events. *Child Development*, 61, 681–96. https://doi.org/10.2307/1130953

Behrens, H. (1993). Temporal reference in German child language: Form and function of early verb use (Unpublished doctoral dissertation). University of Amsterdam.

Bergelson, E., & Swingley, D. (2013). The acquisition of abstract words by young infants. *Cognition*, 127, 391–397. https://doi.org/10.1016/j.cognition.2013.02.011

Berman, R. I., & Slobin, D. I. (1994). *Relating events in narrative: A cross-linguistic developmental study*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Bornstein, M. H., Cote, L. R., Maital, S., Painter, K., Park, S-Y., Pascual, L., Pêcheux, M. G., Ruel, J., Venuti, P., & Vyt, A. (2004). Cross-linguistic analysis of vocabulary in young children: Spanish, Dutch, French, Hebrew, Italian, Korean, and American English. *Child Development*, 75, 1115–1139. https://doi.org/10.1111/j.1467-8624.2004.00729.x

Bowerman, M. (1982). Reorganizational processes in lexical and syntactic development. In E. Wanner & L. R. Gleitman (Eds.), *Language acquisition: The state of the art* (pp. 319–346). Cambridge: Cambridge University Press.

Choi, Y., & Arunachalam, S. (2013). Learning manner and path verbs from the serial verb construction in Korean. In S. Baiz, N. Goldman, & R. Hawkes (Eds.), *BUCLD 37: Proceedings of the 37th annual Boston University Conference on Language Development* (pp. 62–73). Somerville, MA: Cascadilla Press.

Choi, S., & Bowerman, M. (1991). Learning to express motion events in English and Korean: The influence of language-specific lexicalization patterns. *Cognition*, 41, 83–121. https://doi.org/10.1016/0010-0277(91)90033-Z

Comrie, B. (1976). *Aspect*. Cambridge: Cambridge University Press.

Clark, E. V. (1995). *The lexicon in acquisition*. Cambridge: Cambridge University Press.

Clark, E. V., Carpenter, K. L., & Deutsch, W. (1995). Reference states and reversals: Undoing actions with verbs. *Journal of Child Language*, 22, 633–662. https://doi.org/10.1017/S0305000900009983

Cortese, M. J., & Fugett, A. (2004). Imageability ratings for 3,000 monosyllabic words. *Behavior Research Methods, Instruments, & Computers*, 36, 384–387. https://doi.org/10.3758/BF03195585

Dale, P. S., & Fenson, L. (1996). Lexical development norms for young children. *Behavior Research Methods, Instruments, & Computers*, 28, 125–127. https://doi.org/10.3758/BF03203646

Dale, P. S., & Goodman, J. C. (2005). Commonality and individual difference in vocabulary growth. In M. Tomasello & D. I. Slobin (Eds.), *Beyond nature-nurture: Essays in honor of Elizabeth Bates* (pp. 41–78). Mahwah, NJ: Lawrence Erlbaum Associates.

de Lemos, C. (1981). Interactional processes in the child's construction of languages. In W. Deutsch (Ed.), *The child's construction of language* (pp. 57–76). New York, NY: Academic Press.

Fillmore, C. (1970). The grammar of hitting and breaking. In R. Jacobs & P. Rosenbaum (Eds.), *Readings in English transformational grammar* (pp. 120–133). Waltham, MA: Ginn.

Fisher, C. (1996). Structural limits on verb mapping: The role of analogy in children's interpretations of sentences. *Cognitive Psychology*, 31, 41–81. https://doi.org/10.1006/cogp.1996.0012

Forbes, J. N., & Farrar, M. J. (1995). Learning to represent word meaning: What initial training events reveal about children's developing action verb concepts. *Cognitive Development*, 10, 1–20. https://doi.org/10.1016/0885-2014(95)90016-0

Forbes, J. N., & Poulin-Dubois, D. (1997). Representational change in young children's understanding of familiar verb meaning. *Journal of Child Language*, 24, 389–406. https://doi.org/10.1017/S0305000997003127

Gentner, D. (1978). On relational meaning: The acquisition of verb meaning. *Child Development*, 49, 988–998. https://doi.org/10.2307/1128738

Gentner, D. (1982). Why nouns are learned before verbs: Linguistic relativity versus natural partitioning. In S. A. Kuczaj (Ed.), *Language development, Vol. 2: Language, thought, and culture* (pp. 301–334). Hillsdale, NJ: Lawrence Erlbaum Associates.

Gertner, Y., & Fisher, C. (2012). Predicted errors in children's early sentence comprehension. *Cognition*, 124, 85–94. https://doi.org/10.1016/j.cognition.2012.03.010

Gillette, J., Gleitman, H., Gleitman, L., & Lederer, A. (1999). Human simulations of vocabulary learning. *Cognition*, 73, 135–176. https://doi.org/10.1016/S0010-0277(99)00036-0

Gleitman, L. (1990). The structural sources of verb meanings. *Language Acquisition*, 1, 3–55. https://doi.org/10.1207/s15327817la0101_2

Goodman, J. C., Dale, P. S., & Li, P. (2008). Does frequency count? Parental input and the acquisition of vocabulary. *Journal of Child Language*, 35, 515–531. https://doi.org/10.1017/S0305000907008641

Goldfield, B. A. (2000). Nouns before verbs in comprehension vs. production: The view from pragmatics. *Journal of Child Language*, 27, 501–520. https://doi.org/10.1017/S0305000900004244

Gopnik, A., & Meltzoff, A. N. (1986). Relations between semantic and cognitive development in the one-word stage: The specificity hypothesis. *Child Development*, 57, 1040–1053. https://doi.org/10.2307/1130378

Greenfield, P. M., & Smith, J. H. (1976). *Structure of communication in early language development*. New York, NY: Academic Press.

Gropen, J., Pinker, S., Hollander, M., & Goldberg, R. (1991). Affectedness and direct objects: The role of lexical semantics in the acquisition of verb argument structure. *Cognition*, 41, 153–195. https://doi.org/10.1016/0010-0277(91)90035-3

Horvath, S., Rescorla, L., & Arunachalam, S. (2015). Differences in expressive verb vocabulary between typically developing and delayed language learners at two years. Paper presented at the 2015 American Speech-Language-Hearing Association Convention, Denver, CO.

Horvath, S., Rescorla, L., & Arunachalam, S. (in revision). The syntactic and semantic features of two-year olds' verb vocabularies: A comparison of typically developing children and late talkers. *Journal of Child Language*.

van Hout, A. (2018). On the acquisition of event culmination. In K. Syrett & S. Arunachalam (Eds.), *Semantics in Language Acquisition* (pp. 95–122). Amsterdam: John Benjamins.

Husband, E. M. (2011). Rescuing manner/result complementarity from certain death. *Proceedings of the 47th annual Chicago Linguistics Society*. Chicago, IL: CLS.

Huttenlocher, J., Smiley, P., & Charney, R. (1983). Emergency of action categories in the child: Evidence from verb meanings. *Psychological Review*, 90, 72–93. https://doi.org/10.1037/0033-295X.90.1.72

Imai, M., Li, L., Haryu, E., Okada, H., Hirsh-Pasek, K., Golinkoff, R., et al. (2008). Novel noun and verb learning in Chinese-, English-, and Japanese-speaking children. *Child Development*, 79, 979–1000. https://doi.org/10.1111/j.1467-8624.2008.01171.x

Levin, B. (2011). *Verb classes within and across languages*. Handout, Leipzig, April.

Levin, B., & Rappaport Hovav, M. R. (1991). Wiping the slate clean: A lexical semantic exploration. *Cognition*, 41, 123–151. https://doi.org/10.1016/0010-0277(91)90034-2

Levin, B., & Rappaport Hovav, M. R. (1995). *Unaccusativity: At the syntax-lexical semantics interface*. Cambridge, MA: The MIT Press.

Levin, B., & Rappaport Hovav, M. R. (2013). Lexicalized meaning and manner/result complementarity. In B. Arsenijević, B. Gehrke, & R. Marín (Eds.), *Subatomic semantics of event predicates* (pp. 49–70). Dordrecht: Springer.

Li, P. (2001). The CHILDES Parental Corpus. Retrieved from <http://childes.psy.cmu.edu/derived/parentfreq.cdc>

Li, P., & Shirai, Y. (2000). *The acquisition of lexical and grammatical aspect*. Berlin: Walter de Gruyter. https://doi.org/10.1515/9783110800715

Ma, W., Golinkoff, R. M., Hirsh-Pasek, K., McDonough, C., & Tardif, T. (2009). Imageability predicts the age of acquisition of verbs in Chinese children. *Journal of Child Language*, 36, 405–23. https://doi.org/10.1017/S0305000908009008

MacWhinney, B. (2000). *The CHILDES Project: Tools for analyzing Talk* (3rd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.

Maratsos, M. (1990). Are actions to verbs as objects are to nouns? On the differential semantic bases of form, class, category. *Linguistics*, 28, 1351–1380. https://doi.org/10.1515/ling.1990.28.6.1351

Naigles, L. R., & Kako, E. T. (1993). First contact in verb acquisition: Defining a role for syntax. *Child Development*, 64, 1665–1687. https://doi.org/10.2307/1131462

Naigles, L. R., & Hoff-Ginsberg, E. (1998). Why are some verbs learned before other verbs? Effects of input frequency and structure on children's early verb use. *Journal of Child Language*, 25, 95–120. https://doi.org/10.1017/S0305000997003358

Naigles, L. R., Hoff, E., Vea, D., Tomasello, M., Brandt, S., Waxman, S. R., & Childers, J. B. (2009). Flexibility in early verb use: Evidence from a multiple-n diary study. *Monographs of the Society for Research in Child Development*, 74(2), vii–112.

Nelson, K. (1973). Structure and strategy in learning to talk. *Monographs of the Society for Research in Child Development*, 38(1–2), 1–135. https://doi.org/10.2307/1165788

Papaeliou, C. F., & Rescorla, L. (2011). Vocabulary development in Greek children: A cross-linguistic comparison using the Language Development Survey. *Journal of Child Language*, 38, 861–887. https://doi.org/10.1017/S030500091000053X

Papafragou, A., Cassidy, K., & Gleitman, L. (2007). When we think about thinking: The acquisition of belief verbs. *Cognition*, 105, 125–165. https://doi.org/10.1016/j.cognition.2006.09.008

Penner, Z., Schulz, P., & Wymann, K. (2003). Learning the meaning of verbs: What distinguishes language-impaired from normally developing children? *Linguistics*, 41, 289–319. https://doi.org/10.1515/ling.2003.010

R Development Core Team (2012). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

Rappaport Hovav, M., & Levin, B. (2010). Reflections on manner/result complementarity. In M. Rappaport Hovav, E. Doron, & I. Sichel (Eds.), *Lexical semantics, syntax, and event structure* (pp. 21–37). Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199544325.003.0002

Rescorla, L. (1980). Overextension in early language development. *Journal of Child Language*, 7, 321–335. https://doi.org/10.1017/S0305000900002658

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

Rescorla, L., & Achenbach, T. M. (2000). Use of the Language Development Survey (LDS) in a national probability sample of children 18 to 35 months old. *Journal of Speech, Language, and Hearing Research*, 45, 733–743. https://doi.org/10.1044/1092-4388(2002/059)

Rescorla, L., Alley, A., & Christine, J. B. (2001). Word frequencies in toddlers' lexicons. *Journal of Speech, Language, and Hearing Research*, 44, 598–609. https://doi.org/10.1044/1092-4388(2001/049)

Rescorla, L., Nyame, J., & Dias, P. (2016). Vocabulary development in European Portuguese: A replication study using the Language Development Survey. *Journal of Speech, Language, and Hearing Research*, 59, 1484–1490.

Rescorla, L., Frigerio, A., Sali, M. E., Spataro, P., & Longobardi, E. (2014). Typical and delayed lexical development in Italian. *Journal of Speech, Language, and Hearing Research*, 57, 1792–1803. https://doi.org/10.1044/2014_JSLHR-L-13-0242

Rescorla, L., Lee, Y. M. C., Oh, K. J., & Kim, Y. A. (2013). Lexical development in Korean: Vocabulary size, lexical composition, and late talking. *Journal of Speech, Language, and Hearing Research*, 56, 735–747. https://doi.org/10.1044/1092-4388(2012/11-0329)

Rescorla, L., & Safyer, P. (2013). Lexical composition in children with autism spectrum disorder (ASD). *Journal of Child Language*, 40, 47–68. https://doi.org/10.1017/S0305000912000232

Rissman, L. (2015). Cinderella broke and broke: Object deletion and manner-result complementarity. *Proceedings from the 51st Meeting of the Chicago Linguistic Society* (pp. 425–439). Chicago, IL: CLS.

Schulz, P. (2018). Acquisition of telicity. In K. Syrett & S. Arunachalam (Eds.), *Semantics in Language Acquisition* (pp. 123–150). Amsterdam: John Benjamins.

Scott, R., Gertner, Y., & Fisher, C. (2018). Not all subjects are agents: Transitivity and meaning in early language comprehension. In K. Syrett & S. Arunachalam (Eds.), *Semantics in Language Acquisition* (pp. 153–176). Amsterdam: John Benjamins.

Slobin, D. I. (1985). Crosslinguistic evidence for the language-making capacity. *The Crosslinguistic Study of Language Acquisition*, 2, 1157–1256.

Smith, C. S. (1991). *The parameter of aspect*. Dordrecht: Kluwer. https://doi.org/10.1007/978-94-015-7911-7

Snedeker, J., Zeitlin, M., & Crawford, J. (2013). Why are some words learned before others? Predictors of lexical production in infants and children learning English. Paper presented at the 38th annual Boston University Conference on Language Development, Boston, MA.

Tardif, T. (1996). Nouns are not always learned before verbs: Evidence from Mandarin speakers' early vocabularies. *Developmental Psychology*, 32, 492–504. https://doi.org/10.1037/0012-1649.32.3.492

Tardif, T., Gelman, S. A., & Xu, F. (1999). Putting the "noun bias" in context: A comparison of English and Mandarin. *Child Development*, 70, 620–635. https://doi.org/10.1111/1467-8624.00045

Tomasello, M. (1992). *First verbs: A case study in grammatical development*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511527678

Waxman, S., Fu, X., Arunachalam, S., Leddon, E., Geraghty, K., & Song, H. J. (2013). Are nouns learned before verbs? Infants provide insight into a long-standing debate. *Child Development Perspectives*, 7, 155–159. https://doi.org/10.1111/cdep.12032

Yuan, S., Fisher, C., & Snedeker, J. (2012). Counting the nouns: Simple structural cues to verb meaning. *Child Development*, 83, 1382–1399. https://doi.org/10.1111/j.1467-8624.2012.01783.x

SECTION II

Event semantics

CHAPTER 5

On the acquisition of event culmination

Angeliek van Hout
University of Groningen

There is quite a high rate of acceptance of telic-perfective predicates as descriptions of non-culminating events in children learning Germanic and Romance languages. What causes children, much more so than adults, to accept non-culminating interpretations of telic-perfective sentences? In this review, I discuss learners' difficulties in each of three grammatical dimensions that contribute to event culmination: the notion of 'result' as encoded in the lexical semantics of verbs, telicity of verb phrases, and perfectivity of tense-aspect morphology. I conclude that telicity and perfectivity do not cause the non-culmination acceptance patterns. Instead, the learnability challenge for event culmination lies in the acquisition of verb meanings. I sketch several new angles for further research, including the role of agentivity of the subject.

Keywords: verb meaning, telicity, aspect, perfectivity, event culmination, completion entailment, L1 acquisition, scalar semantics, pragmatic inferences

1. Introduction

This chapter reviews research on the acquisition of event culmination in typically-developing children. Telic-perfective sentences (such as, *The child walked to the park in an hour*) entail event culmination (that is, the event was completed). Many studies on child language have, however, discovered an often high rate of acceptance of such sentences as descriptions of non-culminating events in learners of Germanic and Romance languages. This chapter discusses several explanations for children's non-culminating interpretations of telic-perfective sentences in this stage of acquisition.

There are three grammatical dimensions that, together, determine event culmination: the lexical semantics of verbs, the telicity of verb phrases and the perfectivity of tense-aspect morphology. Verb semantics involves features such as result, manner, goal, path, motion and causation, some of which are relevant for event culmination, in particular, the lexical features of result and goal. Telicity is a property of

https://doi.org/10.1075/tilar.24.05hou

event descriptions as expressed by verb phrases; it involves a natural culmination point as part of the verb phrase meaning. Aspect is a property of tense-aspect markers, typically expressed by morphemes on the verb. Tense-aspect markers anchor the run-time of an event on a time line by relating it to a reference time. Perfective aspect asserts the final event boundary – there is a final moment at which a given event ceases and does not continue any longer. All three dimensions interact in the grammar of event culmination. Examples of each will appear below as these topics are introduced.

The notion of event culmination has been addressed in several lines of acquisition research, which have developed more or less independently, though sometimes they intersect. One such line is work on the acquisition of verb meanings, specifically, the question of how children learn about the result component in a verb's lexical representation; another is formed by studies on the acquisition of telicity; and a final line includes studies on the acquisition of perfective versus imperfective aspect. This chapter seeks to connect these three lines of acquisition research, focusing on the culmination entailment associated with telic predicates in combination with perfective aspect. Integrating my own research on event culmination of the past twenty years with the literatures on the acquisition of lexical verb meanings, telicity and aspect, I will point out connections, similarities and contrasts in the acquisition patterns across telic constructions, tense-aspect markers, verb types and languages. I describe in detail when, in certain types of languages, children are more liberal than adults in accepting incomplete situations for telic-perfective sentences. In so doing I focus on the following question for learnability: what can explain this stage in development? The goal is to critically evaluate the explanations that have been offered in the three lines of acquisition research.

This chapter is organized as follows: Section 2 defines the semantic notions of telicity and perfectivity, pointing out how these grammatical dimensions determine event culmination. The next three sections review studies on the acquisition of: the notion of result in lexical verb meanings (Section 3), telicity (Section 4) and perfective aspect (Section 5). Throughout these sections I discuss explanations for children's non-culminating interpretations of telic-perfective sentences. Finally, by integrating new developments in semantic theory about the nature of the subject of the sentence and event culmination – whether it is an intentional Agent or a natural Cause, Section 6 develops a research agenda with novel questions about the acquisition of event culmination.

2. The grammatical dimensions of event culmination

Before turning to acquisition I define the crucial grammatical properties involved in the grammar of event culmination – verb semantics, telicity and perfectivity – and discuss how these are involved in a step-by-step construal of event culmination at different grammatical levels: verb, verb phrase and sentence. Lexical verb meaning is one of the sources of the telicity of verb phrases. Other grammatical elements in the verb phrase can also play a role in determining telicity, in particular, satellite phrases (such as particles and PPs) and the quantization properties of the direct object. The combination of a telic predicate and a perfective aspect marker leads to an entailment of completion: the natural culmination point has been reached. This section will form the grammatical basis for the subsequent acquisition sections which will discuss to what extent the expression of these notions in a given language presents a challenge to the language learner.

Telicity is a property of event descriptions – verbs and verb phrases – and characterizes the temporal contour of events. Telicity falls under the type of aspect that is called "lexical aspect" (also called: "situation type", "inner aspect", "Aktionsart", "aspectual classification of verb classes"); as such it is independent of tense-aspect inflection. It contrasts with grammatical aspect, which is typically expressed by inflectional morphology on the verb (e.g., perfective prefixes in Slavic languages, perfective tenses in the Romance languages), or encoded in verbal constructions and periphrasis (e.g., the perfect tense with auxiliary plus participle in Germanic and Romance languages). Telicity has been used as a feature to classify the meanings of individual verbs (Vendler, 1957), but it is more appropriate to use it as a property of verb phrases (as pointed out by Verkuyl, 1972).

Definitions of telicity use terms such as "*telos*" (goal in Greek), "natural boundedness", "culmination", and "set terminal endpoint", all of which express the idea that a verb or verb phrase is telic if its meaning includes a specific moment toward which the event it describes naturally develops (Comrie, 1976; Dowty, 1979; Klein, 1994; Smith, 1991; Verkuyl, 1993). In mereological approaches (Krifka, 1989; Rothstein, 2004), telicity is analyzed in terms of event quantization: properties that describe the part-whole structure of events, specifically, homogeneity and cumulativity. The essential idea is that a subpart of an event described by a telic predicate does not qualify as an instance of that same event description (telic predicates are non-homogeneous). Nor does the combination of two events described by the same telic predicate qualify as one single instance of that event (telic predicates are non-cumulative). Thus, a telic expression is such that, whenever it is true of some event in the world, it is not true of any subparts of that event (at least, not necessarily, because the described event includes a culmination point and a subpart of the event may lack this moment). For example, crossing a street entails going from one side of the street to the other, and a subpart of that distance (e.g., going from one side to the middle of the road) cannot be described as crossing a street, and so the predicate *cross a street* is telic (quantized). An atelic event description, on the other hand, can refer both to the event as a whole and to any of its subparts. A subpart of an event of sleeping on the sofa also qualifies as sleeping on the sofa, and so the predicate *sleep on the sofa* is atelic (non-quantized). The traditional test for telicity contrasts two temporal modifiers (Dowty, 1979): *in an hour* combines with telic predicates, while *for an hour* combines naturally with atelic predicates, (1).

(1) a. The child slept on the sofa for an hour / * in an hour.
b. The child destroyed the Lego castle *for an hour / in an hour.

Telic predicates constitute a heterogeneous class; there are at least three grammatical sources that contribute to telicity. First, verb semantics plays a role: telicity can be purely lexical, determined by the intrinsic meaning of the verb itself which crucially lexicalizes the endpoint of an event. (1) Illustrates this for the telic verb *destroy* versus the atelic verb *sleep*. Other lexically telic verbs are *kill, cross, break, open* and *close*. To find out if a verb lexicalizes an endpoint, one can test if one can use the verb to describe an action that lacks this endpoint. For instance, an event of dropping a vase and nearly breaking it cannot be described as "break a vase", nor can an event of almost opening a door be described as "open a door".

In addition to the lexical semantics of individual verbs, the computation of telicity depends on other elements in the verb phrase. One is satellite phrases to the verb. For example, in the Germanic languages, verb particles and directional PPs form constructions that lexicalize the culmination point by explicitly encoding a result or a goal, for example, *eat-eat up, blow-blow out, run-run away* (van Hout, 1996)[1]. Verb phrases with a lexically atelic verb become telic when combined with such a PP, (2), or particle, (3)[2].

(2) a. The child biked in the park for an hour / * in an hour.
b. The child biked to school *for an hour / in an hour.

---

1. Across typologically different languages, there are various other kinds of "telicizing" constructions, including telic-reflexive clitic *se* in Spanish (Nishida, 1994), verb-verb compound constructions in languages like Mandarin (Soh & Kuo, 2005), and certain prefixes in Slavic languages (Borik, 2006; Filip, 2008).

2. Some speakers reading find the "for an hour" construction rather acceptable here and allow it to be followed by "but the box never made it all the way" This means that the particle *out* does not add maximal culmination, but merely a scale (here, in the direction of a place outside, not necessarily reaching it).

(3) a. The child carried the heavy box for an hour / * in an hour.
    b. The child carried the heavy box out *for an hour / in an hour.

Furthermore, the semantics of the direct object matters, in particular, the quantization properties of the direct object co-determine the telicity of a transitive verb phrase (verb plus object) (Krifka, 1989; Verkuyl, 1993). This type of telicity applies to incremental theme verbs, including verbs of consumption (*eat, drink*), verbs of creation (*draw, build*), and also some verbs expressing a change of state (*fill*), (4)–(6). The essence is that the quantization of the direct object carries over to the verb phrase: a quantized object gives a telic verb phrase, (4a)–(6a), whereas a non-quantized object (a mass term or bare plural) gives an atelic verb phrase, (4b)–(6b).

(4) a. The child ate a bar of chocolate *for an hour / in an hour.
    b. The child ate chocolate for an hour / * in an hour.

(5) a. The child drew a snow man *for an hour / in an hour.
    b. The child drew snowflakes for an hour / * in an hour.

(6) a. The child filled a bucket *for an hour / in an hour.
    b. The child filled buckets for an hour / * in an hour.

Coining two descriptive terms for these grammatical sources of telicity (van Hout, 2008a), I defined "predicate telicity", illustrated in (2)–(3), in terms of the presence of "… an overt marker on the verbal predicate that carries telicity" (p. 259), and contrasted this with "compositional telicity", illustrated in (4)–(6), as the "…joint syntactic-semantic effects … [of] a verb's lexical property of taking an incremental theme, … [and] the quantificational semantics of the direct object noun phrase (mass or count term)" (p. 259).

In addition to telicity, perfectivity presents another dimension of event culmination. The perfective-imperfective opposition is the grammatical aspect distinction that presents a speaker's perspective on an event. Definitions of perfective aspect employ terms such as "holistic event", "event viewed from the outside" and "event viewed as a whole", and contrast these with imperfective aspect notions such as "ongoing event", "event viewed from the inside" and "viewed with interior composition" (Comrie, 1976; Dahl, 1985; Smith, 1991). Interval-based approaches in formal semantics define perfective and imperfective aspect as a relation between two time intervals. Perfective aspect involves the inclusion of the complete run-time of an event in a reference time interval. This entails that the full event – including its initial and final boundaries – has happened and the event no longer continues (Demirdache & Uribe-Etxebarria, 2000; Kamp & Reyle, 1993; Klein, 1994, among others). Imperfective aspect presents the mirror image of this relation: the reference time is included in the run-time of the event. Thus, imperfective aspect asserts that the event was ongoing at a certain reference time, but does not make claims about its final boundary, as the event may have been interrupted without ever reaching culmination. This means that imperfective aspect "cuts off" the culmination point from a telic predicate: the event may or may not have reached its culmination point, and it may or may not still be continuing. This is the so-called 'imperfective paradox' (Dowty, 1979).

One test for perfectivity makes use of the notion of continuation (Smith, 1991). When a clause with perfective aspect is followed by a subsequent clause claiming that the event continues, there is a contradiction, (7a). With imperfective aspect, on the other hand, continuation is possible, (7b). The incongruity of (7a) reveals that a telic predicate (*cross the street*) in combination with perfective aspect – English simple past – entails culmination. In contrast, the acceptability of continuation of the same predicate with imperfective aspect – the English progressive, (7b) – shows that the culmination, which is a crucial part of the lexical meaning of the verb *cross*, does not need to be reached in the actual world (as long as there is a possible world in which culmination can be conceived, e.g., Landman, 1992). Imperfective aspect thus cancels the culmination entailment.

(7) a. #The child crossed the street. In fact, she is still crossing the street.
b. The child was crossing the street. In fact, she is still crossing the street.

In summary, event culmination presents a puzzle in reasoning, the pieces of which are determined by three types of grammatical elements: the lexical semantics of verbs, the telicity of verb phrases and the perfectivity of tense-aspect markers, as laid out above. The acquisition challenge for the learner is to sort out the effects of all these grammatical elements and their interplay.

The next three sections will discuss whether each of these grammatical sources can be the cause for children's difficulties with event culmination, in particular, their over-acceptance of incomplete situations for telic-perfective sentences. Section 3 discusses studies that investigate the lexical feature of result as encoded in verb meanings. Section 4 provides a summary of event culmination interpretation patterns established in studies on the acquisition of telicity focusing on i) the aspectual roles of telic particles and clitics and ii) direct object quantization. Section 5 reviews event culmination in studies on the acquisition of perfectivity, determining to what extent children know the meaning of perfective aspect markers and their interaction with telicity.

3. Non-culmination patterns and the acquisition of the lexical meaning of verbs

There is a long tradition of research on the acquisition of verb meanings. Many studies focus on lexical features of verbs such as causation, manner, path and motion (Hirsh-Pasek & Golinkoff, 2006; Tomasello & Merriman, 1995, among many others). Nevertheless, features relevant for event culmination, specifically lexical notions of result and end state, have hardly been investigated in this tradition. The original study that addressed verb semantics associated with telicity and event culmination was Gentner (1978) (though she did not label it in these terms). Gentner investigated what children know about the meanings of the verbs *mix* and *stir* in adult English, specifically, whether they associate these verbs with a particular manner (i.e., as an atelic description) or a particular endstate (i.e., as a telic description). While stirring merely refers to a manner of action with no implications about the effect, mixing is supposed to lead to an endstate in which two substances become more homogeneous. In other words, mixing implies some form of culmination, while stirring does not. Nonetheless, Gentner found that 5-year-olds accepted *mix* for situations with no particular endstate (e.g., mixing a substance that was homogeneous from the start), and concludes that the children misinterpreted *mix* as a manner verb, not including an endstate feature. Gentner proposed the Manner Bias, arguing that children initially assume that verbs refer to manners of actions, and not results. In a study with a similar focus, comparing telic change-of-state verbs such as *fill* and atelic manner verbs such as *pour* (without labeling these verbs in telicity terms), Gropen, Pinker, Hollander and Goldberg (1991) find that preschoolers accepted situations without an endstate for change-of-state verbs (*fill*), interpreting them instead as if they were manner verbs (*pour*). Seeing that children misinterpreted change-of-state verbs as manner verbs, ignoring the feature of change that is essential in the lexical meaning of these verbs, Gropen and colleagues argue that this interpretation pattern supports Gentner's Manner Bias.

The Manner Bias provides a straightforward account of children's overly liberal acceptance of non-culminating events: if children initially do not include the notions result or endstate in their lexical-semantic specification of verbs, culmination is irrelevant. Nevertheless, there are several reasons why this account cannot be on the right track. First, the generalization draws on a very limited set of verbs (*mix, fill*). Second, the Manner Bias predicts that children start out assuming that all verbs are manner verbs, i.e., that no verbs include an endstate or result component in their lexical semantics. Yet, the review of the telicity studies in Section 4 will reveal that, although children indeed accept non-culmination to some extent, they also very often reject it. Thus, the lexical representation of certain verbs in child grammars does include an endstate or a result meaning component. Finally, as with all accounts that assume a bias or a learning heuristic, there is the learnability problem of exiting the Manner Bias stage: What could trigger a child to adjust her lexical semantics for a given verb and add an endstate feature to its meaning?

Wittek (2002) also takes issue with the Manner Bias. Investigating German children's acquisition of the notion of endstate in verb meanings, Wittek compares two versions of change-of-state verbs: particle verbs such as *wachmachen* (literally, 'make awake') – which transparently encode the endstate in the particle – versus monomorphemic verbs such as *wecken* ('wake') – which also encode the endstate lexically, but in a non-transparent way. The results (of her experiment 3, see also Figure 1 in Section 4) indeed show an effect of the transparency of encoding: children rejected situations with a zero result more often for transparent particle verbs than for non-transparent verbs (e.g., a girl tries to wake up a man by making an alarm go off next to his ear, but he keeps snoring). Wittek concludes that children have acquired the lexical notion of endstate, at least for particle verbs, and argues that this goes against the Manner Bias. Nevertheless, in order to explain the relatively high rate of acceptance of non-culmination with non-transparent verbs (*wecken* 'wake'), Wittek proposes that children have what she calls a "Weak Endstate interpretation" for these change-of-state verbs: "they seem to interpret these verbs as if they specify that a particular endstate might well come about, but need not" (Wittek, 2002: 88). So, even though the verb's lexical meaning includes an endstate component, this component is optional (the endstate is weak).

Wittek (2002) supports the idea of weak-endstate verbs with a typological argument, pointing to several adult languages in which monomorphemic change-of-state verbs similarly do not imply completion: Mandarin (Talmy, 1991); Japanese (Ikegami, 1985); Hindi (Singh, 1994) and Tamil (Pederson, 1998). In fact, there is also a class of English verbs with weak endstates (e.g., *wipe*, *wash*, *sweep*): even though an interpretation including an endstate is "pragmatically favored" (terminology from Brisson, 1994), culmination can be denied in a "but"-clause, which shows that it is merely an implication, and not an entailment (Talmy, 1991). Wittek (2002) notes that this insight was originally presented, *inter alia*, by Harnish (1976) and Atlas and Levinson (1981). She argues that German children overextend the lexical representations for weak-endstate verbs to a larger class of verbs, thus locating the cause for non-culminating interpretations at a lexical level, and delegating the culmination inference of these verbs to the pragmatic domain.

In their studies on the acquisition of telicity in German, Schulz and colleagues have also stressed the importance of what they call Endstate Orientation as an acquisition strategy (Penner, Schulz & Wymann, 2003; Schulz, this volume; Schulz, Wymann & Penner, 2001). More recently, several other acquisition studies have stressed the role of endstates in young children's event representations, all of which argue that endstates are a basic feature in infants' and children's cognitive repertoire (Lakusta & DiFabrizio, 2016; Wagner & Lakusta, 2009).

In semantic theory, the insight that pragmatics plays a role in event culmination has led to developments beyond the mereological approach of telicity outlined in Section 2. Theories of scalar structures are now used to model the flexible behavior of English "wipe"-type verbs (also called "implied fulfillment" verbs). Hay, Kennedy & Levin (1999) and Kennedy and Levin (2008) advance such a scalar approach to telicity in which event culmination is partly determined by scalar semantics and partly by pragmatics. For predicates whose meaning encodes a closed scale (e.g., *break*), culmination is an entailment – a hard, non-cancelable inference. For other verb types, which do not lexicalize a closed scale (e.g., *wipe*), a culminating interpretation comes about as a conversational implicature – a soft, cancelable inference. The difference between these two kinds is determined by the event-semantic features in the lexical representation of verbs, in particular, whether or not the verb's meaning includes a degree variable, and if so, whether it is associated with an open (*wipe*) or closed (*break*) scale. Rephrasing Wittek's proposal in terms of this scalar theory of event culmination, the cause for children's over-acceptance of non-culminating situations for change-of-state verbs may well lie in an initial misrepresentation of these verbs as having an open scale (or no scale at all), in contrast to the target lexical representation of such verbs, namely with a closed scale. I will return to the lexical semantics of verbs in my discussion of acquisition of telicity studies in the next section.

4. Non-culmination patterns and the acquisition of telicity

In an earlier review of the acquisition of telicity (van Hout, 2008a), I concluded that there are different interpretation patterns depending on the type of verb and construction. I labeled these "predicate telicity" versus "compositional telicity" (see above). The generalization was that children acquire event culmination earlier when the verbal predicate overtly and explicitly encodes a result state, as is the case for particle verbs in the Germanic languages and for verbs marked for perfective aspect in the Slavic languages. This contrasts with cases in which telicity is based on the quantization semantics of the direct object, for example, when an incremental theme verb combines with a count term object (e.g., *eat an apple*). Several studies have examined this contrast further and will be included in the present review. My earlier review was grounded in a mereological theory of telicity in which predicate and compositionally telic constructions are analyzed similarly, based on the notions of homogeneity and cumulativity. Adopting a scalar approach to telicity as described in Section 3 (Hay et al., 1999; Kennedy and Levin, 2008), I will now re-assess the generalization that there are two patterns in the acquisition of telicity, focusing on the roles of particles, types of verbs and types of direct objects.

In the present review I have included the following studies: Hacohen (2010); Hodgson (2001); van Hout (1998); Jeschull (2007); Ogiela (2007); Ogiela, Schmitt & Casby (2014); Penner et al. (2003); Schulz et al. (2001); Schulz & Penner (2002); Schulz & Wittek (2003); Wittek (2002). I limit this review to experimental studies that tested children's interpretation, and have not included studies on children's production of tense-aspect markers in spontaneous speech, as developed in a research line on the so-called Aspect-First hypothesis (for a review, see Wagner, 2012). This choice is motivated by two reasons. First, production does not provide a direct measure of interpretation. Second, without carefully constructed contexts and contrasts, it is very hard to draw firm conclusions about children's knowledge of the meanings of tense-aspect forms and constructions.

The experimental designs systematically varied telic and atelic conditions and contrasted complete and incomplete situations. Test sentences involved one or more telic constructions (used with perfective aspect). The paradigms focused on the contrast between completed versus incomplete or ongoing situations, testing either acceptance of telic-perfective sentences in a truth-value judgment task, or preference with such sentences in a picture-selection task. The materials varied: pictures, movies, acted-out scenes. Several studies used pictures with characters involved in eating or drinking something; one character finishes his piece of food or drink and the other eats or drinks it halfway (van Hout, 1998; Ogiela, 2007; Schulz & Penner, 2002). The test sentences varied transitivity and particle verbs (e.g., *Did he eat*?, *Did he eat cheese*?, *Did he eat the cheese*?, *Did he eat up the cheese*? (see Appendix A for some picture pairs). Other studies used short movie clips contrasting situations in which either a full result was reached or nothing happened at all (e.g., a woman approaches a sleeping man and lets an alarm clock ring; the man either wakes up or does not wake up and keeps snoring, Wittek, 2002). The test sentences varied simplex and particle verbs: *Hat das Mädchen den Mann geweckt* / *wachgemacht* / *wachgeklingelt*? (has the girl the man woken / awake made / awake rung? 'Did the girl wake (up) the man?"). In yet other studies the experimenter acted out scenes of manipulating various objects, with either a partly or fully affected object as a result (e.g., for closing a box: *Hat das Mädchen 'se zugemacht*? (has the girl it closed-made 'Did the girl close it?'), Schulz & Wittek, 2003).

The crucial condition in all studies is the non-culminating situation. Non-culmination involved a partial result in all except one study: the event progressed up to a certain point and the object was partially affected (a half-eaten apple;

a partially opened box). In one study, a zero result was shown: the action did not affect the object; there was no change at all in the object (e.g., for the waking-up item, the man was still fast asleep, Wittek, 2002). Table 1 presents the details of the telicity studies. Six studies contrasted transitive clauses with and without verb particles in three Germanic languages; three additional studies only tested particle verbs; and one study focused on sentences with and without the resultative clitic *se* in Spanish.

Table 1. Overview of nine telicity studies included in review on telicity, specifying language, age range, and conditions: types of situations and types of test sentences[3]

Study	Language	Ages	Situations	Test sentences
van Hout 1998	Dutch	3–5	Full vs. Partial result	(1) Transitive & (2) Particle
van Hout 1998	English	3–5	Full vs. Partial result	(1) Transitive & (2) Particle
Jeschull 2007	English	3–6	Full vs. Partial result	(1) Transitive & (2) Particle
Ogiela 2007	English	3–6	Full vs. Partial result	(1) Transitive & (2) Particle
Schulz & Penner 2002	German	4–5	Full vs. Partial result	(1) Transitive & (2) Particle
Schulz et al. 2001; Penner et al. 2003	German	2–4	Full vs. Partial result	Particle
Schulz & Wittek 2003	German	4–5	Full vs. Partial result	Particle
Wittek 2002	German	4–5	Full vs. Zero result	(1) Transitive & (2) Particle
Hodgson 2001	Spanish	4–11	Full vs. Partial result	(1) Transitive & (2) Clitic *se*

Table 2 presents sample test sentences in the two conditions from these studies. The tense-aspect form used in the test sentences varied across languages, but was always a form with perfective semantics, asserting that the event ended: the present perfect (Dutch, German), simple past (English) and perfective past (Spanish preterito indefinido).

3. Ogiela's (2007) dissertation presents data from adults and children; Ogiela et al. (2014) focuses on the adult results. Extending the work by Schulz et al. (2001), Penner et al. (2003) present a wider range of participant groups.

Table 2. Sample test sentences in two conditions in nine telicity studies: transitive sentences with morphologically simple verbs and transitive sentences with particle verbs, or – in Spanish – with the reflexive clitic *se*. Two studies tested only particle verbs.

Study	Sample test sentence
van Hout 1998	Heeft het paard zijn appel (op)gegeten? (Did the horse eat (up) his apple?)
van Hout 1998	Did the horse eat (up) his apple?
Jeschull 2007	Who drank his coke (up)?
Ogiela 2007	Did the woman push the dogs (over)?
Wittek 2002	Hat das Mädchen den Mann geweckt / wachgemacht? (Did the girl wake (up) the man?)
Schulz-Penner 2002	Hat das Mädchen den Apfel (auf)gegessen? (Did the girl eat the apple?)
Schulz et al. 2001	Hat sie 'se aufgemacht? (Did she open it?)
Schulz-Wittek 2003	Hat das Mädchen 'se zugemacht? (Did the girl close it?)
Hodgson 2001	¿Quien (se) comió la madalena? (Who ate the cupcake?)

Figures 1 and 2 show how often children and adults accepted incomplete situations for telic sentences as reported in the studies above. Four- and 5-year-olds have been grouped together into one age group in some studies while other studies distinguished them as separate groups. Since the latter did not report any major age differences, I summarize here the results of the 5-year-olds, referring to the individual studies for more results of younger and older groups. For each study the pairs of bars compare the transitive condition with morphologically simple verbs versus the particle verb condition in the Germanic languages, and transitives with and without clitic *se* in Spanish. Comparing these nine telicity studies I will now discuss the following issues: rate of acceptance of non-culmination for telic-perfective sentences; age effects; variation across verbs; the form of the direct object.

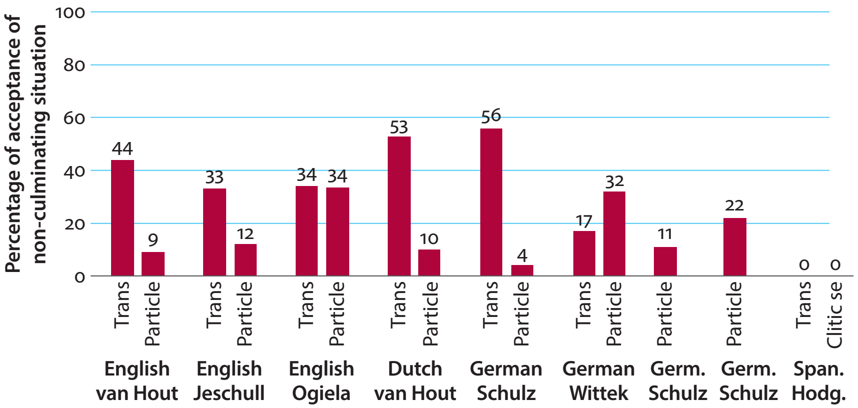


Figure 1. Group means of acceptance of partial or zero result situations by 5-year-olds in nine telicity studies

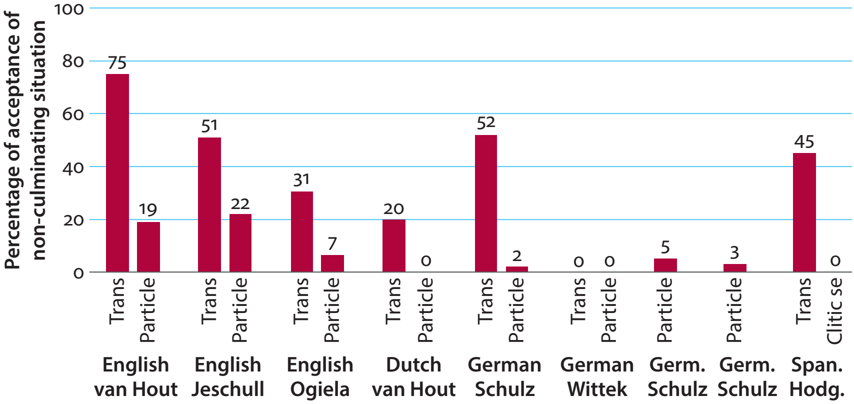


Figure 2. Group means of acceptance of partial or zero result situations by adults in nine telicity studies

The first and most obvious observation is that acceptance of non-culmination *does* exist, and sometimes it is even quite high (75%). This holds across studies, languages and ages. Non-culminating interpretations are therefore not some experimental fluke, but a fact in need of a semantic explanation. There is one clear pattern across studies and languages: there is much less acceptance in the particle and clitic *se* conditions than in the transitive conditions with a morphologically simple verb, for children as well as adults. In Dutch, German and Spanish, adult participants always rejected the particle verb or reflexive *se* clauses for incomplete situations. The different patterns across morphologically simple verbs versus particle verbs support my earlier generalization that predicate telicity is easier to acquire than compositional telicity (van Hout, 2008a). Nevertheless, even particle verbs were accepted for incomplete situations, sometimes to quite a large extent: by English, Dutch and German children (around 10% in Van Hout, 1998; up to over 30% in Ogiela, 2007 and Wittek, 2002) and even by English adults (up to 22% in Jeschull, 2007). This latter finding is unexpected in a scalar theory of telicity since particles encode the maximal end of a scale.

Perhaps surprisingly, there do not seem to be any consistent differences between children and adults across studies. Dutch and German children accepted telic predicates as descriptions of non-culminating situations more often than adults (van Hout, 1998; Schulz et al., 2001; Schulz & Wittek, 2003; Wittek, 2002), whereas in other studies the rate of acceptance was similar (Ogiela, 2007, for English, Schulz & Penner, 2002 for German), or even reverse (Jeschull, 2007, for English; Hodgson, 2001 for Spanish).

Another observation is the variation in acceptance of non-culmination across studies. This is undoubtedly related to the choice of verbs in the test sentences: different verbs triggered different rates of non-culmination acceptance. The studies tested different verbs, mostly incremental theme verbs (including *eat, drink, build, fix*) and change-of-state verbs (including *open, close, break, wake up*). The verbs are listed exhaustively in Appendix B.

Only a few studies discuss variation across verbs (the others do not mention item effects). Wittek (2002) established two patterns among the German particle verbs. For six verbs (German 'close', 'open', 'pick', 'extinguish', 'kill' and 'break'), the zero result situation was mostly rejected. But for two other verbs (German 'fill' and 'wake'), it was mostly accepted: 100% for *ein Glass vollmachen* 'fill a glass'; 60% for *einen Mann wachmachen* 'wake up a man'. Without these two verbs, the percentage of non-culmination acceptance drops from 32% to 13%, a number more similar to the German children and Dutch and English children in the other studies. Wittek notes that the 'glass filling' movie may have been problematic, adding that over-acceptance may also be due to the durativity of the event (the glass filling up little by little as an incremental theme) in contrast to the other items that were punctual change-of-state verbs. She does not offer any suggestions for the acceptance of *wachmachen* 'wake up', a transparent particle verb which she expected to lexically encode an end state.

In contrast to Wittek (2002), Ogiela (2007) and Ogiela et al. (2014) did not find item effects for their particle verbs (*eat up, drink up, push over, carry over*). Nevertheless, Ogiela and colleagues established variation in acceptance of non-culmination for simple, transitive verbs (*eat, drink, build, fix*). Even though all four simple transitive verbs were durative, incremental theme verbs, they had been categorized into two classes in the experimental design. In addition to a telic interpretation, *eat* and *drink* also allow a partitive construal (meaning 'eat/drink of part of NP'), in which case they are atelic and allow non-culmination. *Build* and *fix*, on the other hand, are not partitive. Given this flexibility in telicity, Ogiela and colleagues expected, and indeed found, more acceptance of non-culmination for *eat* and *drink* than for *build* and *fix*. Nevertheless, there was a significant contrast with two inherently atelic verbs in a control condition: *eat* and *drink* were accepted less often for non-culminating situations than *push* and *carry*. So, even though *eat, drink, build* and fix are "equally" telic according to semantic theory, given that they were used in a transitive sentence with a quantized object, there turned out to be a scale of non-culmination acceptance.

One further telicity study, Hacohen (2010) on Hebrew, also established item effects across different verbs. This study was not included in the tables above because it used a different paradigm, and moreover, the children were much older than those in the studies discussed above: school-age children between 7 and 12. Hebrew past tense is not perfective, and so telic sentences in the past tense do not necessarily entail culmination. Instead, Hacohen used the infinitive verb form embedded under ‘tell’, for example, ‘I told Tara to paint the square’. Participants watched video clips to see if the girl did as she had been told. An item analysis revealed that the verbs split into two classes, at least for the adults (in the condition with definite singular objects): acceptance of non-culmination varied between 11–33% for the three incremental-theme verbs (Hebrew ‘paint’, ‘draw’ and ‘peel’), whereas the incomplete situation was always rejected for the two change-of-state verbs (Hebrew ‘close’ and ‘empty’).

The variation across verbs as reported in these three studies – with different patterns of non-culmination acceptance for different verbs – is unexpected in a mereological theory of telicity based on quantization. It can, however, be explained using a scalar approach, as this framework presents a more “fluid” definition of telicity by incorporating effects of verb class. Verbs with a closed scale are unambiguously telic. For verbs with an open scale, on the other hand, the scalar theory offers room for ambiguity and/or underspecification of the endstate. This contrasts with a mereological approach (Krifka, 1989), in which telicity is essentially a fixed notion – defined by quantization and the notions of homogeneity and cumulativity. A mereological theory does not allow for variation across different verb classes, and so verb class effects are unexpected. The finding that predicate telicity is acquired early can be explained in the scalar approach though: for verbs that explicitly encode the closed end of the scale in the form of a particle, children do not (or much less) accept non-culmination. The challenge for learners are therefore verbs that do include an endstate in their lexical specification but do not explicitly (or transparently) encode this meaning component. This is indeed the class that takes longer to be acquired (Wittek, 2002). Moreover, several studies report differences between change-of-state verbs on the one hand and incremental theme verbs on the other (García del Real, 2015; Hacohen, 2010; Ogiela, 2007; Ogiela et al., 2014), with (much) lower rates of non-culmination acceptance for change-of-state verbs. This can be explained within a scalar approach by arguing that incremental-theme verbs involve an open scale, whereas change-of-state verbs involve a closed scale.

A final point of discussion in the telicity studies is the form of the direct object in the test sentences. In most studies the object was a singular NP. Some used test sentences with a cliticized pronoun (Schulz et al., 2001, 2003; Penner et al., 2003); others a definite NP (Hodgson, 2001; Ogiela, 2007; Ogiela et al., 2014, Schulz & Penner, 2002; Wittek, 2003); and yet others an NP possessive pronoun ‘his’ (van Hout, 1998; Jeschull, 2007), as illustrated in Table 2. It is hard to draw any conclusions about object forms across studies, but a few studies systematically varied the direct object and found effects of type of object on acceptance of non-culmination.

Ogiela (2007) and Ogiela et al. (2014) used two kinds of plural NPs – with a definite determiner (*eat the brownies*) or a cardinal number (*eat two brownies*). In the incomplete condition in Ogiela's (2007) study, the actor engaged in a certain kind of action with two object referents, one after the other. For the first referent, the event was complete, and for the other one it was incomplete, for example, eating one brownie completely and the other halfway. Ogiela and colleagues found an interaction with verb type: for the flexible telic verbs *eat* and *drink*, there was less non-culmination acceptance in the conditions with cardinal determiners than with definite determiners. There were no effects for the other verb types though, nor for the particle verbs. The researchers argue that, given that a cardinal number (*two)* more clearly specifies the quantized amount than a definite determiner, it triggers a culminating interpretation more strongly; the latter can also be used on a partitive reading.

The main goal of Hacohen's (2010) Hebrew study was to systematically investigate the effects of different types of direct objects for non-culmination, varying three grammatical dimensions of the object NP: singular/plural ('paint the square / paint the squares'), definite/indefinite ('paint the square / paint a square') and mass/count ('paint cloth / paint a square'). The design had six conditions, four of which were telic (indefinite singular count; definite singular count; definite plural count; definite mass) and two atelic (indefinite mass (bare); indefinite plural (bare)). Overall, children and adults rejected incomplete situations in the telic conditions with count terms: both definite and indefinite noun phrases triggered rejection, and so did singular and plural nouns. This shows that the definite/indefinite and singular/plural dimensions are irrelevant for telicity, as expected by theories on quantization and telicity (Krifka, 1989; Verkuyl, 1993). On the other hand, there was a main effect of the count/mass distinction: children as well as adults accepted the incomplete situation for mass terms and bare plural objects. This suggests that, despite an overall higher acceptance rate of non-culmination in children, the relevance of count/mass for compositional telicity had been acquired.

To summarize, I have provided a comprehensive review of (i) telicity studies that compared presence and absence of particles in the Germanic languages Dutch, German and English, and the telic-reflexive clitic *se* in Spanish, and (ii) studies that varied direct object forms with different quantization properties in English and Hebrew. One conclusion is that predicate telicity, as encoded by particles in the three Germanic languages and reflexive clitic *se* in Spanish, is easier to acquire than compositional telicity, confirming my earlier generalization (van Hout, 2008a). Nevertheless, even particle verbs did not always lead to culmination interpretations; there was some non-negligible acceptance of non-culmination for particle verbs (varying between 0–35%). Moreover, there were individual verb effects: some trigger more non-culmination acceptance than others. Age did not play a consistent role across the studies. As for the aspectual role of direct object quantization, given the variation in different kinds of objects (nouns and determiners) across different studies, in most cases with no systematic comparison, it is premature to draw any conclusions about the effect of direct object quantization on the acquisition of telicity. This remains an under-investigated area, especially with respect to the count/mass distinction.

5. Non-culmination patterns and the acquisition of perfectivity

The culmination entailment of telic-perfective sentences is jointly determined by the interaction of telicity and perfectivity. Children's interpretations are not sufficiently restrictive – they occasionally accept incomplete situations for telic-perfective sentences – which can, in principle, be caused by non-target-like knowledge of either one. So the question is: does acceptance of non-culmination arise from a non-adult-like semantics of telicity (or, at least, certain telic constructions), or rather from a non-adult-like semantics of perfective aspect? If children do not yet know the target semantics of perfective markers, acceptance of non-culmination could indicate incomplete acquisition of grammatical aspect, rather than telicity (lexical aspect). In other words, is there a developmental stage in which "perfective" markers are not yet perfective for children?

In order to determine what children know of perfectivity, I move on to review comprehension studies on the acquisition of grammatical aspect with experimental designs systematically contrasting perfective and imperfective aspect and using telic predicates. There is quite an extensive literature on this topic (for reviews, see van Hout, 2016, and Wagner, 2012). The first comprehension study on the perfective-imperfective distinction was Weist, Wysocka & Lyytinen (1991). Weist and colleagues found that English and Polish children appropriately differentiated the two aspects as early as 2;6. The Finnish children in Weist et al. (1991) did not reliably differentiate partitive and accusative case on the direct object in the aspect task until the age of 6;6 – a finding they explain by appealing to the one-to-many form/meaning mapping of the Finnish object cases.

Using the same task, albeit with somewhat different materials, however, a study by Wagner (2002) did not replicate the findings of Weist et al. (1991) regarding early mastery of the perfective-imperfective distinction: children did not perform at adult-like levels until the age of 5 years. Wagner attributes the age of acquisition gap between her results and those of Weist and colleagues to children's initially more limited knowledge of aspect as relating to the intentions of the agent of the action. The pictures in her picture-selection task only showed the resulting outcome for the object referent with no agent in view, whereas in Weist et al.'s study the pictures showed showed a resulting state for both the object referent (a finished versus an incomplete flower) and the agent (not drawing any more versus still drawing). Wagner concludes that, for young children, information about the state of the agent is crucial for deciding about culmination.

All other acquisition studies – using pictures or movies showing agent and object referents – in several other languages established early differentiation of the perfective-imperfective distinction. In a review about the acquisition of aspect (van Hout, 2016) I summarized this literature. See Appendix C for an overview of comprehension studies on grammatical aspect that compared acquisition of perfective and imperfective aspect in ten languages, specifying the age range tested and the youngest age at which children systematically differentiated the perfective and imperfective condition and showed knowledge of the culmination entailment of perfective aspect. For reasons of space I do not discuss any further details here about verbs and aspect forms (but see van Hout, 2016). The overall conclusion of my previous review is that children acquire the perfective-imperfective distinction early. Specifically, in most studies, the target meaning of the perfective markers in the relevant language is in place at the earliest age tested. A number of the studies listed here did not, however, find fully target-like interpretation and production of *im*perfective aspect up to the age of 6. They did, however, differentiate the two aspects. Seeing, then, that children know the target semantics of perfective markers, I conclude that the cause for the acceptance of non-culmination is *not* incomplete acquisition of grammatical aspect.

This brings us back to telicity and the question: is there a developmental stage at which telic predicates are not telic for children, or certain kinds of telic predicates are not telic? The review of the telicity acquisition patterns in Section 4 revealed that there is a distinction within the class of telic predicates: predicate telicity, as encoded by particles in the Germanic languages and telic-reflexive clitic *se* in Spanish, is easier to acquire than compositional telicity. The fact that child participants are able to reject non-culminating situations for perfective sentences with particle verbs leads to the following conclusion: children know the interaction of telicity and perfectivity (i.e., they compute the culmination entailment of telic-perfective predicates, at least for particle verbs and telic-reflexive clitic *se*). Whichever aspectual theory one assumes for modeling this interaction (Section 2), the mere fact that children are adult-like with certain telic-perfective predicates suggests that they have acquired the interplay of telicity and perfectivity in their grammars. Hence, the cause of the overly liberal non-culmination pattern is *not* immature knowledge of the telicity-perfectivity interface.

Having discussed the roles of particles, direct objects and perfective markers in acquisition, and concluding that none of these can explain acceptance of non-culmination, there is one remaining possible cause: the relevant event type properties encoded in a verb's lexical meaning. Is there a developmental stage at which lexically telic verbs (those that are not overtly encoded as such like particle verbs) are not yet telic for children? In other words, the meanings of verbs such as *cross*, *kill*, *destroy*, and *break* crucially include a result state. Can it be that these verbs, which are telic in the adult lexicon, are not (yet) telic for children?

6. Conclusions and new developments in the acquisition of event culmination

This chapter has presented a comprehensive overview of acquisition work on event culmination in three lines of research, showing that there is quite a high rate of acceptance of non-culmination by children (and adults too). I have sought to explain why, across languages, children accept non-culminating interpretations of telic-perfective sentences. Defining the elements involved in the grammar of culmination – the lexical semantics of verbs, telicity at the verb phrase level, perfective aspect at the sentence level – and discussing several possible explanations, I have concluded that the learnability challenge for event culmination, once perfectivity has been acquired, lies in the acquisition of verb meanings.

Several acquisition studies have pointed out the role of agentivity and intentionality for children's understanding of event culmination, but none have systematically varied these dimensions in the experimental designs. Interestingly, these insights from acquisition are being met by recent developments in crosslinguistic semantic theories about non-culmination in certain (adult) languages, including Mandarin, Japanese, Hindi, Thai and Salish languages (see Demirdache & Martin, 2015, for an overview). It turns out that there is a crucial role for the type of subject: whether the subject is a volitional and intentional Agent or rather a non-volitional and non-intentional natural Cause. Demirdache and Martin (2015) show that sentences with Agents sometimes allow for acceptance of non-culminating situations in these adult languages, but not Causes, revealing a crucial role for Agent control and event culmination.

This insight presents a novel angle on past acquisition studies and shapes new questions for future research, since all previous acquisition studies used Agent subjects. The question is hence, what happens with non-intentional Agents? And with Causer subjects? Demirdache and Martin's (2015) Agent Control hypothesis predicts that using other types of subjects in stimuli will lead to a lower rate of non-culmination acceptance. This novel line of inquiry has just entered acquisition research. Recent studies in various child languages revisit the non-target-like acceptance of non-culmination by manipulating subject-type: Agent versus Causer, in order to find out if non-culmination in child language reflects the Agent Control Hypothesis (van Hout et al., 2017; Strangmann, 2015).

Research on the grammar and acquisition of event culmination over the past few decades has been done within different, more or less independent lines of research. Various linguistic theories have been proposed: mereological and scalar approaches, pragmatics and lexical-semantics. Acquisition studies have investigated verb meanings, telicity and perfectivity. This chapter has tried to connect these different strands in order to reach deeper insights which can pave the way to further research. Given the important role of verb semantics for event culmination in adult grammars, there are several questions for future acquisition research. Which features of verb semantics contribute to different patterns in the acquisition of event culmination, and how? Moreover, given the dual source for deriving culmination in a scalar approach (semantic entailment versus pragmatic implicature), the acquisition question splits as well: How do children acquire event culmination when it is an entailment, and how do they do acquire it when it is an implicature? In other words, do children know that verbs from certain classes yield culminating interpretations more robustly than verbs from other classes? And how does this vary across languages? These questions and others can be approached by systematically comparing verbs from different event-semantic classes and further developing event culmination paradigms such as the ones that have been employed in the telicity and aspect studies surveyed in this chapter.

Acknowledgements

I gratefully acknowledge support from the Netherlands Organization for Scientific Research NWO (GraMALL, Grasping Meaning across Languages and Learners, 2014–2017). I thank my European colleagues in the GraMALL group for many inspiring discussions: Maria Arche, Fabrizio Arosio, Hamida Demirdache, Francesca Foppolo, Isabel García del Real, Fabienne Martin, Nina Kazanina, Jinhong Liu, Oana Lungu, Iris Strangmann. I thank Petra Schulz and Aviya Hacohen for their insightful reviews and Sudha Arunachalam and Kristen Syrett for being wonderful editors.

References

Atlas, J. D., Levinson, S. C. (1981). It-clefts, informativeness, and logical form: Radical pragmatics. In P. Cole (Ed.), *Radical pragmatics* (pp. 1–61). New York, NY: Academic Press.

Borik, O. (2006). *Aspect and reference time*. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199291298.001.0001

Brisson, C. (1994). The licensing of unexpressed objects in English. In K. Beals, J. Denton, R. Knippen, L. Melnar, H. Suzuki, & E. Zeinfeld (Eds.), *Papers from the 30th Annual Regional meeting of the Chicago Linguistic Society* (pp. 90–102). Chicago, IL: Chicago Linguistic Society.

Comrie, B. (1976). *Aspect*. Cambridge: Cambridge University Press.

Dahl, Ö. (1985). *Tense and aspect systems*. Oxford: Blackwell.

Delidaki, S. (2006). *The acquisition of tense and aspect in child Greek* (Unpublished doctoral dissertation). University of Reading.

Demirdache, H., & Uribe-Etxebarria, M. (2000). The primitives of temporal relations. In R. Martin, D. Michaels, & J. Uriagereka (Eds.), *Step by step: Essays on minimalist syntax in honor of Howard Lasnik* (pp. 157–186). Cambridge, MA: The MIT Press.

Demirdache, H., & Martin, F. (2015). Agent control over non-culminating events. In E. Barrajón López, J. L. Cifuentes Honrubia & S. Rodríguez Rosique (Eds.), *Verb classes and aspect* (pp. 185–217). Amsterdam: John Benjamins. https://doi.org/10.1075/ivitra.9.09dem

Dowty, D. (1979). *Word meaning and Montague grammar*. Dordrecht: Reidel. https://doi.org/10.1007/978-94-009-9473-7

Filip, H. (2008). Events and maximalization: The case of telicity and perfectivity. In S. Rothstein (Ed.), *Theoretical and crosslinguistic approaches to the semantics of aspect* (pp. 217–256). Amsterdam: John Benjamins. https://doi.org/10.1075/la.110.10fil

Gagarina, N. (2008). *Stanovlenie Grammatičeskich Kategorij Russkogo Glagola v Detskoj Reči* [Entstehung der grammatischen Kategorien des russischen Verbs im Erstspracherwerb]. St. Petersburg: Nauka.

García del Real, I. (2015). *The acquisition of tense and aspect in Spanish* (Unpublished doctoral dissertation). University of the Basque Country, Vitoria-Gasteiz.

García del Real, I., & Ezeizabarrena, M. (2011). Comprehension of grammatical and lexical aspect in early Spanish and Basque. In S. Ferré, P. Prévost, L. Tuller, & R. Zebib (Eds.), *Selected proceedings of the Romance Turn IV: Workshop on the acquisition of Romance languages* (pp. 82–103). Newcastle upon Tyne: Cambridge Scholars.

Gentner, D. (1978). On relational meaning: The acquisition of verb meaning. *Child Development*, 49, 988–998. https://doi.org/10.2307/1128738

Gropen, J., Pinker, S., Hollander, M., & Goldberg, R. (1991). Syntax and semantics in the acquisition of locative verbs. *Journal of Child Language*, 18, 115–151. https://doi.org/10.1017/S0305000900013325

Hacohen, A. (2010). *On the acquisition of Hebrew compositional telicity* (Unpublished doctoral dissertation). Ben-Gurion University of the Negev. Doctoral dissertation notice in *Language Acquisition*, 18, 84–86.

Harnish, R. M. (1976). Logical form and implicature. In T. G. Bever, J. J. Katz, & D. T. Langendoen (Eds.), *An integrated theory of linguistic ability* (pp. 313–391). New York, NY: Crowell.

Hay, J., Kennedy, C., & Levin, B. (1999). Scale structure underlies telicity in 'degree achievements'. In T. Matthews & D. Strolovitch (Eds.), *Semantics and Linguistic Theory 9*, (pp. 127–144). Ithaca, NY: CLC Publications.

Hirsh-Pasek, K., & Golinkoff, R. M. (2006). *Action meets word: How children learn verbs*. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195170009.001.0001

Hodgson, M. (2001). The acquisition of Spanish reflexive *se* in transitive verb-argument structures and singularity. In B. Skarabela, S. Fish, & A. H-J. Do (Eds.), *BUCLD 26: Proceedings of the 26th annual Boston University Conference on Language Development* (pp. 302–313). Somerville, MA: Cascadilla Press.

Hodgson, M. (2003). The acquisition of Spanish perfective aspect: a study on children's production and comprehension. *ZAS Papers in Linguistics* 29, 2003, 105–117.

van Hout, A. (1996). *Event semantics of verb frame alternations: A case study of Dutch and its acquisition* (Doctoral dissertation). University of Tilburg. Published in 1998 by Garland Publishing, New York in Outstanding Dissertations in Linguistics series.

van Hout, A. (1998). On the role of direct objects and particles in learning telicity in Dutch and English. A. Greenhill, M. Hughes, & H., Littlefield (Eds.), *BUCLD 22: Proceedings of the 22nd annual Boston University Conference on Language Development* (pp. 397–408). Somerville, MA: Cascadilla Press.

van Hout, A. (2005). Imperfect imperfectives: On the acquisition of aspect in Polish. In P. Kempchinsky & R. Slabakova (Eds.), *Aspectual inquiries* (pp. 317–343). Dordrecht: Springer. https://doi.org/10.1007/1-4020-3033-9_14

van Hout, A. (2007). Acquisition of aspectual meanings in a language with and a language without morphological aspect. In H. Caunt-Nulton, S. Kulatilake, & I. Woo (Eds.), *Supplemental Proceedings of the 31st annual Boston University Conference on Language Development*. <http://www.bu.edu/linguistics/APPLIED/BUCLD/supp31.html>

van Hout, A. (2008a). Acquiring telicity cross-linguistically: On the acquisition of telicity entailments associated with transitivity. In M. Bowerman & P. Brown (Eds.), *Crosslinguistic perspectives on argument structure: Implications for learnability* (pp. 255–278). New York, NY: Routledge.

van Hout, A. (2008b). Acquisition of perfective and imperfective aspect in Dutch, Italian and Polish. *Lingua*, 118, 1740–1765. https://doi.org/10.1016/j.lingua.2007.08.011

van Hout, A. (2016). The acquisition of aspect. In J. Lidz, W. Snyder, & J. Pater (Eds.), *The Oxford handbook of developmental linguistics* (pp. 587–610). Oxford: Oxford University Press.

van Hout, A., Arche, M., Demirdache, H., García del Real, I., García Sanz, A., Gavarró, A., Gomez Marzo, L., Hommes, S., Kazanina, N., Liu, J., Lungu, O., Martin, F., & Strangmann, I. M. (2017). Agent control and the acquisition of event culmination in Basque, Dutch, English, Spanish and Mandarin. In M. LaMendola & J. Scott (Eds.), *BUCLD 41: Proceedings of the 41st annual Boston University Conference on Language Development* (pp. 323–332). Somerville, MA: Cascadilla Press.

van Hout, A., & Hollebrandse, B. 2001. On the acquisition of the aspects in Italian. J.-Y. Kim and A. Werle (Eds) *The proceedings of SULA, The semantics of under-represented Languages in the Americas*. UMOP 25, 111–120. GLSA, Amherst.

Ikegami, Y. (1985). Activity – accomplishment – achievement – a language that can't say 'I burned it, but it didn't burn' and one that can. In A. Makkai & A. K. Melby (Eds.), *Linguistics and philosophy: Essays in honor of Rulon S. Wells* (pp. 265–304). Amsterdam: John Benjamins. https://doi.org/10.1075/cilt.42.21ike

Jeschull, L. (2007). The pragmatics of telicity and what children make of it. In A. Belikova, L. Meroni, & M. Umeda (Eds.), *Proceedings of the 2nd Conference on Generative Approaches to Language Acquisition – North America* (pp. 180–187). Somerville, MA: Cascadilla Proceedings Project.

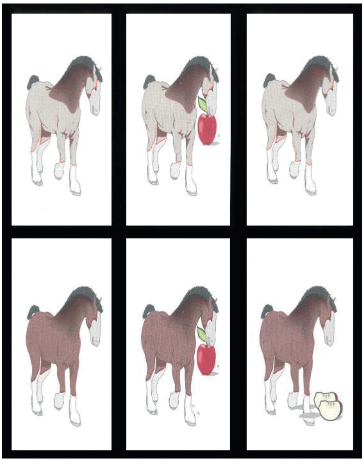
Kamp, H., & Reyle, U. (1993). *From discourse to logic*. Dordrecht: Kluwer.
Kazanina, N., & Phillips, C. (2007). A developmental perspective on the imperfective paradox. *Cognition*, 105, 65–102. https://doi.org/10.1016/j.cognition.2006.09.006
Kennedy, C., & Levin, B. (2008). Measure of change: The adjectival core of degree achievements. In L. McNally & C. Kennedy (Eds.), *Adjectives and adverbs: Syntax, semantics and discourse* (pp. 156–182). Oxford: Oxford University Press.
Klein, W. (1994). *Time in language*. London: Routledge.
Konstantzou, K. (2014). *Acquisition of tense and aspect: Evidence from production and comprehension tasks* (Unpublished doctoral dissertation). University of Athens.
Krifka, M. (1989). Nominal reference, temporal constitution and quantification in event semantics. In R. Bartsch, J. van Benthem, & P. van Emde Boas (Eds.), *Semantics and contextual expressions* (pg. 75–115). Dordrecht: Foris.
Lakusta, L., & DiFabrizio, S. (2017). And, the winner is… A visual preference for endpoints over starting points in infants' motion event representations. *Infancy*, 22, 323–343.
Landman, F. (1992). The progressive. *Natural Language Semantics*, 1, 1–32. https://doi.org/10.1007/BF02342615
Li, P., & Bowerman, M. (1998). The acquisition of lexical and grammatical aspect in Chinese. *First Language*, 18, 311–350. https://doi.org/10.1177/014272379801805404
Matsuo, A. (2009). Young children's understanding of ongoing vs. completion in imperfective and perfective participle. *Linguistics*, 47, 743–757. https://doi.org/10.1515/LING.2009.024
Nishida, C. (1994). Spanish reflexive clitic *se* as an aspectual class marker. *Linguistics*, 32, 425–458. https://doi.org/10.1515/ling.1994.32.3.425
Ogiela, D. (2007). *Development of telicity interpretation: Sensitivity to verb-type and determiner-type* (Unpublished doctoral dissertation). Michigan State University.
Ogiela, D., Schmitt, C., & Casby, M. W. (2014). Interpretation of verb phrase telicity: Sensitivity to verb type and determiner type. *Journal of Speech, Language, and Hearing Research*, 57, 865–875. https://doi.org/10.1044/2013_JSLHR-L-12-0271
Pederson, E. (1998). Event realization in Tamil. In P. Brown & M. Bowerman (Eds.), *Crosslinguistic perspectives on argument structure: Implications for learnability* (pp. 331–355). Mahwah, NJ: Lawrence Erlbaum Associates.
Penner, Z., Schulz, P., & Wyman, K. (2003). Learning the meaning of verbs: What distinguishes language-impaired from normally developing children? *Linguistics*, 41, 289–319. https://doi.org/10.1515/ling.2003.010
Rothstein, S. (2004). *Structuring events*. Oxford: Blackwell. https://doi.org/10.1002/9780470759127
Schulz, P. (2018). Telicity in typical and impaired acquisition. In K. Syrett & S. Arunachalam (Eds.), *Semantics in Language Acquisition* (pp. 123–150). Amsterdam: John Benjamins.
Schulz, P., & Penner, Z. (2002). How you can eat the apple and have it too: Evidence from the acquisition of telicity in German. In J. Costa & M. J. Freitas (Eds.), *Proceedings of the Generative Approaches to Language Acquisition Conference*. Lisbon: Associação Portuguesa de Linguística, Lisboa.
Schulz, P., & Wittek, A. (2003). Opening doors and sweeping floors: What children with Specific Language Impairment know about telic and atelic verbs. In B. Beachley, A. Brown, & F. Colin (Eds.), *BUCLD 27: Proceedings of the 27th annual Boston University Conference on Language Development* (pp. 727–738). Somerville, MA: Cascadilla Press.
Schulz, P., Wymann, K., & Penner, Z. (2001). The early acquisition of verb meaning in German by normally developing and language impaired children. *Brain and Language*, 77, 407–418. https://doi.org/10.1006/brln.2000.2411

Singh, M. (1994). *Perfectivity, definiteness, and specificity: A classification of verbal predicates in Hindi* (Unpublished doctoral dissertation). University of Texas at Austin.
Smith, C. (1991). *The parameter of aspect*. Dordrecht: Kluwer. https://doi.org/10.1007/978-94-015-7911-7
Soh, H. L., & Kuo, J. Y. (2005). Perfective aspect and accomplishment situations in Mandarin Chinese. In H. J. Verkuyl, H. de Swart, & A. van Hout (Eds.), *Perspectives on aspect* (pp. 199–216). Dordrecht: Springer. https://doi.org/10.1007/1-4020-3232-3_11
Stoll, S. (1998). Acquisition of Russian aspect. *First Language*, 18, 351–377. https://doi.org/10.1177/014272379801805405
Strangmann, I. (2015). *The effect of agents and subject causes on the interpretation of non-culminating endstates* (MA thesis). University of Groningen.
Talmy, L. (1991). Path to realization: A typology of event integration. In L. A. Sutton, C. Johnson, & R. Shields (Eds.), *Proceedings of the 17th Annual Meeting of the Berkeley Linguistics Society: General session and parasession on the grammar of event structure* (pp. 480–519). Available online: <http://linguistics.berkeley.edu/bls/previous_proceedings/bls17.pdf >
Tomasello, M., & Merriman, W. E. (1995). *Beyond names for things: Young children's acquisition of verbs*. New York, NY: Psychology Press.
Vendler, Z. (1957). Verbs and times. *The Philosophical Review*, 66, 143–160. https://doi.org/10.2307/2182371
Verkuyl, H. (1972). *On the compositional nature of the aspects*. Dordrecht: Reidel. https://doi.org/10.1007/978-94-017-2478-4
Verkuyl, H., (1993). *A theory of aspectuality*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511597848
Vinnitskaya, I., & Wexler, K. (2001). The role of pragmatics in the development of Russian aspect. *First Language*, 21, 143–186. https://doi.org/10.1177/014272370102106202
Wagner, L. (2002). Children's comprehension of completion entailments in the absence of agency cues. *Journal of Child Language*, 29, 109–125. https://doi.org/10.1017/S0305000901004949
Wagner, L. (2012). First language acquisition. In R. Binnick (Ed.), *The Oxford handbook of tense and aspect* (pp. 458–480). Oxford: Oxford University Press.
Wagner, L., & Lakusta, L. (2009). Using language to navigate the infant mind. *Perspectives on Psychological Science*, 4, 177–184. https://doi.org/10.1111/j.1745-6924.2009.01117.x
Weist, R. M., Wysocka, H., & Lyytinen, P. (1991). A cross-linguistic perspective on the development of temporal systems. *Journal of Child Language*, 18, 67–92. https://doi.org/10.1017/S0305000900013301
Wittek, A. (2002). *Learning the meaning of change of state verbs: A case study of German child language*. Berlin: De Gruyter Mouton. https://doi.org/10.1515/9783110907988

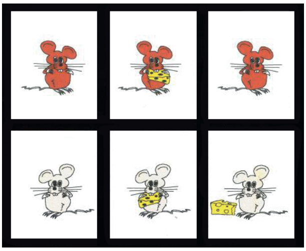
Appendix A. Samples of experimental materials for testing the culmination entailment

Below are some picture series and accompanying test sentences from the truth-value judgment task in van Hout (1998). The pictures were put down in front of the child one by one, while the experimenter told a short story about two animals or two people. In the middle picture the agent is eating or drinking, and in the final picture the agent has either finished his food or drink (culminating situation) or has left half of it (non-culminating situation). There were four conditions in this experiment; each is illustrated here.

Particle verb condition: *Did the light brown horse eat up his apple? Did the dark brown horse eat up his apple*?

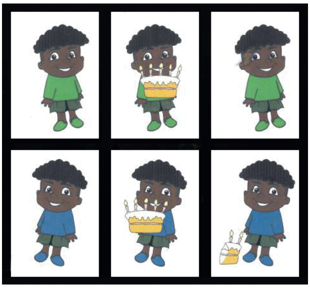


Transitive condition: *Did the red mouse eat his cheese? Did the white mouse eat his cheese*?



Transitive with mass term condition: *Did the boy with the white cap drink coke? Did the boy with the red cap drink coke*?

Intransitive condition: *Did the boy in green eat? Did the boy in blue eat*?



Appendix B. Overview of verbs tested in nine telicity studies

This table lists, per study and per language, all verbs that were tested in the transitive and particle verb conditions, and for Spanish in the transitive and clitic *se* condition.

Study	Language	Condition	Verbs
van Hout 1998	Dutch	transitive	eten 'eat', drinken 'drink'
		particle	opeten 'eat up', opdrinken 'drink up'
van Hout 1998	English	transitive	eat, drink
		particle	eat up, drink up
Jeschull 2007	English	transitive	eat, drink, fold, wrap
		particle	eat up, drink up, fold up, wrap up
Ogiela 2007	English	transitive	eat, drink, build, fix
		particle	eat up, drink up, push over, carry over
Schulz & Penner 2002	German	transitive	essen 'eat', trinken 'drink'
		particle	aufessen 'eat up', austrinken 'drink up'
Schulz et al. 2001; Penner et al. 2003	German	transitive	–
		particle	aufmachen 'open'
Schulz & Wittek 2003	German	transitive	–
		particle	aufmachen 'open', zumachen 'close', abmachen 'take off', anmachen 'turn on'
Wittek 2002	German	transitive	schliessen 'close', knacken 'crack', pflücken 'pick', löschen 'extinguish', füllen 'fill', töten 'kill', wecken 'wake', zerbrechen 'break'
		particle	zumachen 'close', aufmachen 'open', abmachen 'pick', ausmachen 'extinguish', vollmachen 'fill', totmachen 'kill', wachmachen 'wake', kaputmachen 'break'
Hodgson 2001	Spanish	transitive	limpiar 'clean', comer 'eat, beber 'drink', hacer 'do', leer 'read'
		particle	limpiarse 'clean', comerse 'eat', beberse 'drink', hacerse 'do', leerse 'read'

Appendix C. Overview of grammatical aspect acquisition studies

Reprint of Table 25.6 (p. 607) from Ch.25 'Lexical and Grammatical Aspect' by Angeliek van Hout in "Oxford Handbook of Developmental Linguistics", edited by Lidz, J. W. Snyder, W. & Pater, J. (2016). By permission of Oxford University Press (www.oup.com).

This is a summary of aspect comprehension studies in ten languages, copied from my review in van Hout (2016); some references have been updated. Age PERF: Youngest age at which children systematically differentiated perfective vs. imperfective aspect and show knowledge of the culmination entailment of perfective aspect.

Language	Study	Ages tested	Age PERF
Basque	García del Real & Ezeizabarrena, 2011	5;0–5;11	5[a]
Chinese	Li & Bowerman, 1998	3;11–6;4	5[b]
Finnish	Weist et al., 1991	2;6–6;6	–[c]
Dutch	van Hout, 2007, 2008b	3–5	3[d]
English	Matsuo, 2009	1;6–6;8	5[e]
	Wagner, 2002	1;11–5;7	5[e]
	Weist et al., 1991	2;6–6;6	2;6[f]
Italian	van Hout & Hollebrandse, 2001; van Hout 2008b	3;0–6;1	4
Spanish	García del Real, 2015;	5;0–5;11	5[a]
	Hodgson, 2003	3;0–8;0	5
Greek	Delidaki, 2006	3;0–6;5	3;0
	Konstantzou, 2014	4;0–6;11	4[a]
Polish	van Hout, 2005, 2008	2;0–4;11	3
	Weist et al., 1991	2;6–6;6	2;6[f]
Russian	Gagarina, 2008	3;0–6;11	3
	Kazanina & Philips, 2007	2;10–6;9	3
	Stoll, 1998	2;0–6;11	2;6
	Vinnitskaya & Wexler, 2001	3;0–6;5	3

Notes:

a This study did not test children younger than the indicated age.

b This reflects the data for their lexical verb class categories "Accomplishment Resultative" and "Locative".

c Not even the oldest Finnish learners at 6;6 interpreted accusative and partitive object case appropriately in order to distinguish complete versus incomplete situations.

d The Dutch 3-year-olds correctly rejected the perfect for incomplete situations by 63%, which was different from chance, but far from ceiling.

e The three studies with English learners establish different ages for the acquisition of the culmination entailment of perfective aspect (English *-ed* forms in simple past or in a prenominal participle). In contrast to Weist et al. (1991), Wagner (2002) and Matsuo (2009) only showed the state of the object referent at the end of the event, not the agent of the action. The presence of agent-oriented information may be required for the younger children to be able to succeed in this task (Wagner, 2002).

f Weist et al.'s results show that the English and Polish learners correctly differentiate the perfective-imperfective distinction at the age of 2;6. However, since no percentages correct are given for each aspect separately, we cannot tell from their report how firmly these young learners have acquired the culmination entailment of the perfective.

CHAPTER 6

Telicity in typical and impaired acquisition

Petra Schulz
Goethe-University Frankfurt

This chapter discusses children's knowledge of the syntactic-semantic interface of different types of telicity in typical and impaired acquisition. It maintains that telicity can be semantic or pragmatic, depending on whether event completion is entailed or implicated. It further delineates the *Event Structural Bootstrapping* account, arguing that *Endstate Orientation* – rather than a Manner Bias – guides children's acquisition of verb meanings. Findings from studies testing children's early verb production and comprehension are provided as evidence for a strong *Endstate Orientation* in typical development, but not in Specific Language Impairment. These data speak for modularity in the domain of semantics and for the presence of selective impairments in verb semantics. Cross-linguistic implications for further research are formulated in the conclusion.

Keywords: telicity, Specific Language Impairment (SLI), typical language development, comprehension, German, semantic impairment, verb (particle), verb semantics, event structural bootstrapping, Endstate Orientation

1. Introduction

How do children learn the meaning of verbs? Although this question has instigated much research for more than 30 years (e.g., Gentner, 1982), how the child succeeds in this task is still a matter of debate. Compared to object-referring nouns like *apple* or *door*, the lexical representation of a verb is complex and comprises several distinct components like core meaning, argument selection, and event type, all of which have to be acquired. Moreover, unlike in learning object labels, children cannot rely on joint visual attention in learning labels for actions or situations. Events are typically fleeting, and the production of the label for the event often does not coincide with the time at which the event occurs. Additionally, the relation between verb and event is underdetermined, for a verb usually refers to a specific aspect of a scene. The same event could, for example, be described as *running into the garden, entering the garden*, or *arriving home*. Finally, in contrast to object labels, verbs are

https://doi.org/10.1075/tilar.24.06sch

subject to considerable cross-linguistic variation in how the different event types are encoded in syntax and lexicon (e.g., English *He ran into the garden* vs. French *Il est entré dans le jardin en courant*).

Given these challenges, we might expect that verbs are difficult to learn. Words referring to events, however, occur very early in children's speech, and typically developing (TD) children master event structural concepts like telicity at a young age. This is in stark contrast to impaired acquisition: Children with Specific Language Impairment (SLI) (sometimes also referred to as Developmental Language Disorder, DLD) exhibit persistent difficulty with crucial aspects of telicity. Focusing on German, this chapter summarizes recent findings on children's knowledge of the syntactic-semantic interface of different types of telicity in typical and impaired acquisition. It starts with a summary of the concepts of semantic and pragmatic telicity in Section 2. Section 3 outlines the acquisitional perspective on telicity and introduces the *Event Structural Bootstrapping* account that argues that *Endstate Orientation* – rather than a Manner Bias – universally guides children's early acquisition of verb meanings. Section 4 discusses findings from children's early verb production. Section 5 describes findings from a set of comprehension studies testing semantic and pragmatic telicity in both TD and SLI children. The chapter ends with a conclusion and cross-linguistic implications for further research.

Related findings in Dutch and English, referred to in this chapter, suggest that the specific findings for German hold for other Germanic languages as well (see also van Hout, 2018). How the acquisition strategy of *Endstate Orientation* may be borne out in non-Germanic languages is briefly outlined in the conclusion.

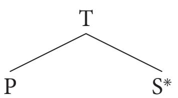
2. Semantic and pragmatic construal of telicity

Languages across the world use a common set of event components for their verbs. I suggest that the event component 'endstate', involved in telicity, plays a crucial role in children's acquisition of verb meaning, and that this initial preference for endstates is universal. As event components are encoded differently in typologically different languages (Talmy, 1991), I predict cross-linguistic variation to exist regarding how telic event expressions are initially realized by the child. This assumption is in line with Maguire et al. (2010), who hypothesize that children initially show common, possibly universal verb construal, and only later demonstrate language specific tendencies. In contrast to Maguire et al. (2010), who argue for an initial preference for Path interpretations, and to Gentner (1978), who argues for a Manner Bias in children's interpretation of verbs, I argue for an initial *Endstate Orientation*.

2.1 Event structure of verbs

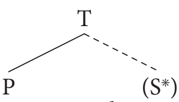
Verbs designate events including states such as ‘being happy’ and actions like ‘walking’, ‘sweeping’, or ‘opening’, that all differ regarding their internal temporal make-up (Dowty, 1979; Vendler, 1957). Events like ‘opening’ have an endpoint built into them leading to a natural culmination point, while events such as ‘being happy’ or ‘walking’ do not have an endpoint allowing the event to continue indefinitely or to stop at any moment in time. Verbs designating events with an endpoint are referred to as telic, and verbs designating events without such an endpoint are referred to as non-telic. Within Pustejovsky’s (1995) model of event typology, events can be further classified according to their complexity. States (S) like ‘being happy’ and processes (P) like ‘walking’ are simple events; they are referred to with atelic verbs. Complex events involve a transition (T) from a process to a state. Telic verbs like *open* designate so-called endstate-oriented transitions, i.e. the endstate subevent is the head-of-event, illustrated in (1a). To capture the intuition that some verbs such as *sweep* designate process-oriented transitions, in which the process subevent P is more prominent than the resulting state, Pustejovsky (1995) suggests that in this case P is marked as head-of-event. Diverging from this assumption, I postulate the structure in (1b). Like in endstate-oriented transitions, in process-oriented transitions the endstate is marked as head-of-event, but unlike in endstate-oriented transitions, in process-oriented transitions the endstate is optional. This structure reflects the insight that process-oriented transitions are in fact ambiguous between atelic processes and telic endstate-oriented transitions.

(1) a.



Endstate-oriented transition
open the door

b.



Process-oriented transition
sweep the floor

Telic interpretations of verbs can be construed semantically or pragmatically: Semantic telicity arises through entailment and pragmatic telicity though implicature (Arunachalam & Kothari, 2010; Filip, 2008; 2014; Hay, Kennedy, & Levin, 1999; Jeschull, 2007). The telic interpretation of verbs like *open*, which designate endstate-oriented transitions, arises through entailment. (2a) for example entails 'the present is open' and this entailment is not cancellable, as seen in (2b).[1]

(2) a. Jill opened the present.
b. Jill opened the present, #but it is still wrapped.

Telic interpretations of surface-contact verbs like *sweep* and of degree achievements such as *cool the beer* (Filip, 2008) arise through a generalized conversational implicature, as illustrated in (3). The implicature of (3a) that the flor is clean is cancellable, as seen in (3b).

(3) a. Jill swept the floor.
b. Jill swept the floor, but it is still not clean.

The temporal adverbial modification test (Dowty, 1979) confirms the different status of natural culmination in predicates like *open* vs. *sweep: in an hour/in a minute* combines with telic predicates, while the durational adverb *for hours/for minutes* combines with atelic predicates ((4a) vs. (4b)). *Open* is telic, whereas *sweep* has both a telic and an atelic reading.

(4) a. Jill opened the present in a minute/ * for minutes.
b. Jill swept the floor in an hour / for hours.

In summary, atelic predicates designate processes and states; semantically telic predicates designate endstate-oriented transitions like 'opening' and pragmatically telic predicates designate process-oriented transitions like 'sweeping'.

2.2 Encoding telicity in German: A note on verb particles and particle verbs

Languages differ as to how event types are encoded by syntax and in the lexicon. In German, like in English and Dutch, verb particles play a prominent role in marking telicity by contributing to the semantics of the complex verb containing the particle. We can distinguish between different basic types of verb particles regarding their telicity properties (for German, see Schulz, Wymann, & Penner, 2001, for English, see Walková, 2013). Telic verb particles (also referred to as resultative) mark the prominent endstate of a transition, like *auf* 'open' in *aufmachen* 'open' or *aus* 'out' in *ausmachen* 'turn off'. Note that in some cases complex verbs retain their atelic reading despite presence of a particle that is telic by default (e.g., *aussehen*, AUS.

1. All examples contain perfective aspect, since imperfective aspect such as progressive in English cancels the completion entailment (Dowty, 1979):

(i) Jill was opening the present; in fact she is still opening the present.

see ‘to look’). Atelic verb particles exclusively mark the process, like *rum* ‘around’ in *rumlaufen* ‘walk around’. Finally, ambiguous verb particles mark either the process or the endstate of a transition, like *rauf* ‘up’ and *runter* ‘down’ in *raufgehen* ‘walk up’ und *runtergehen* ‘walk down’. This corresponds to the observation by Levin and Rappaport Hovav (1995) that verbs of directed motion such as *ascend* and *descend* are not necessarily telic. Importantly, verb particles like *auf* occur as complex verb alternations of simplex verbs (e.g., *aufessen* ‘eat up’ vs. *essen* ‘eat’) or in particle verbs (e.g., *aufmachen* ‘open’), which lack a simple verb alternation.[2]

2.3 Inherent and compositional telicity

Like in English and Dutch, in German, (a)telicity of the verb is generally determined by its lexical semantics or via the interaction with the morpho-syntactic context the verb appears in (see van Hout, 1996, 1998, 2000, 2008). Inherently telic verbs such as *open*, *find*, and *arrive* designate events with a natural culmination point that is part of the verb meaning. These inherently telic verbs are all cases of semantic telicity (i.e. telicity arises through entailment, see Section 2.1). In contrast, compositional telicity of the predicate, resulting from the interaction of an atelic verb with the event-semantic properties of other morpho-syntactic elements in the sentence (Krifka, 1989, 1998), can either arise through entailment or through implicature. Among the elements triggering an event-type shift are directional phrases such as *walk into the house*, resultative phrases such as *laugh yourself silly*, complex verb formation as in particle verbs and presence of a ‘quantized’ NP[3] as in *eat the apple*. Recent work has shown that the type of subject (intentional or non-intentional agent) contributes to telicity as well (Demirdache & Martin, 2015). In the following I focus on those structures that have been prominent in acquisition research: predicates containing telic verb particles and/or quantized NPs, combined with intentional agents as the subject’s referents.

Telic verb particles trigger telicity via entailment. They have also been referred to as strong telicity markers, because adding them to an atelic process verb obligatorily results in an event-type shift to telic (Schulz et al., 2001; Penner, Schulz, & Wymann, 2003; van Hout, 1998). This is exemplified in (5) and (6) with the process verbs *essen* ‘eat’ and *zeichnen* ‘draw’ and their telic particle counterparts. In the telic particle verb sentences (5a) and (6a) it is infelicitous to cancel the entailment that the culmination point is reached. This is in contrast to their process counterparts (5b) and (6b).

2. Note that unlike their English counterparts, German particle verbs often do not have a clear adjectival base (see Hay et al., 1999). The adjective describing the endstate of *aufmachen* ‘open’ for example is *offen,* not *auf.*

3. A predicate P is quantized iff: if P(x) and y < x, then not P(y) (Krifka, 1998).

(5) a. Er hat auf-gegessen, (#aber es ist noch was übrig).
he has up-eaten. PST.PTCP (#but it is still some left)
'He ate it up, (#but there is something left).'
b. Er hat gegessen, (aber es ist noch was übrig).
he has eaten. PST.PTCP (#but it is still some left)
'He ate, but there is still something left.'

(6) a. Sie hat das Haus ab-gezeichnet,
she has the. ACC.SG house .ACC.SG off-drawn. PST.PTCP
(#aber es ist nicht fertig).
(#but it is not done)
'She copied the house, (#but it is not completely drawn.).'
b. Sie hat gezeichnet, aber es ist noch nicht fertig.
she has drawn. PST.PTCP but it is still not done
'She drew, but it is not yet finished.

Certain quantized NPs trigger telicity via implicature. They have also been referred to as weak telicity markers, because combination with a process verb does not always result in a telic predicate, as shown in (7) (=English (3b)). The implicated culmination point can be cancelled without rendering the sentences infelicitous, despite the presence of the quantized NP.

(7) a. Jill fegte den Boden, aber er ist nicht sauber.
Jill swept the. ACC.SG floor. ACC.SG but he is not clean
'Jill swept the floor, but it is not clean.'
b. Jill ging den Hügel hinauf, aber erreichte die
Jill walked the .ACC.SG hill .ACC.SG up but reached the
Spitze nicht.
top not
'Jill walked up the hill, but didn't reach the top.'

The role of quantized NPs for telicity with verbs of consumption is more complex. In earlier approaches (see Krifka, 1989) quantized NPs have been assumed to trigger semantic telicity with *eat* and *drink*. This is illustrated for *eat* with the contrast between the intransitive verb (8a) and the bare, non-quantized NP (8b) vs. the quantized NP (9c). The durational adverbial combines with the atelic predicates in (8a) and (8b), but only marginally with the specific definite NP in (8c).

(8) a. Sie hat (eine Stunde lang) gegessen.
she has (an hour long) eaten. PST.PTCP
'She ate (for an hour).'
b. Sie hat (eine Stunde lang) Käse gegessen.
she has (an hour long) cheese. ACC.SG eaten. PST.PTCP
'She ate cheese (for an hour).'

c. Sie hat (?eine Stunde lang) den Käse
she has (?an hour long) the. ACC.SG cheese. ACC.SG
gegessen.
eaten. PST.PTCP
'She ate the cheese (?for an hour).'

More recent approaches argue that verbs of consumption may select a quantized NP without automatically triggering an event type shift (Filip, 2014; Hayet al., 1999). This is illustrated in (9). This finding is crucial for acquisition studies, which have often used verbs of consumption with definite NPs as test items in experiments.

(9) a. Bill ate the apple bite by bite for ten minutes (and still didn't finish it) (Jackendoff, 1996: 308)
b. She ate the sandwich but as usual she left a few bites. (Hay et al., 1999: 139)

We hence conclude that certain quantized NPs, e.g. specific definite NPs, are indeed weak telicity markers, involving pragmatic telicity, at least when selected by verbs of motion, surface-contact, and consumption (see Section 5.3. for empirical data confirming this assumption for adults and children). Note that not all definite NPs behave alike. Measurement phrases, for example, entail telicity (e.g., #Mary ate two apples, but finished eating only one.). Based on the notions of entailment and implicature, Table 1 summarizes the different types of telicity.

Table 1. Types of telicity and their encoding in German

Type of telicity	Source	Encoding	Examples
Inherent	Entailment	Lexical semantics of the verb	Morphologically simple verbs (e.g., *finden* 'find'), morphologically complex verbs (e.g., *aufmachen* 'open')
Compositional	Entailment	Resultative verb particles (strong telicity marker)	auf- (e.g., *aufessen*, eat up'), aus- (e.g., *austrinken*, drink up'), ab- (e.g., *abzeichnen*, draw')
	Implicature	Certain quantized NPs (weak telicity marker)	Specific definite NPs (e.g., *den Apfel essen* 'eat the apple', *den Boden fegen* 'sweep the floor', *den Berg hinaufgehen* 'walk up the hill')

In summary, telicity in German is encoded either by the lexical semantic properties of the verb (semantic telicity) or formed compositionally, for example by adding a resultative particle or a quantized NP complement to an atelic verb. Telicity in resultative particles arises through entailment of the culmination point (semantic telicity), while telicity in certain quantized NPs arises through implicature of the culmination point (pragmatic telicity).

3. The acquisitional perspective

3.1 The acquisition task of mastering telicity

Given the complex relationship between the lexical semantics of the verb determining its basic event-semantic property and the morpho-syntactic contexts that may trigger event type shifting, the acquisition task of the child is manifold. In order to determine whether a predicate is telic or not, she has to know the lexical event-semantic properties of the individual verbs. That is, she has to distinguish between inherently telic verbs and verbs that can alter their event type. She also has to learn whether the natural culmination point of the event designated by a predicate is entailed or implicated (i.e. whether the telicity marker is weak or strong). Finally, she has to learn the function of the various morpho-syntactic elements in her language (e.g., verb particles and quantized NPs) and the mapping between these elements and their role in determining the predicate's event-type.

3.2 Learning strategy for verbs: Event Structural Bootstrapping

How can the child succeed in this complex acquisition task? In previous work my colleagues and I have proposed the *Event Structural Bootstrapping* account, suggesting that typically developing children, faced with the task of verb learning, initially focus on the verb's event structure rather than on its core meaning or its argument selection (Penner et al., 2003; Schulz et al., 2001; Schulz, Penner, & Wymann, 2002). We have argued that TD children proceed in a 'piecemeal' fashion in order to overcome the input ambiguity in the domain of the verb lexicon: They first focus on event expressions that encode endstate-oriented transitions and only later on the other event types (processes, states, and process-oriented transitions). That is, children are argued to initially show an *Endstate Orientation* guiding their early learning of verb meaning (Schulz et al., 2002). Endstate-oriented transitions are a safe starting point for learning verb meaning because their event structure is unambiguous. Inherently telic verbs like *open* only designate events with their natural culmination point reached, unlike process verbs like *sweep* that are used with events where a culmination point has been reached or not, and unlike processes and states, which lack a culmination point altogether (see Section 2.1).

Within the class of telic predicates, particle verbs of the type *aufmachen* (AUF.make, 'open') and *zumachen* (ZU.make, 'close') meet the requirement of an unambiguous event structure best, since the internal hierarchy of the transition is optimally transparent. The resultative particle *auf* unambiguously marks the endstate as head-of-event, and the dummy light verb *machen* 'make' lexically marks the process subevent as less prominent.

Supporting evidence for children's early strategy of *Endstate Orientation* comes from studies by Lakusta and colleagues. Infants' (age 16 months) preference for looking at endpoint events over starting point events (Lakusta & DiFabrizio, 2016), as well as their bias to encode goals over the source when describing events (Lakusta et al., 2016), suggest that infants consider the endpoint an essential component of events. Likewise, 5-year-old children's preference for change-of-state over motion scenes when asked to match a novel verb to one of two scenes (Kelly & Rice, 1994) points to the prominence of change-of-states in children's perception of events. In other words, the strategy of *Endstate Orientation* draws on infants' supposedly universal conceptual preference for endpoints, which supports their subsequent linguistic encoding (see Wagner & Lakusta, 2009, for a similar view). In a similar vein, Maguire et al. (2010) propose that children initially show common, possibly universal verb construal.

As *Endstate Orientation* draws on the interaction of cognitive preferences and linguistic encoding, aspects of the learning environment such as frequency of a certain structure in the input to the child should play a minor role in explaining the order of verb acquisition. Our account furthermore predicts that the strategy of *Endstate Orientation* is generalizable across different languages and, crucially, that it extends to languages that encode goal/endstate differently. That is, across typologically different languages, children are expected to first focus on endstates in both production and comprehension. Which verbal elements the child realizes first to encode the endstate is determined by how a given language encodes endstate-oriented transitions.

Note that the strategy of *Endstate Orientation* is in stark contrast to the *Manner Bias* proposed by Gentner (1978). According to the Manner Bias, at the outset of verb acquisition language learners assume that verbs denote manners of actions; notions like 'endstate' and 'culmination point' are not part of their lexical-semantic representation of verbs (see also Gropen et al., 1991; for a critical evaluation of the manner bias, see van Hout).

3.3 Predictions for the acquisition of German

Within the account of *Event Structural Bootstrapping* specific predictions for verb acquisition can be derived. Most importantly, children are expected to master differences between event types early. Regarding production, children adhering to *Endstate Orientation* are predicted to acquire event expressions referring to endstate-oriented transitions early and to first realize the prominent subevent of telic particle verbs, e.g., in German resultative particles like *auf* 'open' or *zu* 'closed', before producing full particle verbs like *aufmachen* 'open' or morphologically simple telic verbs like *öffnen* 'open', and also before non-telic verbs designating processes, states, and process-oriented transitions.

Regarding comprehension, children are expected to be sensitive to the difference between telic and atelic verbs early on. Inherently telic verbs are predicted to be mastered early, as well as compositionally telic verbs with strong telicity markers such as *aufessen* 'eat up', because the contrast to the atelic *essen* 'eat' provides clear evidence for the learner regarding its event-type. The role of quantized NPs as weak telicity markers may be acquired later, as they provide unreliable form-function cues for the learner by implicating but not entailing a telic reading.

4. Telicity in production

German-speaking children start using verbs already in their second year of life (Kauschke, 2000). Simplex verbs are generally the first verbs to appear, but both simplex and particle verbs are used frequently already before age 2. In a case study of Simone, *aufmachen* 'open', *kaputtmachen* 'break', and *abmachen* 'take off' were among the most frequent particle verbs (Behrens, 1998). Isolated verb particles play an important role in early verb acquisition as they often assume the function of a full verb and occur already in the single-word period. This is illustrated in (10) and (11) for German (see Penner et al., 2003).

(10) Child (1;03 years):
Auf! (trying to open an umbrella)
open
'open it'

(11) Child (1;05 years):
Aus! (trying to take her sweater off)
out
'take it off'

In previous work, we investigated children's early verb lexicon between the ages of 12 to 24 months in a longitudinal design, based on data from the parent report RELATIONAL WORD INVENTORY (RWI, Schulz, 2002), which assesses production of verb particles, and from spontaneous speech corpora. The parental report data from 47 German-speaking children showed that isolated verb particles are first used between the ages of 14 and 18 months (Schulz, 2005). Most children log into the verb lexicon with isolated verb particles only, few start out with simple verbs and isolated verb particles simultaneously. Particle verbs occur sometime later, around 18 months of age. Regarding the event type of the first verb particles, 90% of the children logged into the verb lexicon with particles that in the adult system are telic as *aus* 'out', *zu* 'closed', *auf* 'open', *an* 'on', and *ab* 'off'.

From a subset of the children above (n = 43) between the ages of 14 and 18 months spontaneous speech data were analyzed regarding frequency of verb particles, particle verbs, and simplex verbs (Schulz, 2005). Eighty-six percent of the children's spontaneous verb expressions were isolated telic verb particles such as *auf* 'open' or *ab* 'off', compared to 2% ambiguous verb particles, 9% simplex verbs, and 3% particle verbs. In line with previous findings from Penner et al. (2003), the verb particles were used in the function of verbs. In summary, the data from both parent report and spontaneous speech confirms that telic verb particles are produced early and frequently in German.

In line with the *Event Structural Bootstrapping* account, Kieburg and Schulz (2010) found that parents' verb input did not determine children's order of acquisition of event expressions. Using a longitudinal design, the authors examined whether children's early preference for telic verb particles could be explained by word frequency in the ambient language. The analysed data comprised about 5000 utterances from three mothers recorded when their children were between 14 and 20 months old. The order of verb acquisition and the composition of the verb lexicon in 1- to 2-year-olds was not significantly correlated with word frequency, i.e. total token, relative token and type frequency, in parental input. Moreover, the verb particles used by the mothers most frequently (i.e. *her* 'here', *hin* 'there', *rein* 'in', *weg* 'off') differed from those initially produced by their children as documented in the parental report RWI. These were *auf* 'open', *ab* 'off', *aus* 'off', *an* 'on', *zu* 'closed', the same as the resultative verb particles reported in the study of the 47 children mentioned above. These data support the assumption that *Endstate Orientation* guides young children in building their productive verb lexicon.

5. Telicity in comprehension

Under the *Event Structural Bootstrapping* account delineated in Section 3, typically developing children are expected to master the difference between telic and atelic verbs from early on. Children who exhibit Specific Language Impairment (SLI), in contrast, may show deficits in this area.

5.1 A note on children with SLI and semantic deficits

With a prevalence of 6 to 10%, SLI constitutes one of the most frequent developmental disorders and has been the topic of much linguistic research (see the overview in Leonard, 2014; for an overview of research on SLI in German, see Hamann, 2015). Children with SLI exhibit language difficulties without co-occurring cognitive or neurological deficits, or hearing impairments severe enough to explain the language impairment (Leonard, 2014). Children with SLI are delayed in their onset of speech and in their subsequent language development. Problems with morphology and syntax have been proposed to constitute the core characteristic of SLI and have been studied most extensively. Morpho-syntactic deficits have been reported for the majority of children with SLI and have often been found to persist up to school age. However, SLI is a heterogeneous disorder: The specific profiles of language deficits vary regarding the location of the deficits as well as regarding the severity of impairment (Friedmann & Novogrodsky, 2008; Schulz, 2010; van der Lely, 2005). Beyond morpho-syntax, impairments have also been reported for phonology, lexicon, and pragmatics (e.g., Bishop, 1997; Conti-Ramsden & Botting, 2006; Leonard, 2014). Finally, selective deficits have been found for specific linguistic subdomains including morpho-syntax, phonology, and lexicon (e.g., Friedmann & Novogrodsky, 2008; van der Lely, 2005).

Semantic impairments have received less attention and have mostly been considered in concert with pragmatic or lexical deficits. The few existing studies on the semantic abilities of children with SLI suggest that deficits may occur at the word, sentence, and discourse level (Botting & Adams, 2005; Roeper, 2004). Despite these findings, the difficulties children may have with semantics have not traditionally been seen as central to SLI, and only recently has it been argued that children with SLI may also exhibit isolated or co-occurring semantic deficits (Penner et al., 2003; Schulz, 2010; Schulz & Roeper, 2011; see also Hamann, 2015).

The few studies investigating telicity acquisition in SLI have mainly tested production (Kelly & Rice, 1994; Ingham, Fletcher, Schelletter & Sinka, 1998; Watkins & Rice, 1991). Focusing on resultative VPs in English, Watkins and Rice (1991) asked English-speaking children to describe video scenes and found that children with SLI used fewer resultative particles like *off* in *kick off the shoe*, compared to their TD peers. The authors conclude that children with SLI experience problems with the syntactic and semantic properties of particles that cannot be accounted for by problems referring to functional categories. Kelly and Rice (1994) found that five-year-old SLI children, unlike their TD peers, did not show any preference in applying novel verbs to either a motion or a change-of-state scene. A related study by Ingham et al. (1998) found that 6-year-old English-speaking children with SLI have difficulty using complex resultative VPs like *shake the ball out (of the tree)* and preferred simple VPs such as *shake the ball*, when asked to describe video scenes. These studies indicate that resultative verb types may be difficult for children with SLI to acquire. In a similar vein, a comprehension study on Spanish reported that children with SLI did not use the telicity of predicates to guide their understanding of grammatical aspect and tense (Grinstead, McCurley, Pratt, Obregon, & Flores, 2013). I hypothesize that SLI children's difficulty with telic verbs results from an unstable *Endstate Orientation*: They do not reliably represent the endstate as being entailed by semantically telic verbs. That is, SLI children are expected to fail to linguistically encode their preference for endstates, which I suppose is unimpaired at the conceptual level.

5.2 Inherent telicity

Three sets of studies investigated how and at what age German-speaking TD children and children with SLI master inherent semantic telicity.

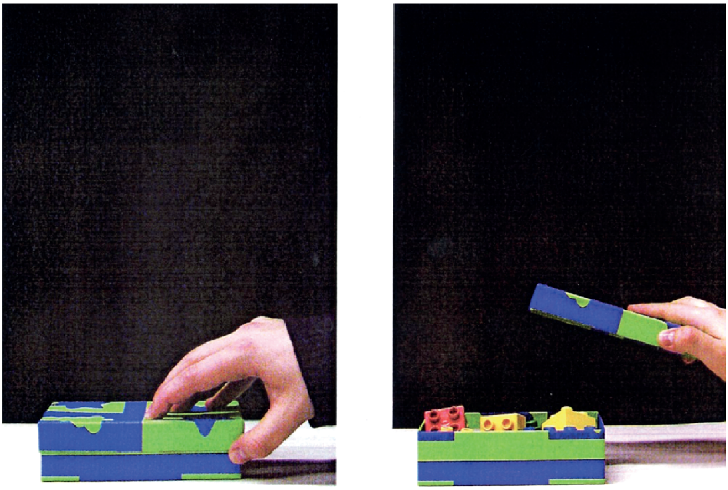
5.2.1 *Inherent telicity: The case of particle verbs*

Previous research indicates that, starting at 3 years of age, TD children exhibit an adult-like interpretation of inherently telic verbs like *aufmachen* ‘open’, *zumachen* ‘close’, *abmachen* ‘remove’, and *ausmachen* ‘extinguish’ (Wittek, 2002). That is, in a truth-value-judgment-task, they consistently reject telic verbs for incomplete events. Extending this research, Schulz and Penner (Penner et al., 2003; Schulz et al., 2001, 2002) investigated TD and SLI children’s comprehension of inherently telic particle verbs. Their study was based on the assumption that *Endstate Orientation* guides TD children’s comprehension from early on (see Section 3.2), while children with SLI, lacking *Endstate Orientation*, should not reliably recognize that the endstate is entailed by telic verbs. The authors focused on clearly endstate-oriented transitions; the telic particle verb *aufmachen* ‘open’ was chosen for the theoretical reasons provided above (see Section 2.2) and because it appears early in children’s speech (see Section 4).

Schulz et al. (2001) tested 16 3- and 4-year-old children with SLI (mean 3;10) and compared their comprehension abilities with that of 16 younger TD children (mean 2;10); 16 adults were included as controls. The children with SLI met the standard inclusion and exclusion criteria: (a) they were diagnosed by speech therapists as suffering from receptive and expressive language deficits, (b) cognitive functioning was reported to be within normal range for age, and (c) there was no report of hearing impairments (see Leonard, 2014). All SLI children were enrolled in programs for children with language disorders, and none had received any therapy focusing specifically on verb meanings. The TD children exhibited age-appropriate speech, language, social, and cognitive functioning according to preschool teacher and parent reports.

Using a truth-value judgment task (TVJ, Crain & McKee, 1986) eight *yes/no*-questions assessed whether children know that the meaning of *aufmachen* ‘open’ entails the endstate ‘be open’. Two-picture sequences were shown to the child depicting different instances of opening a container, with and without event completion. Examples for both test conditions (endstate and no-endstate) are given in Figure 1a and 1b.

a. Endstate condition
Test question: *Hat sie'se aufgemacht*? 'Did she open it?'
Target answer: Yes.



b. No-endstate condition
Test question: *Hat sie'se aufgemacht*? 'Did she open it?'
Target answer: No.

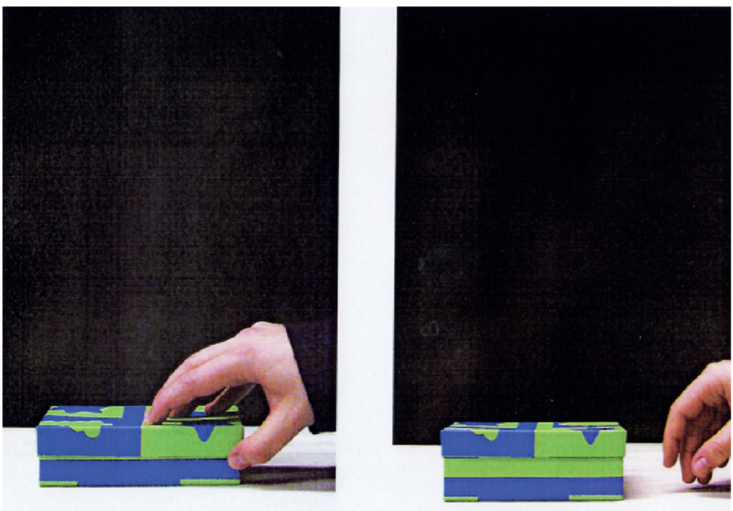


Figure 1. Example item for inherent telicity in the two test conditions.

In the relevant no-endstate condition, the 3-year-olds with SLI performed at chance level (46% correct rejections), while the 2-year-old TD children interpreted the telic verb target-like in most cases (78% correct rejections). Schulz et al. (2002) tested an additional group of 16 3-year-old TD children (mean 3;10 years), who were chronologically age-matched to the group of children with SLI, so that for each child with SLI there was a child in the age-matched TD group within 1 month of age. As expected, the TD same-age peers performed at ceiling in the no-endstate condition with telic verbs (89% correct rejections). Are SLI children's frequent *yes* responses (54%) due to a general yes-bias? This seems implausible, for only children who were able to respond to *yes/no*-questions appropriately were included in the analysis. Moreover, children with SLI, just like their TD peers, performed at ceiling in the control trials requiring a *no*-response.[4] Could it be that children with SLI merely had difficulty inferring from a photograph whether a container is open or still closed? This is unlikely, because all children were first given ample opportunity to manipulate all containers depicted in the picture sequences and experience how they open and close.

To find out whether SLI children's difficulty with inherent telicity is persistent, Penner et al. (2003) tested a group of 16 SLI children between the ages of 5 and 8 years (mean 6;09 years). As illustrated in Figure 2, which summarizes the results for all five groups tested, even at nearly 7 years of age, in only 64% of the cases did the SLI children correctly reject the telic verb *aufmachen* in the crucial no-endstate condition.

The individual interpretation patterns demonstrate SLI children's difficulties even more clearly. While 50% of the TD children at age 2 and 69% of the TD children at age 3 performed above chance (4/4 correct) with inherent telicity, only 25% of the 3-year-olds with SLI and 44% of the 6-year-olds with SLI did so. Notably, children with SLI do not simply interpret the telic verb as atelic: Only 7 out of the 32 of children with SLI overall consistently accepted telic verbs in the no-endstate condition. Extending Penner et al.'s (2003) proposal, I assume that this pattern results from an incorrect event-semantic representation of telic verbs, in which the head-of-event is optional, i.e. the endstate may be present or absent (see (1a) vs. (1b)). This lack of an *Endstate Orientation* would point to a semantic impairment that is not a result of difficulties in the syntactic, phonological, or pragmatic module, an impairment specific to the verbs' event structure (Schulz, 2010). However, as children's knowledge of atelic verbs and general lexical abilities were not assessed in this study, low performance in telicity could in principle also reflect a broad

4. In the control trials the events displayed in the picture-sequences did not match the verb used in the question.

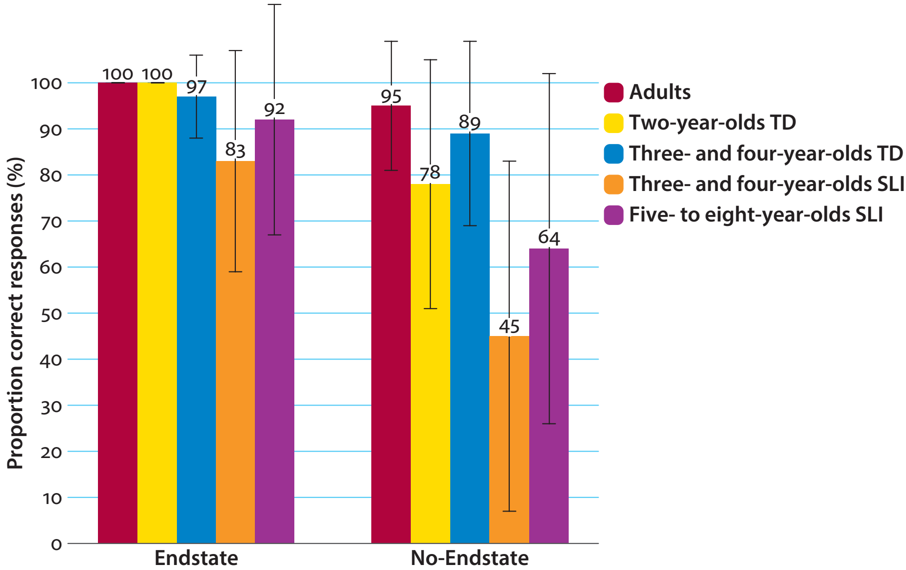


Figure 2. Inherent telicity: Mean proportion of correct responses in the Endstate and the No-Endstate condition by participant group. Error bars indicate standard deviation.

lexical deficit. Moreover, as picture sequences were used, the studies described so far leave open the possibility that SLI children's difficulty with telicity results from domain-general difficulties, for example, with processing complex events presented in pictures. These aspects were addressed in follow-up work described below.

5.2.2 *Difficulty with inherently telic verbs: Result of a general lexical deficit?*

Two subsequent studies explored whether SLI children's difficulties with telicity result from a general lexical deficit (Schulz & Kiese-Himmel, 2006; Schulz & Wittek, 2003; see also Schulz, 2010). Schulz and Kiese-Himmel (2006) tested children with SLI on telic verbs and on their active vocabulary. Telicity interpretation was assessed using the task by Schulz et al. (2001). Children's active vocabulary was tested with a standardized test that required naming pictures depicting nouns and verbs (AWST, Kiese-Himmel & Kozielski, 1996). Participants in the study were 20 5-year-old children with SLI, diagnosed in a speech therapy clinic. They all met the exclusionary criteria for SLI and exhibited expressive language deficits in morphosyntax and phonology; eight children also showed lexical deficits. Performance on the standardized vocabulary test was in the lower normal range (mean T-value: 44.0); individual T-values confirmed the expert classification regarding lexical deficits. Confirming the previous findings, the SLI children showed chance performance on telic verbs in the no-endstate condition (63% correct). Importantly, there was no correlation between performance on the telicity task and on the vocabulary test. In addition, children who mastered telicity did not have a larger overall vocabulary or verb vocabulary than children who failed telicity. These results suggest that difficulties with telicity are independent of a general lexical deficit.

Schulz and Wittek (2003) found that SLI children's difficulty with telicity is restricted to telic verbs and that atelic verb comprehension is unimpaired. The atelic verbs tested were *malen* 'draw', *fegen* 'sweep', *wischen* 'wipe', *bauen* 'build', *schneiden* 'cut', *bürsten* 'brush', *pusten* 'blow, and *puzzeln* 'do a puzzle'. In the no-endstate condition half completed events were shown (e.g., for 'Did she draw?' a half drawn car). Sixteen 5-year-old children with SLI and 16 chronologically age-matched TD children participated in this study. The children with SLI met the typical exclusionary criteria for SLI and were enrolled in a preschool program for children with language disorders. Children with SLI, just like their same-age TD peers, correctly accepted atelic verbs in the no-endstate condition (94% correct), indicating an adult-like interpretation of atelic verbs.

5.2.3 *Difficulty with inherently telic verbs: Just a matter of experimental method?*

One may object that use of picture sequences is not ideal for depicting changes of states and that this may have affected SLI children's performance. That is, SLI children's difficulty with telicity may result from processing complex events presented in pictures, which could point to domain-general difficulties rather than to selective deficits with telicity. The study by Schulz and Wittek (2003) (see Section 5.2.2.) was able to evaluate this possibility, because their TVJ task involved acting out the scenes instead of using picture sequences. All scenes were acted out by a puppet in front of the child by the first experimenter. After the performance, the puppet was seated away from the scene to clearly indicate that the action had stopped. A second experimenter then asked a *yes/no*-question. This way the (in)completeness of the event was made explicit and task demands for the child were lowered. The telic verbs tested were *aufmachen* 'open', *zumachen* 'close', *anmachen* 'turn on', and *abmachen* 'take off'. In the eight relevant test trials, event culmination did not occur (e.g., a container was manipulated without opening it). As expected, the TD children performed well on telic verbs in the no-endstate condition. The children with SLI showed chance performance on telic verbs in the no-endstate condition (53% correct). An analysis by individual confirmed that telic verbs were mastered by 80% of the TD children, but only by 50% of the children with SLI, indicating that the endstate as head-of-event in telic verbs is optional for many children with

SLI. Recall that this is exactly the representation for processes like *draw* that need not result in a culmination point such as a completed picture of a flower.

The findings from this TVJ task using props are in line with the results reported above for picture-sequences and provide clear evidence that lowering the task demands did not improve SLI children's ability to interpret telic verbs. Schulz and Wittek's (2003) study then suggests that SLI children's problems with telicity cannot be accounted for by domain-general problems with processing complex events, but are related to domain-specific semantic difficulties.

5.2.4 *Inherent telicity: Interim summary*

If children exhibit an *Endstate Orientation*, they should recognize that verbs like *aufmachen* 'open' designate events with a natural culmination point that is entailed by the verb meaning. *Endstate Orientation* was shown to guide TD children's comprehension from early on (i.e. from age 2 onwards), the youngest age where testing with TVJ was possible. Children with SLI were found to not reliably recognize that the endstate is entailed by inherently telic verbs up to 8 years of age. A comparison with children's interpretation of atelic verbs revealed that SLI children's difficulties with event types do not extend to atelic verbs, which were correctly assigned an event structural representation without an entailed endstate. Moreover, a comparison of SLI children's knowledge of inherent telicity with vocabulary knowledge indicates no general difficulties in the lexical domain, pointing to a genuine impairment in the lexical semantics module. Depicting the events via acted out scenes rather than via picture sequences did not improve SLI children's performance. Hence the difficulty with inherently telic verbs is not caused by general difficulties with processing complex events in picture sequences, pointing to a domain-specific problem. SLI children's difficulty is argued to result from an instable *Endstate Orientation*, which leads to a representation of endstate-oriented transitions with an optional rather than an obligatory endstate as head-of-event. This mis-setting in the event-semantic representation, which applies to telic but not to atelic verbs, points to a deficit in the sense of de Villiers (2003): A deficit is attested if a piece of grammar is missing or if something is mis-set in the structure. Whether morphologically simple inherently telic verbs like *finden* 'find' follow the same pattern reported here for particle verbs is open.

5.3 Compositional telicity

In Section 2.3, it was argued that compositional telicity results from adding resultative particles or quantized noun phrases to an atelic process verb. The function of strong telicity markers like *auf* in *aufessen* 'eat up' was predicted to be acquired as early as that of inherent telic particle verbs, while the function of weak telicity markers such as certain quantized noun phrases in triggering telicity may be acquired later.

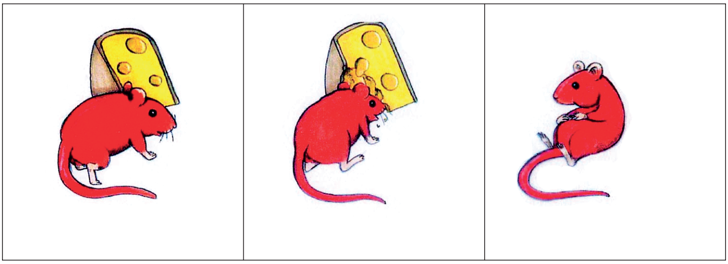
Previous research on the acquisition of telicity investigated comprehension of particle verbs and quantized noun phrases across different languages (see van Hout, 2018). Van Hout (1998, 2000) found that the particle verbs *eat up* and *drink up*, and their Dutch counterparts, were correctly restricted to telic interpretations by most 4- and 5-year-olds. As for the interpretation of quantized noun phrases, children up to age 5 allowed predicates with definite NPs such as *eat his cheese* or *drink his tea*, and their Dutch equivalents, to refer to events with and without event culmination. Since 4- and 5-year-olds in these studies correctly interpreted predicates with bare noun phrases such as *eat cheese* or *drink tea* as atelic, it is unlikely that they were simply unaware of the presence of the determiner. According to van Hout (1998, 2008), children are more lenient than adults in accepting non-culmination with definite NPs, because there is no overt telicity marker on the verb. It may be that the child first learns overt compositional ('predicational' in van Hout's terms) telicity markers such as resultative particles and only later becomes sensitive to the role of the quantized noun phrase in encoding weak telicity.

Two sets of studies tested strong and weak telicity markers in German TD and SLI children within the same experiment (Schulz & Ose, 2008; Schulz & Penner, 2002; Schulz & Wenzel, 2005). Section 5.3.1. discusses the findings on strong telicity markers, Section 5.3.2. those on weak telicity markers.

5.3.1 *Strong telicity markers*

Adapting the TVJ design by van Hout (1996), Schulz and Penner (2002) tested children's comprehension of verbs of consumption, contrasting atelic intransitive *essen* 'eat' and *trinken* 'drink' with their telic particle counterparts *aufessen* 'eat up' und *austrinken* 'drink up'. Unlike in Dutch and English, where the particle verb items were presented with a definite NP (e.g., *The mouse ate up his cheese*), in German, intransitive forms were used (e.g., *Die Maus hat aufgegessen* 'the mouse ate up'), which is grammatical in German, but not in English or Dutch. This way it was possible to assess the effect of the particle independent of the presence of a quantized noun phrase. Participants saw eight picture-sequences depicting different events of eating and drinking, either with or without event completion. Examples of both conditions are given in Figures 3a and 3b.

a. Endstate condition
Test question 1: *Hat die Maus gegessen*? Did the mouse eat?
Target answer: Yes.
Test question 2: *Hat die Maus aufgegessen?* 'Did the mouse eat up?'
Target answer: Yes.



b. No-endstate condition
Test question 1: *Hat die Maus gegessen?* 'Did the mouse eat?'
Target answer: Yes.
Test question 2: *Hat die Maus aufgegessen?* 'Did the mouse eat up?'
Target answer: No.

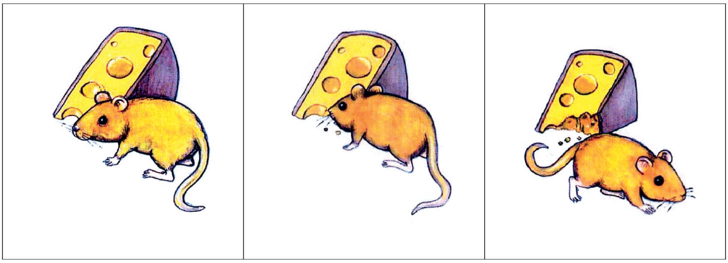


Figure 3. Example item for compositional telicity in the two test conditions.

Control questions asking about details of the story were added to provide an equal number of possible *yes* and *no* responses. Twenty-four German-speaking TD children aged 4 to 6 (mean 5;04) and 24 adults participated in the study. In the no-endstate condition, children, like adults, interpreted both atelic verbs (97% correct *yes* responses) and telic particle verbs (96% correct *no* responses) as predicted. A follow-up study (Schulz & Ose, 2008) with 39 3-year-olds and a new group of 18 children 4- and 5-year-olds, who were all classified as TD via a standardized language test, substantiated the findings of Schulz and Penner (2002). Moreover, they provided evidence that, already at 3 years of age, TD children know that intransitive verbs like *eat* and *drink* can designate events without completion (87% correct *yes* responses) and that the telic particles *auf* in *aufessen* 'eat up' and *aus* in *austrinken* 'drink up' create an event-type shift from atelic to telic (87% correct *no* responses). In summary, TD children's knowledge of strong telicity markers as event-type shifters is – like their knowledge of the event type of inherently telic particle verbs – present from the earliest ages tested to date.

Using the same design, Schulz and Wenzel (2005) tested 13 5-year-old children with SLI who had been diagnosed with receptive and expressive disorders by a speech therapist and had received below-average test scores in at least two subtests of a standardized language test. In the no-endstate condition, children with SLI, like their same-age TD peers, interpreted both the atelic verbs (100% correct *yes* responses) and the telic particle verbs (96% correct *no* responses) target-like. This is *prima facie* surprising given the persistent difficulty of SLI children with inherently telic verbs, reported in Section 5.2. It may be that the event structure of compositionally telic particle verbs is easier to acquire than that of inherently telic particle verbs, because the alternating forms *essen* 'eat up' and *aufessen* 'eat up' map one-to-one to the different event-types atelic and telic. If this analysis is on the right track, this data would point to a very specific deficit within the domain of verbal semantics. In other words, it may be that the lack of *Endstate Orientation* is restricted to lexical but not to compositional semantics. Alternatively, it may be that the presentation of both forms within the same experiment primed learners with SLI to pay attention to the presence or absence of the particle. Data from more children with SLI, across different designs, are needed to substantiate the findings and to decide between these two possibilities. Notably, this response pattern clearly speaks against a Manner Bias in SLI.

5.3.2 *Weak telicity markers*

Regarding their role in marking telicity, weak telicity markers such as quantized noun phrases are expected to be mastered later as they do not provide an obligatory cue for an event-type shift from atelic to telic. Adopting van Hout's design (1996), Schulz and Penner (2002) tested this prediction for German. Comprehension of quantized noun phrases was assessed via *yes/no*-questions like *Hat die Maus den Käse gegessen*? 'Did the mouse eat the cheese?' either with or without event completion (see Figures 3a and 3b). In the no-endstate condition, adults and the five-year-olds, who also took part in the study mentioned above, accepted predicates with quantized NPs in half of the cases (children: 56% *yes* responses, adults: 52% *yes* responses). Given that the adults' responses reflect target-like interpretation, this finding clearly confirms the assumption that with quantized NPs telicity arises though implicature. An analysis of adults' typical verbal comments (e.g., *Yes, but she left a little*) suggests that they take the quantized NP to mark the specific object mentioned previously in the discourse (i.e. the cheese mentioned in the story compared to some other cheese or food not mentioned). Judging from children's typical verbal comments (e.g., *Yes, a little*) they seem to interpret quantized NPs as specific but not as obligatorily triggering an event-type shift from atelic to telic, just like the adults.

A follow-up study (Schulz & Ose, 2008) involved a new group of 4- and 5-year-olds and a group of 3-year-olds, as well as a new group of adults. In addition to the two conditions above, transitives with quantized NPs (e.g., *den Käse essen* '*eat the cheese*') and with bare nouns (e.g., *Käse essen* 'eat cheese') were tested. Figure 4 summarizes the results across all sentence types for the crucial no-endstate condition. As expected, bare nouns were accepted with non-culminated events by all groups. Even more strongly than in the study by Schulz and Penner (2002), quantized NPs were accepted with non-culmination often (adults: 73%, 3-year-olds: 86%, 4-and 5-year-olds: 69%), with no difference between the groups. Whether this difference between the two studies is accidental or is related to the slightly different design cannot be answered conclusively. The latter but not the former study included bare nouns, which provide a contrast to quantized NPs, but may have also raised participants' sensitivity to the presence of a noun in general. Note that, in contrast to the English and Dutch findings by van Hout (1998, 2000), German-speaking children were not found to be more lenient than adults in accepting non-culmination with weak telicity markers. This may be due to the fact that the quantization is less clear in specific definite NPs than in measure phases (*a cup of tea*) or possessives (*his tea*).

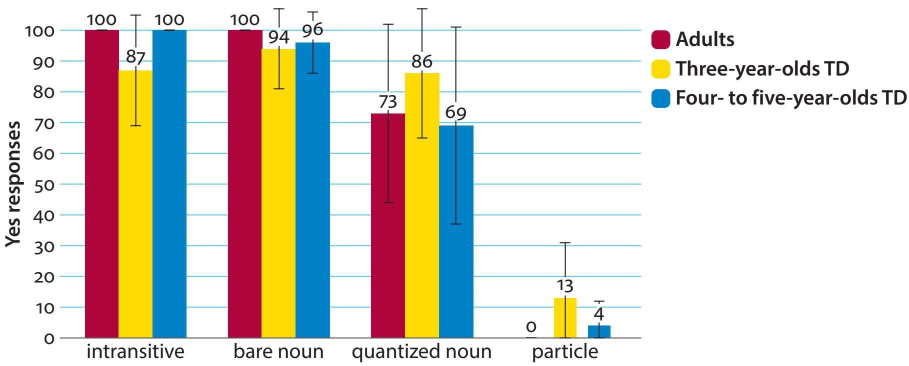


Figure 4. Mean proportion of *Yes* responses in the No-Endstate condition across sentence types and participant groups. Error bars indicate standard deviation

We can conclude that German-speaking TD children are sensitive to weak compositional markers of telicity, and interpret transitive sentences with a quantized NPs as telic or atelic, just like the adult participants in our study do. What about children with SLI? Schulz and Wenzel (2005) also tested the 5-year-old SLI children with bare nouns and quantized NPs. Again their performance was not different from the same-age TD peers, correctly accepting bare nouns (92% *yes* responses) in the no-endstate condition and accepting quantized NPs in this condition (81% *yes* responses) at the same rate as their same-age TD peers.[5]

5. Recall that the children with SLI did not exhibit a yes bias (see Section 5.3.1).

In summary, starting at age 3 German-speaking TD children differentiate between strong and weak telicity markers in an adult-like manner. Telic particles like *auf* 'open' and *aus* 'off' are interpreted as obligatory event-type shifters from atelic to telic, and quantized NPs are interpreted as ambiguous between telic and atelic with verbs of consumption. That is, by age 3 children know that completion is derived by implicature in the latter and by entailment in the former case. It remains to be seen whether these findings extend beyond verbs of consumption to other verb classes including verbs of motion and surface-contact.

6. Conclusion and outlook

The acquisition studies on inherent and compositional telicity summarized in this chapter were based on the assumption that event completion can be entailed (semantic telicity) or implicated (pragmatic telicity). Inherently telic verbs and predicates with strong compositional telicity markers such as verb particles entail event completion, while weak telicity markers such as quantized NPs implicate event completion.

The account of *Event Structural Bootstrapping* outlined here argues that TD children initially focus on the verb's event structure. In their early acquisition of verb semantics, across languages TD children are argued to exhibit the strategy of *Endstate Orientation*, which is rooted in children's early conceptual and visual preference for endpoints. Children following an *Endstate Orientation* are predicted to focus on telic verbs designating endstate-oriented transitions, that obligatorily mark the endstate as head-of-event (see Penner et al., 2003; Schulz et al., 2001, 2002). The strategy of *Endstate Orientation* was contrasted with the Manner bias (Gentner, 1978) claiming that children focus on manner of actions. I hypothesized that children with SLI lack an *Endstate Orientation* and hence should not recognize that the endstate is entailed by inherently telic verbs. The specific predictions for acquisition made by this account were evaluated in several production and comprehension studies in German.

Our results from parental report and spontaneous speech studies revealed that TD children acquire verbs early, exhibiting a clear preference for resultative verb particles like *auf* 'open' and *aus* 'off' as their first event expressions. Accordingly, these data confirm that German TD children adhere to *Endstate Orientation* in building their productive verb lexicon.

Our studies on comprehension showed that at the youngest age tested so far, 2 years of age for inherently telic particle verbs and 3 years of age for compositional telicity, TD children distinguish between semantic and pragmatic telicity. More specifically, TD children reject inherently telic predicates and predicates with strong telicity markers for incomplete events (i.e. semantic telicity) and, like adults, often accept predicates with weak telicity markers such as quantized NPs for incomplete events (i.e. pragmatic telicity). Notably, children who have been diagnosed with SLI, based on standard morpho-syntactic tests, were found to exhibit persistent difficulty with inherently telic verbs. These problems in SLI children were argued to result from an event-semantic representation of complex events in which the endstate as head-of-event is optional, leading to chance performance. The event-structural deficit found for German-speaking children with SLI points to a persistent deficit in the domain of (lexical) semantics that so far has not been at the center of SLI research. Results from a study on compositional telicity indicate that German-speaking children with SLI, in contrast to their persistent difficulty with inherently telic verbs such as *aufmachen* 'open', did not have difficulty with verbs like *aufessen* 'eat up' und *austrinken* 'drink up'. Note that this response pattern also speaks against a Manner Bias in SLI. Taken together, these data from SLI and TD children serve to illustrate how modular acquisition and how selective impairment can be and thus argue for modularity in the domain of verbal semantics as well.

Recently this line of research has been applied to diagnostics. A group of 3- to 8-year-old children, diagnosed with SLI using a standardized test with a focus on morpho-syntax, were reported to have lower test scores than TD children in the semantic telicity subtest of the standardized language test Lise-DaZ (Schulz & Tracy, 2011). This finding suggests that verb semantic deficits may even be robust enough to surface in a standardized language test that has to adhere to constraints different from those for an experiment.

Concluding, the account of *Event Structural Bootstrapping* and, specifically, the strategy of *Endstate Orientation* delineated in this chapter, invite several general predictions: (i) cross-linguistically, endstate-oriented transitions are acquired early, (ii) verb semantics can be selectively impaired in children with SLI, (iii) verb semantic deficits exist in children with SLI across languages, but may surface differently, depending on how the language marks telicity.

Using categories of decompositional representation (Dowty, 1979) or the level of lexical conceptual structure (e.g., Pustejovsky, 1995), the representation of a telic predicate can be roughly stated as [CAUSE… [BECOME [STATE]]]. In Germanic languages, telic verb particles transparently encode the endstate and are hence expected to play a prominent role in early acquisition. In non-Germanic languages the endstate may not be lexicalized (e.g., absence of resultative adjective constructions in Romance). The strategy of *Endstate Orientation* then results in a further prediction: (iv) in non-Germanic languages young learners realize the endstate as a single lexeme, i.e. as the complex head [CAUSE [BECOME [STATE]]]. More research across languages and acquisition types is required to test these four predictions. This may contribute to both the field of acquisition and the field of theoretical semantics.

Acknowledgements

This research was supported by grants from the German Research Council to Zvi Penner (SFB No. 471; FOR 381) and from the LOEWE program for excellency from the State of Hesse to Petra Schulz (Project MILA, IDeA Center). I thank Naama Friedmann, Angeliek van Hout, Manfred Krifka, Ana Pérez-Leroux, Tom Roeper, Esther Rinke and Emanuela Sanfelici for helpful discussion of data and theory. The comments of the three anonymous reviewers and the editors greatly helped to improve the manuscript. Finally I gratefully acknowledge the children and adults who participated in the studies reported here.

References

Arunachalam, S., & Kothari, A. (2010). Telicity and event culmination in Hindi perfectives. In P. M. Bertinetto, A. Korhonen, A. Lenci, A. Melinger, S. Schulte im Walde, & A. Villavicencio (Eds.), *Proceedings of the 2nd interdisciplinary Workshop on the Identification and Representation of Verb Features* (pp. 16–19). Pisa: University of Pisa.

Behrens, H. (1998). How difficult are complex verbs? Evidence from German, Dutch and English. *Linguistics*, 36, 679–712. https://doi.org/10.1515/ling.1998.36.4.679

Bishop, D. V. M. (1997). *Uncommon understanding*. Hove: Psychology Press.

Botting, N., & Adams, C. (2005). Semantic and inferencing abilities in children with communication disorders. *International Journal of Language and Communication Disorders*, 40, 49–66. https://doi.org/10.1080/13682820410001723390

Conti-Ramsden, G., & Botting, N. (2006). Specific Language Impairment. In K. Brown (Ed.), *Encyclopedia of Language and Linguistics* (pp. 629–632). Oxford: Elsevier. https://doi.org/10.1016/B0-08-044854-2/00844-0

Crain, S., & McKee, C. (1986). Acquisition of structural restrictions on anaphora. In S. Berman, J. W. Choe, & J. McDonough (Eds.), *Proceedings of the Northeastern Linguistic Society*. Amherst, MA: GLSA.

Demirdache, H., & Martin, F. (2015). Agent control over non culminating events. In E. Barrajón López, J. L. Cifuentes Honrubia, & S. Rodríguez Rosique (Eds.), *Verb classes and aspect* (pp. 185–217). Amsterdam: John Benjamins. https://doi.org/10.1075/ivitra.9.09dem

de Villiers, J. G. (2003). Defining SLI: A linguistic perspective. In Y. Levy & J. Schaeffer (Eds.), *Language competence across populations: Toward a definition of specific language impairment* (pp. 425–447). Hillsdale, NJ: Lawrence Erlbaum Associates.

Dowty, D. R. (1979). *Word meaning and Montague grammar*. Dordrecht: Reidel. https://doi.org/10.1007/978-94-009-9473-7

Filip, H. (2008). Events and maximalization: The case of telicity and perfectivity. In S. Rothstein (Ed.), *Theoretical and crosslinguistic approaches to the semantics of aspect*, (pp. 217–256). Amsterdam: John Benjamins. https://doi.org/10.1075/la.110.10fil

Filip, H. (2014). Weak and strong telicity via maximization Delimit Event. Presentation at Workshop on Identification and Delimitation of Events, Université Paris 7, March, 2014.

Friedmann, N., & Novogrodsky, R. (2008). Subtypes of SLI: SySLI, PhoSLI, LeSLI, and PraSLI. In A. Gavarró & M. J. Freitas (Eds.), *Language acquisition and development* (pp. 205–217). Newcastle upon Tyne: Cambridge Scholars.

Gentner, D. (1978). On relational meaning: The acquisition of verb meaning. *Child Development*, 49, 988–998. https://doi.org/10.2307/1128738

Gentner, D. (1982). Why nouns are learned before verbs: Linguistic relativity versus natural partitioning. In S. Kuczaj (Ed.), *Language development, Vol. 2: Language, thought, and culture*. (pp. 301–334). Hillsdale, NJ: Lawrence Erlbaum Associates.

Grinstead, J., McCurley, D., Pratt, T., Obregon, P, & Flores, B. (2013). The semantics of the tense deficit in child Spanish SLI. In M. Becker, J. Grinstead, & J. Rothman (Eds.), *Generative linguistics and acquisition: Studies in honor of Nina M. Hyams* (pp. 107–128). Amsterdam: John Benjamins. https://doi.org/10.1075/lald.54.05gri

Gropen, J., Pinker, S., Hollander, M. & Goldberg, R. (1991). Syntax and semantics in the acquisition of locative verbs. *Journal of Child Language*, 18, 115–151. https://doi.org/10.1017/S0305000900013325

Hamann, C. (2015). Language impairment in German children. In A. Stavrakaki (Ed.), *Specific Language Impairment. Current trends in research* (pp. 215–251). Amsterdam: John Benjamins. https://doi.org/10.1075/lald.58.10ham

Hay, J., Kennedy, C. & Levin, B. (1999). Scale structure underlies telicity in 'degree achievements'. In T. Matthews & D. Strolovitch (Eds.), *Proceedings of Semantics and Linguistic Theory 9* (pp. 127–144). Ithaca, NY: CLC Publications.

van Hout, A. (1996). *Event semantics of verb frame alternation. A case study of Dutch and its acquisition* (Doctoral dissertation). University of Tilburg. (Published in 1998 by Garland, NY.)

van Hout, A. (1998). On the role of direct objects and particles in learning telicity in Dutch and English. In A. Greenhill, M. Hughes, H. Littlefield, & H. Walsh (Eds.), *BUCLD 22: Proceedings of the 22nd annual Boston University Conference on Language Development* (pp. 397–408). Somerville, MA: Cascadilla Press.

van Hout, A. (2000). Event semantics in the lexicon-syntax interface: Verb frame alternations in Dutch and their acquisition. In C. L. Tenny & J. Pustejovsky (Eds.), *Events as grammatical objects. The converging perspectives of lexical semantics and syntax* (pp. 239–282). Stanford CA: CSLI.

van Hout, A. (2008). Acquiring telicity cross-linguistically: On the acquisition of telicity entailments associated with transitivity. In M. Bowerman & P. Brown (Eds.), *Crosslinguistic perspectives on argument structure: Implications for learnability* (pp. 255–278). New York, NY: Routledge.

van Hout, A. (2018). On the acquisition of event culmination. In K. Syrett & S. Arunachalam (Eds.), *Semantics in Language Acquisition* (pp. 95–121). Amsterdam: John Benjamins. https://doi.org/10.1075/tilar.24.05hou

Ingham, R., Fletcher, P., Schelletter, C., & Sinka, I. (1998). Resultative VPs and specific language impairment. *Language Acquisition*, 7, 87–111. https://doi.org/10.1207/s15327817la0702-4_2

Jackendoff, R. (1996). The proper treatment of measuring out, telicity, and perhaps even quantification in English. *Natural Language and Linguistic Theory*, 14, 305–354. https://doi.org/10.1007/BF00133686

Jeschull, L. (2007). The pragmatics of telicity and what children make of it. In A. Belikova, L. Meroni, & M. Umeda (Eds.), *Proceedings of the 2nd Conference on Generative Approaches to Language Acquisition North America* (pp. 180–187). Somerville, MA: Cascadilla Proceedings Project.

Kauschke. C. (2000). *Der Erwerb des frühkindlichen Lexikons–eine empirische Studie zur Entwicklung des Wortschatzes im Deutschen*. Tübingen: Narr.

Kelly, D., & Rice, M. (1994). Preferences for verb interpretation in children with specific language impairment. *Journal of Speech and Hearing Research*, 37, 182–192. https://doi.org/10.1044/jshr.3701.182

Kieburg, A., & Schulz, P. (2010). Input factors in early verb acquisition: Do word order variability and word frequency of verbs matter? In M. Anderssen, K. Berentzen, & M. Westergaard (Eds.), *Optionality in the input. Papers from the GLOW 30 Workshop* (pp. 95–127). Dordrecht: Springer.

Kiese-Himmel, C., & Kozielski, P. M. (1996). *Aktiver Wortschatztest für drei- bis sechsjährige Kinder. AWST 3–6*. (2nd ed.). Weinheim: Beltz.

Krifka, M. (1989). Nominal reference, temporal constitution, and quantification in event semantics. In R. Bartsch, J. van Benthem, & P. van Emde Boas (Eds.), *Semantics and Contextual Expression*, (pp. 75–115). Dordrecht: Foris.

Krifka, M. (1998). The origins of telicity. In S. Rothstein (Ed.), *Events and grammar* (pp. 197–235). Dordrecht: Kluwer. https://doi.org/10.1007/978-94-011-3969-4_9

Leonard, L. P. (2014). *Children with Specific Language Impairment*. Cambridge, MA: The MIT Press.

Levin, B., & Rappaport Hovav, M. (1995) *Unaccusativity: At the syntax-lexical semantics interface*. Cambridge, MA: The MIT Press.

Lakusta, L., & DiFabrizio, S. (2016). And, the winner is…a visual preference for end points over starting points in infants' motion event representations. *Infancy*, 22, 323–343.

Lakusta, L., Muentener, P., Petrillo, L., Mullanaphy, N., & Muniz, L. (2016). Does making something move matter? Representations of goals and sources in causal motion events. *Cognitive Science*, 41, 814–826.

Maguire, M. J., Hirsh-Pasek, K., Golinkoff, R. M., Imai, M., Haryu, E., Vanegas, S. Okada, H., Pulverman, R., & Sanchez-Davis, B. (2010). A developmental shift from similar to language specific strategies in verb acquisition: A comparison of English, Spanish, and Japanese. *Cognition*, 114, 299–319. https://doi.org/10.1016/j.cognition.2009.10.002

Penner, Z., Schulz, P., & Wymann, K. (2003). Learning the meaning of verbs: What distinguishes language impaired from normally developing children? *Linguistics*, 41, 289–319. https://doi.org/10.1515/ling.2003.010

Pustejovsky, J. (1995). *The generative lexicon*. Cambridge, MA: The MIT Press.

Roeper, T. (2004). Diagnosing language variations: Underlying principles for syntactic assessment. *Seminars in Speech and Language*, 25, 41–55. https://doi.org/10.1055/s-2004-824825

Schulz, P. (2002). Relational Word Inventory (RWI) (Unpublished manuscript). Humboldt University, Berlin.

Schulz, P. (2005, July). Are children's first event expressions telic? Evidence from child German. Paper presented at 10th International Congress for the Study of Child Language, Berlin.

Schulz, P. (2010). Some notes on semantics and SLI. In A. Castro, J. Costa, M. Lobo, & F. Pratas (Eds.), *Language acquisition and development. Proceedings of Generative Approaches to Language Acquisition 2009* (pp. 391–406). Newcastle upon Tyne: Cambridge Scholars.

Schulz, P., & Kiese-Himmel, C. (2006). Verbverstehen und expressiver Wortschatzumfang bei sprech-/sprachentwicklungsgestörten Kindern. *L.O.G.O.S. Interdisziplinär*, 4, 244–252.

Schulz, P., & Ose, J. (2008, February). Semantics and pragmatics in the acquisition of telicity. Paper presented at 29. Jahrestagung der Deutschen Gesellschaft für Sprachwissenschaft (DGfS), Bamberg.

Schulz, P., & Penner, Z. (2002). How you can eat the apple and have it too: Evidence from the acquisition of telicity in German. In J. Costa & M. J. Freitas (Eds.), *Proceedings of the GALA' 2001 Conference on Language Acquisition* (pp. 239–246). Lisboa: Associação Portuguesa de Linguística, Lisboa.

Schulz, P., Penner, Z., & Wymann, K. (2002). Comprehension of resultative verbs in normally developing and language impaired children. In F. Windsor, M. L. Kelly, & N. Hewlett (Eds.), *Investigations in clinical phonetics and linguistics* (pp. 115–129). Mahwah, NJ: Lawrence Erlbaum Associates.

Schulz, P., & Roeper, T. (2011). Acquisition of exhaustivity in wh-questions: A semantic dimension of SLI? *Lingua*, 121(3), 383–407. https://doi.org/10.1016/j.lingua.2010.10.005

Schulz, P., & Tracy, R. (2011). *Linguistische Sprachstandserhebung – Deutsch als Zweitsprache (LiSe-DaZ)*. Göttingen: Hogrefe.

Schulz, P., & Wenzel, R. (2005, July). Acquisition of compositional telicity by German-speaking children with specific language impairment. Paper presented at the Symposium 'Cross-linguistic acquisition of telicity', 10th International Congress for the Study of Child Language, Berlin.

Schulz, P., & Wittek, A. (2003). Opening doors and sweeping floors: What children with specific language impairment know about telic and atelic verbs. In B. Beachley, A. Brown, & F. Colin (Eds.), *BUCLD 27: Proceedings of the 27th annual Boston University Conference on Language Development* (Vol. 2, pp. 727–738). Somerville, MA: Cascadilla Press.

Schulz, P., Wymann, K., & Penner, Z. (2001). The early acquisition of verb meaning in German by normally developing and language impaired children. *Brain and Language*, 77, 407–418. https://doi.org/10.1006/brln.2000.2411

Talmy, L. (1991). Path to realization: A typology of event integration. In L. A. Sutton, C. Johnson, & R. Shields (Eds.), *Proceedings of the 17th Annual Meeting of the Berkeley Linguistics Society: General session and parasession on the grammar of event structure* (pp. 480–519). Available online: <http://linguistics.berkeley.edu/bls/previous_proceedings/bls17.pdf>

van der Lely, H. K. J. (2005). Domain-specific cognitive systems: Insight from grammatical specific language impairment. *Trends in Cognitive Sciences*, 9, 53–59. https://doi.org/10.1016/j.tics.2004.12.002

Vendler, Z. (1957). Verbs and times. *The Philosophical Review*, 66, 143–160. https://doi.org/10.2307/2182371

Wagner, L., & Lakusta, L. (2009). Using language to navigate the infant mind. *Perspectives on Psychological Science*, 4, 177–184. https://doi.org/10.1111/j.1745-6924.2009.01117.x

Walkková, M. (2013). The aspectual function of particles in phrasal verbs (Unpublished doctoral dissertation). University of Groningen.

Watkins, R., & Rice, M. (1991). Verb particle and preposition acquisition in language-impaired preschoolers. *Journal of Speech and Hearing Research*, 34, 1130–1141. https://doi.org/10.1044/jshr.3405.1130

Wittek, A. (2002). *Learning the meaning of change-of-state verbs: A case study of German child language*. Berlin: Mouton de Gruyter. https://doi.org/10.1515/9783110907988

SECTION III

Syntactic structure and semantic meaning

CHAPTER 7

Not all subjects are agents

Transitivity and meaning in early language comprehension

Rose M. Scott[i], Yael Gertner[ii] and Cynthia Fisher[ii]
[i]University of California Merced / [ii]University of Illinois at Urbana-Champaign

Children use syntax to guide sentence comprehension and verb learning. We explored the nature of the meanings children infer from syntactic evidence by examining the types of event-roles they can link with the subjects and objects of transitive verbs. In two experiments, 23-month-olds heard a novel verb in a transitive sentence while viewing pairs of events in which one participant acted on another without producing a clear effect (Experiment 1) or one participant moved relative to another without contacting it (Experiment 2). In both cases, children looked longer at the event in which the subject referent played a more prominent role. These findings suggest that children map a highly abstract conceptual-semantic asymmetry onto the syntactic difference between subjects and objects.

Keywords: syntactic bootstrapping, verb learning, language acquisition, thematic roles, transitivity, word order

1. Introduction

Children use syntax to guide sentence comprehension and verb learning (e.g., Arunachalam, 2013; Arunachalam & Waxman, 2010; Naigles, 1990; Yuan, Fisher, & Snedeker, 2012). This syntactic bootstrapping procedure has been proposed to solve otherwise intractable problems of ambiguity that face attempts to infer verb meanings from world events alone (e.g., Fisher, Hall, Rakowitz, & Gleitman, 1994; Gleitman, Cassidy, Nappa, Papafragou, & Trueswell, 2005). The structure-mapping account of the origins of syntactic bootstrapping (Fisher, Gertner, Scott, & Yuan, 2010; Gertner & Fisher, 2012; Yuan et al., 2012) further proposes that some syntactic guidance for verb interpretation is available from the start of multi-word sentence comprehension. On this account, syntactic bootstrapping begins with an innate bias towards one-to-one mapping between the number of nouns in a sentence and the

https://doi.org/10.1075/tilar.24.07sco

number of core participant-roles in a conceptualization of an event. Equipped with this bias, even toddlers assign appropriately different meanings to unknown verbs essentially by *counting the nouns* in simple sentences (Jin, 2015; Yuan et al., 2012). As a result, a new verb that appears in two-noun transitive sentences is interpreted as describing a two-participant relation, whereas a verb that appears in one-noun intransitive sentences is assumed to describe one participant's role[1].

But learners go beyond counting the nouns, to determine how their language links semantic roles with syntax – thus specifying who does what to whom. For example, on hearing *The zum mecked the zard*, adults infer not only that *mecking* describes a two-participant relation, but also that *the zum* plays a more agentive role in that relation than *the zard* (Kako, 2006). Children make strikingly similar inferences, using word order to interpret novel verbs (e.g., Dittmar, Abbot-Smith, Lieven, & Tomasello, 2008; Fernandes et al., 2006; Gertner & Fisher, 2012; Gertner, Fisher, & Eisengart, 2006). For instance, Gertner et al. (2006) presented 21-month-olds with a pair of caused-motion events. In one event, a girl bent a boy forward and back by pushing and pulling on his shoulders; in the other, the boy turned the girl in a swivel chair. While viewing these events, children heard "The girl is gorping the boy" or "The boy is gorping the girl." Children looked longer at the event in which the subject of their test sentence played the agent's role. Thus toddlers possess some general knowledge of how their language links semantic roles to syntax: Like adults, when children encounter an unknown verb in a transitive sentence, they infer both that the verb's meaning involves two participant-roles and that the subject referent plays a privileged role, such as an agent's role, in the event.

The experiments just cited presented children with causal events as referents for transitive sentences. These studies therefore establish that children, like adults, readily link transitive subjects and objects with causal agents and patients. However, as we shall see, the meanings of transitive verbs extend far beyond causal action. In this paper we explore the meanings children infer from syntax, by asking what types of event-roles children can link with the subjects and objects of transitive verbs.

The answers to this question will shed light on toddlers' representations of sentence form and meaning, and their expectations about links between the two. Theoretical accounts of language acquisition differ in their assumptions on both these points. A core controversy concerns whether abstractions such as 'subject' or 'agent' are available to guide early acquisition (e.g., Gleitman et al., 2005; Pinker,

1. Note that this single-participant semantic structure could be a component of an event involving other participants (Arunachalam & Waxman, 2010; Fernandes, Marcus, Di Nubila, & Vouloumanos, 2006; Yuan et al., 2012). This follows from a rock-bottom fact about language, that verbs describe construals of events, not the events themselves (e.g., Gleitman et al., 2005; Levin & Rappaport Hovav, 2005; Pinker, 1989).

1989), or whether they emerge from initially concrete representations of language experience (e.g., Abbot-Smith, Lieven, & Tomasello, 2008; Ambridge & Lieven, 2015). Accounts that posit early abstract representations further vary in their assumptions about whether children begin with detailed innate links between syntax and semantics (e.g., Pinker, 1989), or whether many such links are learned from experience (e.g., Chang, Dell, & Bock, 2006).

1.1 Abstract vs. concrete representations of language experience

The structure-mapping account proposes both early abstract representations of syntax and semantics, and simple innate links between the two – namely, the bias to align nouns one-to-one with participant-roles. Early abstract representations play two roles in this account. First, they give children access to the proposed one-to-one mapping bias: To align nouns with participant-roles, children must represent diverse sentences in terms of their number of nouns, and diverse events in terms of their number of core participants. Second, these representations provide an abstract format for new learning, permitting the quick detection and generalization of syntactic–semantic patterns in the native language. For example, we assume children must learn the subject-verb-object word order of English, along with its role in signaling agent-patient roles. Equipped with appropriately abstract representations of form and meaning, children should detect such patterns readily, and promptly extend them to new verbs.

This early-abstraction assumption is at odds with constructivist accounts of language acquisition, which hold that early linguistic representations are word-specific (e.g., Abbot-Smith, et al., 2008; Ambridge & Lieven, 2015). According to these accounts, a child who heard (1) would at first represent this sentence in terms specific to the words it contains and the referent situation, perhaps noting that the "breaker" is named before, and the "thing-broken" after, the word *break*. This representation could not guide the child's interpretation of word order in a sentence containing a different verb, as in (2), with different event-roles ("spinners" and "spinnees") and different representations of sentence positions (before and after *spin*). On the constructivist view, children accumulate and compare word-specific representations, gradually constructing the abstract semantic and syntactic representations that allow them to detect abstract patterns including that, in English, agents are specified before, and patients after, transitive verbs. This abstraction process is predicted to be slow because it requires children to generalize across a host of irrelevant features. In contrast, if young children are biased to represent sentence form and meaning in usefully abstract terms, then the word-order pattern shared by (1) and (2) should readily be detected and extended to other verbs.

(1) The girl broke the window.

(2) The girl is spinning the top.

Strong support for early abstraction comes from the experimental evidence cited above, showing that before age two, children use word order to interpret novel verbs (e.g., Dittmar et al., 2008; Fernandes et al., 2006; Gertner & Fisher, 2012; Gertner et al., 2006). Those findings imply that young children are not limited to word-specific sentence representations. Instead, they represent linguistic experience in abstract terms, allowing them to quickly detect abstract syntactic-semantic patterns in their native language.

1.2 The semantics of subjects and objects

Deep questions remain regarding what kinds of similarities children readily detect across words and sentences, and therefore what generalizations they can apply to the interpretations of new verbs. The answers to these questions depend on the nature of early representations of sentence form and meaning, and children's expectations about links between the two. The structure-mapping account assumes no built-in links between syntax and semantics beyond the one-to-one mapping bias itself; but this is in part because the account focuses on structural cues that become available as soon as the child can identify some nouns. Once children can identify the grammatical subject of the sentence, they may gain access to more refined built-in links between syntax and semantics (e.g., Fisher et al., 1994; Fisher & Song, 2006; Hartshorne et al., 2016; Lidz, Gleitman, & Gleitman, 2003).

For instance, one proposal is that children initially and preferentially link transitivity with *causation* (Hartshorne, Pogue, & Snedeker, 2015; Kline, Snedeker, & Schulz, 2017), assuming that the subjects and objects of transitive verbs are linked with the agents and patients of causal meanings. This proposal stems in part from cross-linguistic evidence that causal meanings make good transitive verbs. Although far from all transitive verbs describe causal events, verbs that describe events in which an agent causes a change in a patient, such as change-of-state (1) or caused-motion verbs (2), are prototypical transitive verbs that are uniformly transitive across languages (e.g., Hopper & Thompson, 1980). In contrast, verbs that describe other types of two-participant events, such as verbs denoting contact actions without a specified effect (e.g., *hit, tap, hug*), or verbs of perception (e.g., *see, hear*), exhibit greater variability in how they are realized syntactically (e.g., Levin & Rappaport Hovav, 2005; Tsunoda, 1985). Hartshorne et al. (2015) and Kline et al. (2017) propose that the strong cross-linguistic tie between transitivity and causality results from learners' built-in biases: If children are biased to interpret transitive sentences as describing causal events, then they should have more trouble learning the meanings, and semantic-syntactic linking properties, of transitive verbs that do not have causal meanings.

An alternative possibility, proposed by Fisher and Song (2006; see also Landau & Gleitman, 2015), is that young children approach language ready to map a more abstract conceptual-semantic asymmetry onto the syntactic difference between subjects and non-subjects. The impetus for this proposal stems from two considerations. First, despite the variety of transitive meanings, diverse linguistic analyses suggest that the semantic roles linked to subject and object positions share an abstract semantic similarity (e.g., Dowty, 1991; Grimshaw, 1990). For example, languages link a broad class of prominent event-roles (including causal agents as noted above, but also other roles entailing sentience, volitional involvement, or motion) with the subjects of transitive verbs. Second, there is evidence that the syntactic distinction between subject and non-subject or complement phrases conveys a broad asymmetry in semantic prominence that is not limited to the subject and direct object of transitive verbs (e.g., Gleitman, Gleitman, Miller, & Ostrin, 1996; Landau & Gleitman, 2015). We next discuss each of these considerations in turn.

1.2.1 *Proto-agents and proto-patients*

Dowty (1991) laid out a contrast between two broad prototype role categories, *proto-agent* and *proto-patient*, each associated with a set of semantic entailments. Proto-agent role entailments include volition, sentience, causing a change in another participant, and movement relative to another participant; proto-patient entailments include undergoing a change of state, being causally affected by another participant, and being stationary relative to another participant's movement. For each transitive verb, the argument with the most proto-agent entailments is realized as the subject, and the argument with the most proto-patient entailments as the object. Causal action verbs exhibit a stark asymmetry in these entailments. For example, the subjects of transitive *break* and *spin* in (1) and (2) have many proto-agent entailments and few or no proto-patient entailments, while the reverse is true of the objects. A related asymmetry applies to the non-causal verbs in (3) and (4). The subject of *hug* is not a causal agent, but still has more proto-agent entailments than does the object (volition, movement relative to another participant); similarly, the subject of *see* must be sentient, whereas the object need not be.

(3) The girl hugged the boy.

(4) The girl saw the boy.

Dowty (1991) argued that these proto-roles result from conceptual categories that emerge early in cognitive development. That is, languages link to subject position the broad class of roles that can be characterized as proto-agents because children view them as similar. Subsequent research in infant cognition has yielded powerful evidence for (at least the building blocks of) suitably broad role categories. For example, pre-linguistic infants identify entities as psychological agents, thus suitable targets for psychological reasoning, if those entities appear to perceive and respond volitionally to changes in their environment (see Baillargeon, Scott, & Bian, 2016, for a review). On this view, the link between transitivity and causality may not be specific to true causality after all, but may instead be a special case of a more general propensity to link transitive subjects with agentive (causal, mobile, or volitional) roles in events.

1.2.2 *Beyond transitives: The asymmetry of syntax*

Dowty's (1991) proto-role proposal was formulated to explain the argument-linking patterns of transitive verbs, but a related semantic asymmetry characterizes the subjects and non-subjects of sentences that do not contain transitive verbs. For example, Talmy (1983) argued that the syntactic asymmetry between subject and non-subject arguments (not just transitive objects) maps onto a perspective-dependent semantic asymmetry that is somewhat independent of the event roles played by each participant. To illustrate, both sentences in (5) describe the same state of affairs; but the predicates *above* and *below* permit different choices as to which participant is construed as the conceptual figure, whose position relative to a ground or reference object is the point of the sentence. The choice of figure vs. ground depends not only on the state of affairs, but also on the focus of the speaker's attention, and what is viewed as changeable vs. stable in the referential context (e.g., Nappa, January, Gleitman, & Trueswell, 2004; Osgood & Bock, 1977). On this view, the choice of subject vs. non-subject serves as an abstract semantic framing device, highlighting one participant as a potentially changeable figure relative to a stable ground object.

(5) a. The star is above the circle.
b. The circle is below the star.

1.2.3 *The semantic prominence of subjects: Prior evidence from adults and children*

Thus, the subjects and objects of transitive verbs can be described as linked with roles bearing more of the semantic entailments of proto-agents (such as causation, motion, or volitional involvement; Dowty, 1991), or roles played by entities that are more readily viewed as the conceptual figure in a figure-ground relation (Talmy, 1983). These theoretical descriptions differ in important ways, but share the assumption that grammatical subjects are linked to arguments that are prominent in a ranking of conceptual-semantic roles. These proposals raise the intriguing possibility that children might have access to a highly abstract default interpretation for transitive verbs. For example, even when the referential scene makes a true causal interpretation unlikely, children might link transitive word order with an asymmetry in movement, volitional involvement, or suitability as a conceptual figure.

Considerable evidence suggests adults can access this abstract default interpretation. First, when given the example with which we began, *The zum mecked the zard*, adults rated *the zum* as more likely to act intentionally, or to cause a change, and rated *the zard* as more likely to undergo a change of state, or to be stationary relative to another participant (Kako, 2006). These patterns could indicate a preference for causal interpretations in particular, or the unfettered attribution of independent proto-agent entailments to the subjects of novel transitive verbs, and proto-patient entailments to their objects, just as Dowty (1991) would predict. Second, adults assign asymmetrical roles to subject and non-subject referents even when the verbs (or other predicates) in those sentences have inherently symmetrical meanings (Gleitman et al., 1996). For example, *meet* describes a symmetrical relationship (two people meet to the same degree), yet adults rated sentence (6a) as more sensible than (6b). Similarly, two objects are equally *near* each other, yet adults found (7a) more sensible than (7b). When interpreting symmetrical predicates combined with nonsense nouns (*The dax met the zum*), adults rated the referents of nouns in object position as bigger, more important, and less mobile than those in subject position. This asymmetrical interpretation must result from the syntax of the sentence, not from the symmetrical predicates themselves. Just as Talmy (1983) would predict, the subject can be interpreted as a potentially moveable figure relative to a stable ground object.

(6) a. My sister met Meryl Streep.
b. Meryl Streep met my sister.

(7) a. The bicycle is near the building.
b. The building is near the bicycle.

Experimental studies of children's interpretations of novel verbs have typically presented children with causal events as target referents for transitive sentences (e.g., Arunachalam, Escovar, Hansen, & Waxman, 2013; Fisher, 2002; Gertner et al., 2006; Gertner & Fisher, 2012; Lidz et al., 2003; Naigles, 1990; Naigles, Fowler, & Helm, 1992; Naigles, Gleitman, & Gleitman, 1993; Yuan & Fisher, 2009; Yuan et al., 2012). The results of these studies establish that even toddlers under two years old, like adults, readily map transitive sentences onto causal events, and readily interpret transitive word order relative to such events, linking transitive subjects and objects with causal agents and patients.

Given these early successes, we might suppose that children prefer causal interpretations of transitive syntax, and have trouble accessing a more abstract default interpretation. Two findings have been offered as support for this hypothesis. First, Kline et al. (2017) found that 4-year-olds mapped novel verbs in transitive (but not intransitive) sentences onto events that could be construed as causal (e.g., an actor claps, then a toy spins), as opposed to similar events in which the perception of causality was disrupted by a spatio-temporal gap between action and effect or by reordering the events so that the effect came first (the toy spins, then the actor claps). Thus, children prefer to link transitive sentences with causal events, at least over certain kinds of non-causal events. Second, Hartshorne et al. (2015) found that 4-year-olds more reliably used word order to determine who did what to whom in sentences containing causal emotion verbs: After hearing a story, children more accurately evaluated the truth of sentences containing emotion verbs denoting the causation of emotion (*Tiger frightened Lion* vs. *Lion frightened Tiger*) than those denoting non-causal emotional states (*Tiger loves Lion* vs. *Lion loves Tiger*). Hartshorne et al. argued that 4-year-olds did not yet understand the argument-role linkings of non-causal emotion verbs such as *love*, because they expect transitive verbs to have causal meanings. Another possibility, however, is that children had trouble with non-causal emotion verbs in this task partly because these verbs denote states rather than actions; the actions in 'frighten' stories may have suggested stronger causal links between propositions than did the states in 'love' stories, facilitating story comprehension and recall (Bransford & Johnson, 1972).

Other data suggest that children can also map transitive verbs onto non-causal meanings. For example, invented verbs in transitive frames (*The duck is kradding the bunny!*) prompted 29-month-olds to look longer at non-causal contact events (e.g., a duck touched a bunny's head) as opposed to synchronous events (both duck and bunny wheeled their arms), relative to their peers who heard the verbs in intransitive frames or in isolation (Naigles & Kako, 1993). This suggests that toddlers viewed non-causal contact events as coherent two-participant events, thus plausible referents for a novel verb in a transitive frame (see also Naigles, 1996; Scott & Fisher, 2009). However, these studies did not test whether children systematically mapped transitive subjects and objects onto participant-roles in non-causal contact events.

Other evidence suggests older preschoolers can use word-order to map the arguments of novel verbs onto roles in non-causal events. Fisher and Song (2006) showed 3-year-olds non-causal motion or location scenes (e.g., a toy car rolling up to a flashlight) accompanied by transitive sentences containing ambiguous pronouns and nonsense verbs (e.g., *It's pilking it.*); scene participants varied in animacy.

When asked, “Which one is pilking the other one?”, children tended to select moving rather than immobile objects and animate rather than inanimate objects. These choices are just what we should expect if children preferred to map subjects onto event roles with more proto-agent properties (e.g., volition, movement relative to another participant), and direct objects onto roles with more proto-patient properties. Perhaps most strikingly, 4-year-olds assign asymmetrical interpretations to the subject versus non-subject arguments of symmetrical predicates (Chestnut & Markman, 2016): Children heard sentences with nonsense nouns and symmetrical predicates (*A blicket is like a toma; The plig is next to the fem.*), and were asked to choose the referents of the novel nouns in a picture. Children’s choices revealed that, like adults, they aligned subjects and non-subjects with an abstract figure-ground asymmetry, and thus inferred that subject referents were less typical (a zebra is like a horse, rather than the reverse) or smaller and more movable (the bicycle is next to the building) than non-subject referents. Taken together, these data show that 3- and 4-year-olds can interpret the syntactic distinction between subject and non-subject in terms of highly abstract categories of semantic roles – perhaps proto-agent vs. proto-patient, or figure vs. ground.

1.3 The present research

Building on this evidence, we asked whether much younger children, just under two years old, could also use word order to map novel transitive verbs onto two-participant relationships beyond prototypical causality. In two verb-learning experiments, we tested toddlers in a preferential-looking task in which they heard a single novel verb in a transitive sentence, while viewing a pair of candidate referent events. In Experiment 1, we took a modest step away from the prototypical caused-motion events examined in previous studies of early sensitivity to word order (e.g., Gertner et al., 2006): children saw events in which one participant acted on another but produced no clear change of position or posture in the recipient of this action (see Figure 1). In Experiment 2, children saw events in which one participant simply moved relative to another without contacting it. If toddlers, like older children and adults, can align broad semantic categories such as proto-agent vs. proto-patient with transitive subjects and objects, then they should look longer at the target event – the event in which the subject of the test sentence plays a role with more of the features of proto-agency, such as volitional action and motion relative to another participant.

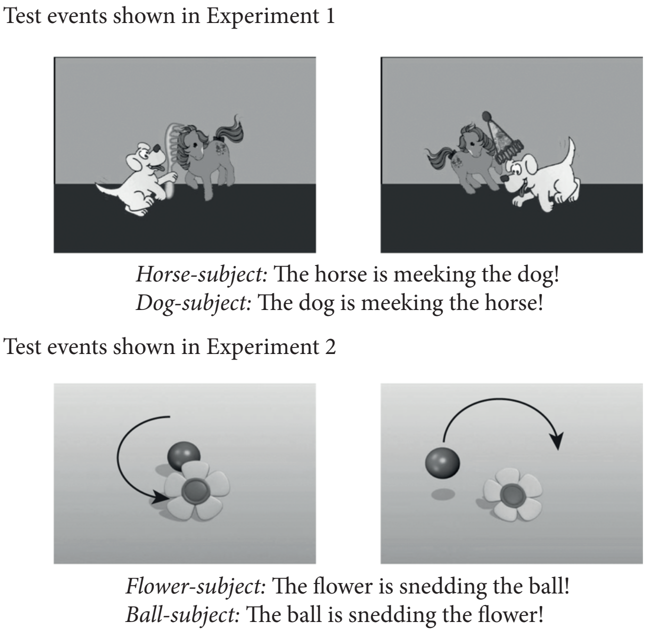


Figure 1. Event-pairs and sentences for the novel-verb test trials in Experiments 1 and 2. Arrows superimposed on the event-pair for Experiment 2 sketch the path of motion in these events

2. Experiment 1

In Experiment 1, we tested 23-month-olds' comprehension of English word-order in reference to two novel 'grooming' events. In one event, a horse placed a hat on a dog's head; in the other, the dog passed a comb over the horse's mane (Figure 1). These events were accompanied by a transitive sentence containing an invented verb. For half the children, the horse was the subject of the sentence ("The horse is meeking the dog!"); for the other half, the dog was the subject ("The dog is meeking the horse!").

These events involved the action of one participant on another, but differed from the caused-motion events examined in previous work in that they involved no change in the position or posture of the recipient of action. English has transitive verbs that describe dressing (*dress, garb*) or grooming (*groom, curry*). But such verbs are few, and many transitive verbs describing acts of bodily care specify a body part rather than the whole body as direct object (*brush his hair*; Levin, 1993). The events labeled by such verbs could be construed as causal; after all, the goal is typically to bring about a dressed or groomed state. But unlike causal transitive verbs such as *break* or *spin*, many such verbs do not entail the achievement of any particular intended outcome. For example, consider the fictional character Harry Potter, who is known for his notoriously messy hair. One could comb his hair in vain, never achieving a well-groomed state.

These are contact-action events, like those that Naigles and Kako (1993) contrasted with synchronous-action events. Their data, discussed above, showed that 2-year-olds readily viewed non-causal contact events as coherent two-participant events, and thus plausible referents for a novel verb in a transitive as opposed to an intransitive frame. However, Naigles and Kako did not test whether children systematically mapped transitive subjects and objects onto participant-roles in such events. Although these contact-action events lacked the clear causal effects shown in most previous work, their participants' roles still suggest a strong asymmetry in proto-agent vs. proto-patient entailments, and in likely construal as figure vs. ground. For example, the proto-agent entailments of the actor's role in these events are volitional involvement, and motion relative to another participant. If 23-month-olds' interpretation of transitive word-order extends beyond the prototypical caused-motion events examined thus far, then they should be able to use word order in a transitive sentence to identify the target event, and thus look longer at the event in which the subject of their test sentence performs the action.

2.1 Method

2.1.1 *Participants*

Twenty-four 23-month-olds participated (mean 23.1, range 22.0–24.0, 13 male, 11 female). All were native speakers of English. Five additional children were excluded due to inattentiveness (1), failure to complete the experiment (1), or because their average match proportion (see *Coding* below) was more than 2.5 SD from the mean of their condition (3). Children's productive vocabulary was measured using the short form of the Bates-Macarthur CDI, Level 2 (Fenson, Pethick, Renda, Cox, Dale, & Reznick, 2000). Vocabulary scores ranged from 17 to 80 with a median of 52.

2.1.2 *Apparatus*

Children sat on a parent's lap facing two 20-inch television screens placed about 30 inches away. The screens were 12 inches apart and at child's-eye level. Soundtracks played from a central speaker. A camera hidden between the two monitors recorded children's eye movements during the experiment. Parents wore opaque sunglasses, preventing them from biasing their children's responses.

2.1.3 *Materials and procedure*

Stimulus materials were simple animated videos involving a cartoon horse and dog. Events were shown in synchronized pairs and accompanied by a soundtrack recorded by a native English speaker. The procedure had three phases: character-familiarization, practice, and test.

In the character-familiarization phase, the horse was shown moving across the screen (6s) and was labeled twice ("There's a horse! Look at the horsie!") while the other screen remained blank. Following a 2-s interval, the dog was introduced on the other screen in the same manner ("There's a dog! Look at the doggie!"). This was followed by two 6-s trials, separated by 2-s blank-screen intervals. In each trial, the horse and dog appeared on their respective screens simultaneously. In the first trial, children were asked to "Find the dog!"; in the second trial they were instructed to "Find the horse!"

In the practice phase, two familiar transitive verbs were presented accompanied by familiar events. The first familiar verb, *feed*, was accompanied by a feed event in which the horse fed the dog with a bottle, and an approach event, in which the horse approached the dog and stopped a short distance away. Children were introduced to these events in two previews. In the first preview, children saw the feed event on one screen (6s) accompanied by neutral audio (e.g., "Look here! Watch this!") while the other screen remained blank. Following a 2-s blank-screen interval, the approach event was previewed in the same manner on the other screen. Next, during a 6-s blank-screen interval, children heard, "Now watch! The horse is gonna feed the dog!" Children then saw the feed and approach events simultaneously (6s) and heard the sentence, "The horse is feeding the dog" repeated twice. This was followed by a 6-s blank-screen interval in which children heard, "The horse fed the dog. Find feeding." Children then saw both videos again (6s) and heard, "The horse is feeding the dog. Find feeding." Following a 3-s blank-screen interval this procedure was repeated with a second familiar verb, *lick*; one event showed the dog licking the horse and the other showed the dog moving away from the horse.

The character-familiarization and practice phases were designed to familiarize the children with the two characters, the wording of the test trials, and the fact that one of the two events matched the soundtrack on each trial. *Horse, dog, feed*, and *lick* were used because these words are likely to be familiar to 2-year-olds (Dale & Fenson, 1996). Practice events were paired so that the same character was the actor on both screens; thus, children had to use knowledge of the verb, rather than word order, to identify the matching screen in the practice trials.

Finally, in the test phase, children were presented with the novel verb *meek* accompanied by a pair of test events. In the hat event, the horse placed a hat on the dog's head. In the comb event, the dog passed a comb once over the horse's mane. The left/right position of the actor in the test events was counterbalanced within sentence condition using a "flip" operation in the video-editing software.

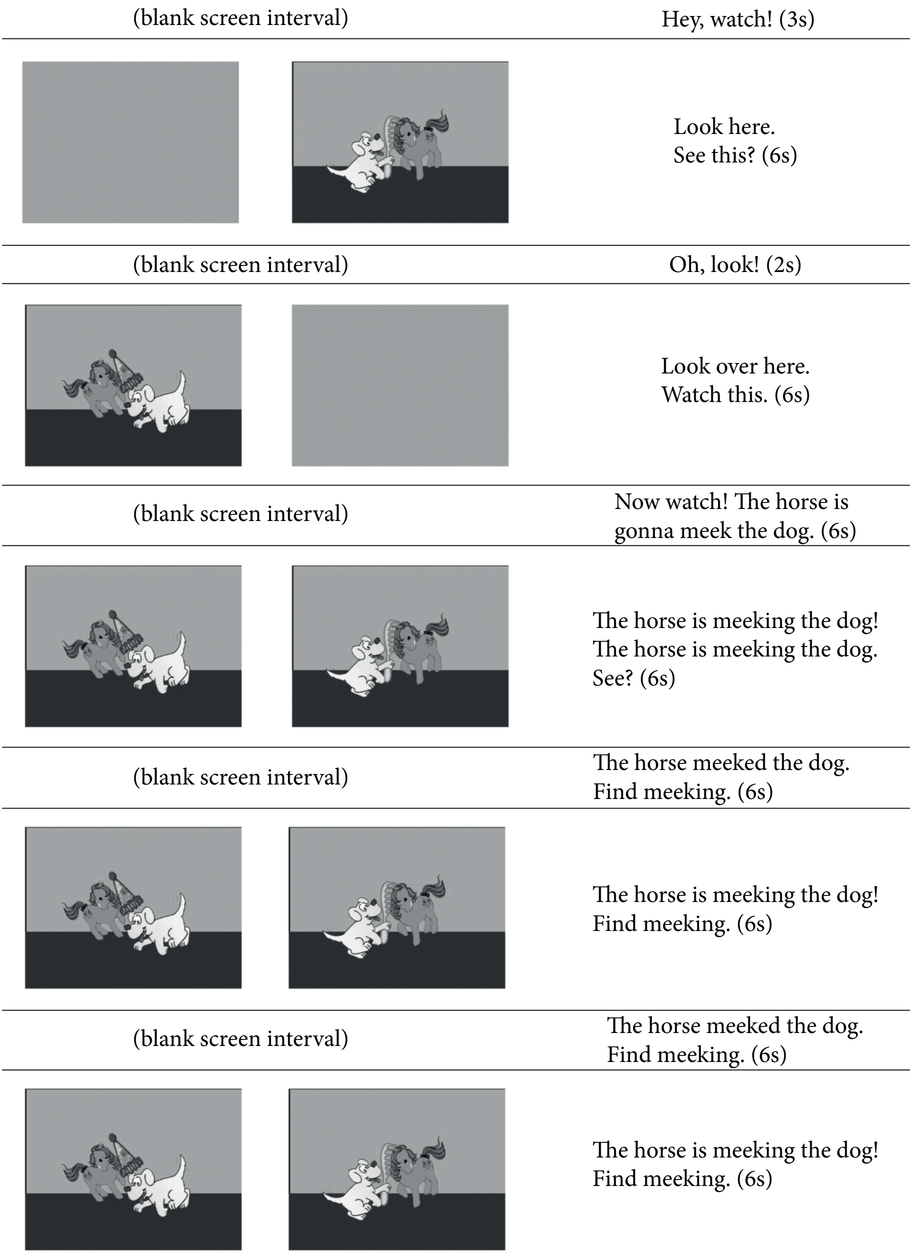


Figure 2. The sequence of events during the test phase in Experiment 1

The sequence of events within the test phase is shown in Figure 2. The test events were first previewed separately, as described above for the practice phase. Next, during a 6-s blank-screen interval, children heard the novel verb *meek* in a transitive sentence. Half of the children heard, "The horse is gonna meek the dog" and half heard "The dog is gonna meek the horse." This was followed by three 6-s test trials in which children saw the two test events and heard the novel verb used in sentences with the appropriate subject (*horse* or *dog*). Test trials were separated by 6-s blank-screen intervals in which children heard additional novel-verb sentences. In total, children heard the novel verb presented 7 times, before, during and after 3 presentations of the test events. Multiple verb forms (*gonna meek, meeking, meeked*) were included to provide ample evidence that the new word was a verb, and to suit the timing of stimulus sentences relative to presentation of the events.

The left-right positioning of events in the character-familiarization, practice, and test phases was counter-balanced with test-sentence subject and with the left-right side of the actor in the test events.

2.1.4 *Coding*

We coded where children looked (left-screen, right-screen, away) frame by frame from silent video. To assess reliability, 6 children's data were independently coded by a second coder. The two coders agreed on the children's direction of gaze for 95% of video frames. Trials in which children looked away from the two screens for more than 67% of the trial were eliminated from analyses (1/96 practice trials, 5/72 test trials). We analyzed the proportion of looking time to the event that matched the audio out of the total time spent looking at either test event, averaged across the three test trials. Analyses based on mean looking times to the matching and non-matching events rather than proportions revealed the same pattern of significant effects as the main analyses reported below.

Preliminary analyses of children's looking-time performance in the test trials revealed no effects of sex, sentence subject, or whether the child's vocabulary or performance in the practice trials was above or below the median, all Fs < 2.07, all ps $> .16$[2]. We therefore collapsed across these factors in subsequent analyses.

2. Preliminary analyses revealed a marginal effect of actor side, $F(1, 22) = 3.30$, $p = .08$. Children looked marginally longer at the matching event if the actor was on the right side of the event ($M = .68$, $SD = .19$) than if the actor was on the left side of the event ($M = .54$, $SD = .20$).

2.2 Results and discussion

Averaged across the three 8s test trials, children looked significantly longer at the matching video, the one in which the subject of the sentence performed the contact action, than expected by chance ($M = .61$, $SD = .20$), $t(23) = 2.58$, $p = .017$, $d = .54$. Children's preference for the matching event was not significantly correlated with their vocabulary score, $r(24) = -.22$, $p = .30$. Examination of individual children's performance indicated that 7/11 of the children with vocabulary scores above the median and 8/13 children with vocabulary scores at or below the median looked longer at the matching event.

Even though our two test events did not depict a clear causal effect, 23-month-olds used their knowledge of English word order to arrive at a sensible interpretation of a novel transitive verb. They might have done so by inferring that the subject of the sentence referred to an event participant whose role had more proto-agent properties (volition, movement relative to another participant), while the object of the sentence referred to the participant whose role had more proto-patient properties (stationary relative to another participant).

3. Experiment 2

In Experiment 1 we showed for the first time that 23-month-olds could use their understanding of English word order to interpret transitive sentences that depart from the prototypical caused-motion events examined in earlier work. However, the contact-action events of Experiment 1 constituted only a modest step away from the clear caused-motion events examined in previous studies. In Experiment 2, we sought to confirm and extend our finding by asking whether 23-month-olds could systematically map a transitive sentence onto non-causal events that involved *no contact* between event participants, but simply the motion of one participant relative to the other. As in Experiment 1, children watched pairs of animated events (Figure 1). In one event, a ball jumped back and forth over a flower while the flower tilted gently in place; in the other event, the flower moved around the ball in a circle as the ball bounced gently in place. While viewing these events, half of the children heard "The ball is snedding the flower", and half heard "The flower is snedding the ball".

The roles of the ball and flower in these events lacked most of the features of prototypical causality. Neither participant acted on or touched the other, and neither appeared to cause a change in the other's motion. Nonetheless, one participant in each event still exhibited more proto-agent properties than the other because it underwent a salient motion relative to the other participant.

English has transitive verbs that could describe similar actions (*hurdle* and *circle*, though children are unlikely to know these verbs), as well as other transitive motion verbs that express the path of motion relative to a ground object (*enter the room, cross the street*). Such verbs are atypical transitive verbs (i.e., they have few proto-agent and proto-patient entailments), and transitive verbs of this sort are atypical motion descriptions in English. English motion verbs tend to express manner of motion, with path of motion expressed via prepositional phrases (*jumping over the flower, going around the ball*; Naigles & Terrazas, 1998; Papafragou, Massey, & Gleitman, 2006; Talmy, 1985). However, prior evidence suggests that both preschoolers and adults can interpret transitive verbs as referring to motion relative to a reference object (Hohenstein, Naigles, & Eisenberg, 2004; Naigles & Terrazas, 1998; Wagner, 2010). For example, Hohenstein et al. (2004) presented 3.5-year-olds with directed motion events (e.g., a woman skipping toward a tree) accompanied by transitive verbs ("She's kradding the tree") or intransitive verbs with a prepositional phrase ("She's kradding toward the tree"). Syntactic structure guided children's extensions of these verbs to new events. Children who heard transitive sentences looked longer at new test events that preserved the path (tree-ward) but not the manner (skipping) of the original event, relative to those who heard intransitive sentences. This suggests that older children and adults readily interpreted directed-motion events as two-participant events and thus suitable referents for a transitive verb; but these studies did not test whether children systematically mapped transitive subjects and objects onto the moving participants and reference objects in such events.

If 23-month-olds, like older children and adults, map subjects and objects onto broad semantic role classes such as proto-agent vs. proto-patient, or conceptual figure vs. ground, then they should look longer at the target event in this task – the event in which the subject of their test sentence is the more active, mobile participant.

Experiment 2 was also designed to address an alternative interpretation of the word-order findings of Experiment 1. In Experiment 1, the test phase was preceded by a practice phase in which children heard two familiar transitive verbs with the test characters (the horse and dog) named in subject and object position. Dittmar et al. (2008) argued that previous results showing early sensitivity to syntax (Gertner et al., 2006) could be attributed to within-experiment learning from practice items that were similar to the novel-verb test items. To rule out this possibility, in Experiment 2 we removed the familiar-verb practice trials altogether.

3.1 Method

3.1.1 *Participants*

Twenty-four 23-month-olds participated (mean 22.8, range 21.8–23.9, 13 male, 11 female). All were native speakers of English. Three additional children were excluded due to inattentiveness (1), because their productive vocabulary was below the fifth percentile for their age (1), or because their average match proportion was more than 2.5 SD from the mean of their condition (1). Children's productive vocabulary scores, measured as in Experiment 1, ranged from 16 to 91 with a median of 55.

3.1.2 *Apparatus*

The apparatus was identical to that used in Experiment 1.

3.1.3 *Materials and procedure*

Stimulus materials were animated videos involving a ball and a flower, created using the 3D animation software Maya®. Events were shown in synchronized pairs and accompanied by a soundtrack recorded by a native English speaker.

The procedure had two phases: character-familiarization and test. In the character-familiarization phase, a ball was shown in the center of one screen (7s) and was labeled twice ("There's a ball! Look at the ball!") while the other screen remained blank. Following a 2-s interval, a flower was introduced on the other screen in the same manner ("There's a flower! Look at the flower!"). This was followed by two 7-s trials, separated by 2-s blank-screen intervals. In each trial, the ball and flower appeared simultaneously on their respective screens. In the first trial, children were asked to "Find the flower!"; in the second trial they were instructed to "Find the ball!"

Next, in the test phase, children were presented with the novel verb *sned* accompanied by a pair of test events. In the hurdle event, the ball jumped back and forth over the flower while the flower tilted gently in place; in the circle event, the flower circled the ball while the ball bounced gently in place.

The sequence of events within the test phase is shown in Figure 3. The test events were first introduced separately in two previews. In the first preview, children saw the hurdle event on one screen (7s) accompanied by neutral audio (e.g., "Look here! See this?"); the other screen remained blank. Following a 2-s blank-screen interval, the circle event was previewed in the same manner on the other screen. Next, during a 6-s blank-screen interval, children heard the novel verb *sned* in a transitive sentence. Half of the children heard "The ball is gonna sned the flower," and half heard "The flower is gonna sned the ball." This was followed by three 7-s test trials in which children saw the two test events and heard the novel verb used in sentences with the appropriate subject (ball or flower). Test trials were separated by 6-s blank-screen intervals in which children heard additional novel-verb sentences in the past tense. As in Experiment 1, children heard the novel verb presented 7 times and in three different tenses.

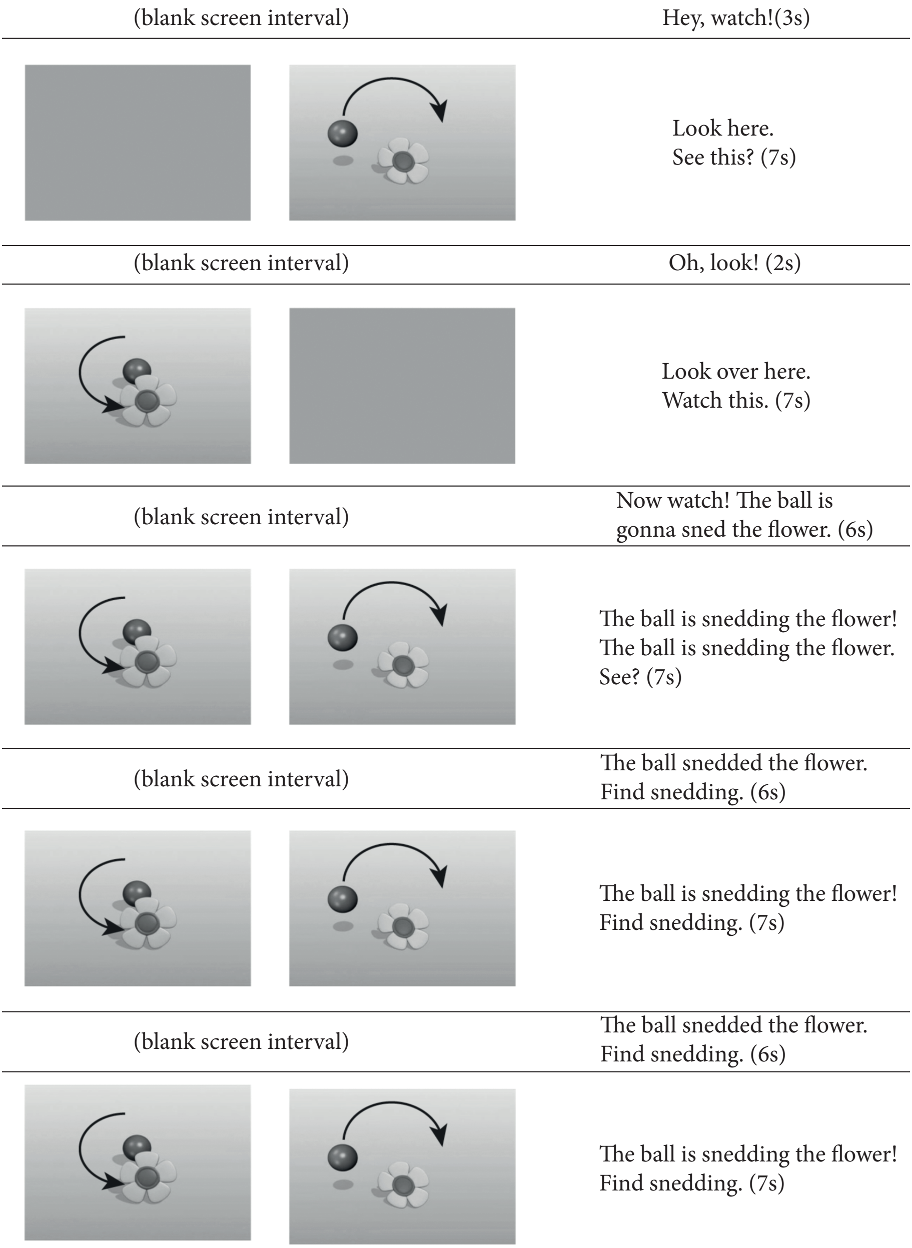


Figure 3. The sequence of events during the test phase in Experiment 2

The left-right positioning of the sequence of events in the character-familiarization and test phase was counter-balanced with test-sentence subject.

3.1.4 *Coding*

As in Experiment 1, we coded where children looked (left-screen, right-screen, away), frame by frame from silent video. To assess reliability, 4 children's data were independently coded by a second coder. The two coders agreed on the children's direction of gaze for 97% of coded video frames. We analyzed the proportion of looking time to the event that matched the audio out of the total time spent looking at either test event, averaged across the three test trials. Analyses based on mean looking times to the matching and non-matching events rather than proportions revealed the same pattern of significant effects as the main analyses reported below.

Preliminary analyses of children's looking time performance in the test trials revealed no effects of sex, sentence subject, or whether the child's vocabulary or performance in the practice trials was above or below the median, all Fs < 1. We therefore collapsed across these factors in subsequent analyses.

3.2 Results and discussion

Averaged across the three test trials, children looked significantly longer at the matching video, the one in which the subject of the sentence moved more relative to the other participant, ($M = .57$, $SD = .14$) than expected by chance, $t(23) = 2.56$, $p = .018$, $d = .50$. Children's proportion of looking time to the matching video was not significantly correlated with their vocabulary scores, $r(24) = .09$, $p = .67$. Examination of individual children's performance indicated that 6/11 of the children with vocabulary scores above the median and 10/13 children with vocabulary scores at or below the median looked longer at the matching event.

Even though our two test events depicted non-causal motion events, 23-month-olds again used their knowledge of English word-order to arrive at a sensible interpretation of a transitive sentence containing a novel verb.

4. General discussion

In two verb-learning experiments, we probed what kinds of asymmetries in event roles children map onto word order in transitive sentences. To do so, we created stimulus events lacking key features of the prototypical caused-motion events examined in previous studies. In our events, although there was no causal agent, one participant's role entailed more of the semantic entailments of proto-agency than the other (motion or volitional involvement; Dowty, 1991). In both experiments, despite the absence of 'real' causal agents, 23-month-olds interpreted word order systematically: they looked longer at the event in which the subject of the sentence was the more mobile participant. This suggests that 23-month-olds readily link transitive word order with broad event-role categories that extend beyond the prototypical agent-patient case to include cases in which one character contacts another, or simply moves relative to another. Our findings add to what we know about early interpretation of transitive word order in three ways.

First, children's success in our word-order tasks confirms and extends prior evidence that 2-year-olds can link transitive sentences with contact-action events (Naigles & Kako, 1993), and that older preschoolers and adults can link transitive sentences with directed-motion events (e.g., "She's kradding the tree!"; Hohenstein et al., 2004; Naigles & Terrazas, 1998; Wagner, 2010). The systematic responses in our task suggest that by 23 months, children can view contact-action and directed-motion events as *two-participant events*, thus appropriate referents for a transitive verb.

Second, the present studies were the first to show that toddlers mapped transitive subjects and objects systematically onto participant-roles in non-causal events. In both experiments, children linked the subject of their test sentence with the participant in the event that initiated the contact (Experiment 1), or that moved more saliently relative to the other participant (Experiment 2). This suggests that they linked a highly abstract semantic-role asymmetry with the subject and object of the transitive test sentences.

Third, in Experiment 2 children interpreted word-order systematically despite the absence of familiar-verb practice trials. The test phase was preceded only by a character-familiarization phase; thus, children had no opportunity to learn shallow sentence-interpretation strategies within the experiment that could have supported test-item performance. This tells us that the 23-month-olds brought to the task some knowledge of the semantic significance of English word order; this knowledge was robust enough to influence interpretation of a novel verb, and abstract enough to guide children's attention toward an event in which the subject of the sentence was not a causal agent, but merely mobile.

The early recruitment of this abstract grouping of semantic roles constrains theories of early syntax acquisition. Constructivist accounts propose that abstract semantic-role categories emerge through language experience: children learn to see diverse event-roles as similar because their language expresses them similarly (e.g., Abbot-Smith et al., 2008). Early-abstraction accounts propose that language learning builds on the child's own tendency to view broad classes of semantic roles as similar, and to assume any language will treat them similarly (e.g., Chang et al., 2006; Gleitman et al., 2005; Pinker, 1989). Toddlers' willingness to link abstract role categories with transitive word order challenges the constructivist account, but is easily handled by an early-abstraction account.

Similarly, the abstract nature of the semantic roles children infer from syntax has consequences for early syntactic bootstrapping. Via structure-mapping, once children can identify some nouns, they infer that sentences containing two nouns refer to two-participant construals of events. Given abstract representations of form and meaning, children generalize what they learn about how one verb arranges its arguments in the syntax to other verbs. The present results suggest that 2-year-olds readily link broad classes of semantic roles with word order in transitive sentences, and thus that early extensions of syntactic knowledge across verbs may be correspondingly broad. For example, if proto-agent and -patient roles form part of the starting point for language acquisition, then any learning about *hitting* and *breaking* should promptly guide learning about *hugging* and *seeing*.

Dowty's (1991) proto-role proposal could easily explain our results. At the core of this proposal is a claim that proto-agent and proto-patient role categories emerge from non-linguistic cognitive development, and thus are available to guide the inferences of young learners. Causal verbs exhibit the strongest asymmetry in proto-agent vs. proto-patient entailments, but our stimulus events were designed to strip away some of these entailments. Nonetheless, one participant in each event still made a better proto-agent than the other because it appeared to volitionally contact the other (in Experiment 1), or underwent a salient motion relative to the other participant (in Experiment 2).

However, given the arguments and evidence reviewed in the Introduction, we suspect this is unlikely to be the whole story, and that children might have access to a default interpretation of the subject-object asymmetry that extends beyond the semantic-role entailments of proto-agents and proto-patients. Recall that 4-year-olds and adults preferred interpretations such as 'the zebra is like the horse' over 'the horse is like the zebra' (Chestnut & Markman, 2016; Gleitman et al., 1996). Because *is like* has a symmetrical meaning (and thus its two arguments play the same roles), the perceived semantic asymmetry in these sentences must come from the syntax, not from the semantic role entailments of the predicate. As Talmy (1983) noted, the choice of subject vs. non-subject depends not only on the event-dependent participant-roles associated with the verb, but also on which of the participants filling those roles makes a better conceptual figure in a figure-ground relation. Because more mobile event-participants make better conceptual figures as well as proto-agents, Talmy's figure-ground asymmetry could also explain our results. Disentangling these two explanations will require investigating the early learning of a wide range of predicate meanings, including spatial terms and symmetrical predicates (e.g., Landau & Gleitman, 2015). Whatever the result of these investigations, the current findings suggest that young children have access to a flexible range of default interpretations for transitive verbs. Even when the referential scene makes a causal construal unlikely, toddlers systematically map transitive word order onto a broad asymmetry in semantic roles.

Acknowledgments

This research was supported by the National Institute of Child Health and Human Development (HD054448) and the National Science Foundation (BCS-1348522). We thank Renée Baillargeon for helpful comments. Correspondence concerning this article should be addressed to Rose Scott, Department of Psychology, University of California, Merced, 5200 North Lake Rd., Merced, CA, 95343. Electronic mail may be sent to rscott@ucmerced.edu.

References

Abbot-Smith, K., Lieven, E. V. M., & Tomasello, M. (2008). Graded representations in the acquisition of English and German transitive constructions. *Cognitive Development*, 23, 48–66. https://doi.org/10.1016/j.cogdev.2007.11.002

Ambridge, B., & Lieven, E. (2015). A constructivist account of child language acquisition. In B. MacWhinney & W. O'Grady (Eds.), *The handbook of language emergence* (pp. 478–510). Hoboken, NJ: John Wiley & Sons.

Arunachalam, S. (2013). Two-year-olds can begin to acquire verb meanings in social impoverished contexts. *Cognition*, 129, 569–573. https://doi.org/10.1016/j.cognition.2013.08.021

Arunachalam, S., Escovar, E., Hansen, M. A., & Waxman, S. (2013). Out of sight, but not out of mind: 21-month-olds use syntactic information to learn verbs even in the absence of a corresponding event. *Language and Cognitive Processes*, 28, 417–425. https://doi.org/10.1080/01690965.2011.641744

Arunachalam, S. & Waxman, S. R. (2010). Meaning from syntax: Evidence from 2-year-olds. *Cognition*, 114, 442–446. https://doi.org/10.1016/j.cognition.2009.10.015

Baillargeon, R., Scott, R. M., & Bian, L. (2016). Psychological reasoning in infancy. *Annual Review of Psychology*, 67, 159–186. https://doi.org/10.1146/annurev-psych-010213-115033

Bransford, J. D., & Johnson, M. K. (1972). Contextual prerequisites for understanding: Some investigations of comprehension and recall. *Journal of Verbal Learning and Verbal Behavior*, 11, 717–726. https://doi.org/10.1016/S0022-5371(72)80006-9

Chang, F., Dell, G. S., & Bock, K. (2006). Becoming syntactic. *Psychological Review*, 113, 234–272. https://doi.org/10.1037/0033-295X.113.2.234

Chestnut, E. K., & Markman, E. M. (2016). Are horses like zebras, or vice versa? Children's sensitivity to the asymmetries of directional comparisons. *Child Development*, 87, 568–582.

Dale, P. S., & Fenson, L. (1996). Lexical development norms for young children. *Behavior Research Methods, Instruments, & Computers*, 28, 125–127. https://doi.org/10.3758/BF03203646

Dittmar, M., Abbot-Smith, K., Lieven, E., & Tomasello, M. (2008). Young German children's early syntactic competence: A preferential looking study. *Developmental Science*, 11, 575–582. https://doi.org/10.1111/j.1467-7687.2008.00703.x

Dowty, D. (1991). Thematic proto-roles and argument selection. *Language*, 67, 547–619. https://doi.org/10.1353/lan.1991.0021

Fenson, L., Pethick, S., Renda, C., Cox, J. L., Dale, P. S., & Reznick, J. S. (2000). Short-form versions of the MacArthur communicative development inventories. *Applied Psycholinguistics*, 21, 95–115. https://doi.org/10.1017/S0142716400001053

Fernandes, K. J., Marcus, G. F., Di Nubila, J. A., & Vouloumanos, A. (2006). From semantics to syntax and back again: Argument structure in the third year of life. *Cognition*, 100, B10–B20. https://doi.org/10.1016/j.cognition.2005.08.003

Fisher, C. (2002). Structural limits on verb mapping: The role of abstract structure in 2.5-year-olds' interpretations of novel verbs. *Developmental Science*, 5, 56–65. https://doi.org/10.1111/1467-7687.00209

Fisher, C., Gertner, Y., Scott, R. M., Yuan, S. (2010). Syntactic bootstrapping. *Wiley Interdisciplinary Reviews: Cognitive Science*, 1, 143–149.

Fisher, C., Hall, D. G., Rakowitz, S., & Gleitman, L. R. (1994). When it is better to receive than to give: Syntactic and conceptual constraints on vocabulary growth. *Lingua*, 92, 333–375. https://doi.org/10.1016/0024-3841(94)90346-8

Fisher, C. & Song, H. (2006). Who's the subject? Sentence structure and verb meaning. In K. Hirsh-Pasek & R. Golinkoff (Eds.), *Action meets word: How children learn verbs* (pp. 392–426). Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195170009.003.0016

Gertner, Y., & Fisher, C. (2012). Predicted errors in children's early sentence comprehension. *Cognition*, 125, 85–94. https://doi.org/10.1016/j.cognition.2012.03.010

Gertner, Y., Fisher, C., & Eisengart, J. (2006). Learning words and rules: Abstract knowledge of word order in early sentence comprehension. *Psychological Science*, 17, 684–691. https://doi.org/10.1111/j.1467-9280.2006.01767.x

Gleitman, L. R., Cassidy, K., Nappa, R., Papafragou, A., & Trueswell, J. C. (2005). Hard words. *Language Learning and Development*, 1, 23–64. https://doi.org/10.1207/s15473341lld0101_4

Gleitman, L. R., Gleitman, H., Miller, C., & Ostrin, R. (1996). Similar, and similar concepts. *Cognition*, 58, 321–376. https://doi.org/10.1016/0010-0277(95)00686-9

Grimshaw, J. (1990). *Argument structure*. Cambridge, MA: The MIT Press.

Hartshorne, J. K., O'Donnell, T. J., Sudo, Y., Uruwashi, M., Lee, M., & Snedeker, J. (2016). Psych verbs, the linking problem, and the acquisition of language. *Cognition*, 157, 268–288. https://doi.org/10.1016/j.cognition.2016.08.008

Hartshorne, J. K., Pogue, A., & Snedeker, J. (2015). Love is hard to understand: The relationship between transitivity and caused events in the acquisition of emotion verbs. *Journal of Child Language*, 42, 467–504. https://doi.org/10.1017/S0305000914000178

Hohenstein, J. M., Naigles, L. R., & Eisenberg, A. R. (2004). Keeping verb acquisition in motion: A comparison of English and Spanish. In D. G. Hall & S. R. Waxman (Eds.), *Weaving a lexicon* (pp. 569–602). Cambridge, MA: The MIT Press.

Hopper, P. J., & Thompson, S. A. (1980). Transitivity in grammar and discourse. *Language*, 56, 251–295. https://doi.org/10.1353/lan.1980.0017

Jin, K. S. (2015). The role of syntactic and discourse information in verb learning (Unpublished doctoral dissertation). University of Illinois at Urbana-Champaign.

Kako, E. (2006). Thematic role properties of subjects and objects. *Cognition*, 101, 1–42. https://doi.org/10.1016/j.cognition.2005.08.002

Kline, M., Snedeker, J., & Schulz, L. (2017). Linking language and events: Spatiotemporal cues drive children's expectations about the meanings of novel transitive verbs. *Language Learning and Development*, 13, 1–23.

Landau, B., & Gleitman, L. R. (2015). Height matters. In I. Toivonen, P. Csúri, & E. van der Zee (Eds.), *Structures in the mind: Essays on language, music, and cognition in honor of Ray Jackendoff* (pp. 187–210). Cambridge, MA: The MIT Press.

Levin, B. (1993). *English verb classes and alternations: A preliminary investigation*. Chicago, IL: The University of Chicago Press.

Levin, B., & Rappaport-Hovav, M. (2005). *Argument realization*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511610479

Lidz, J., Gleitman, H., & Gleitman, L. (2003). Understanding how input matters: Verb learning and the footprint of Universal Grammar. *Cognition*, 87, 151–178. https://doi.org/10.1016/S0010-0277(02)00230-5

Naigles, L. (1990). Children use syntax to learn verb meaning. *Journal of Child Language*, 17, 357–374. https://doi.org/10.1017/S0305000900013817

Naigles, L. R. (1996). The use of multiple frames in verb learning via syntactic bootstrapping. *Cognition*, 58, 221–251. https://doi.org/10.1016/0010-0277(95)00681-8

Naigles, L., Fowler, A., & Helm, A. (1992). Developmental changes in the construction of verb meanings. *Cognitive Development*, 7, 403–427. https://doi.org/10.1016/0885-2014(92)80001-V

Naigles, L., Gleitman, H., & Gleitman, L. (1993). Children acquire word meaning components from syntactic evidence. In E. Dromi (Ed.), *Language and cognition: A developmental perspective* (pp. 104–140). Norwood, NJ: Ablex.

Naigles, L. R., & Kako, E. T. (1993). First contact in verb acquisition: Defining a role for syntax. *Child Development*, 64, 1665–1687. https://doi.org/10.2307/1131462

Naigles, L. R., & Terrazas, P. (1998). Motion-verb generalizations in English and Spanish: Influences of language and syntax. *Psychological Science*, 9, 363–369. https://doi.org/10.1111/1467-9280.00069

Nappa, R. L., January, D., Gleitman, L. R., & Trueswell, J. (2004). Paying attention to attention: Perceptual priming effects on word order. In *Proceedings of the 17th Annual CUNY Sentence Processing Conference*, New York.

Osgood, C. E., & Bock, J. K. (1977). Salience and sentencing: Some production principles. In S. Rosenberg (Ed.), *Sentence production: Developments in research and theory* (pp. 89–140). Hillsdale, NJ: Lawrence Erlbaum Associates.

Papafragou, A., Massey, C., & Gleitman, L. (2006). When English proposes what Greek presupposes: The cross-linguistic encoding of motion events. *Cognition*, 98, B75–B87. https://doi.org/10.1016/j.cognition.2005.05.005

Pinker, S. (1989). *Learnability and cognition: The acquisition of argument structure*. Cambridge, MA: The MIT Press.

Scott, R. M. & Fisher, C. (2009). Two-year-olds use distributional cues to interpret transitivity-alternating verbs. *Language and Cognitive Processes*, 24, 777–803. https://doi.org/10.1080/01690960802573236

Talmy, L. (1983). How language structures space. In H. L. Pick & L. P. Acredolo (Eds.), *Spatial orientation: Theory, research, and application* (pp. 225–282). New York, NY: Plenum Press. https://doi.org/10.1007/978-1-4615-9325-6_11

Talmy, L. (1985). Lexicalization patterns: Semantic structure in lexical forms. In T. Shopen (Ed.), *Language typology and syntactic description* (pp. 57–149). Cambridge: Cambridge University Press.

Tsunoda, T. (1985). Remarks on transitivity. *Journal of Linguistics*, 21, 385–396. https://doi.org/10.1017/S0022226700010318

Wagner, L. (2010). Inferring meaning from syntactic structures in acquisition: The case of transitivity and telicity. *Language and Cognitive Processes*, 25, 1354–1379. https://doi.org/10.1080/01690960903488375

Yuan, S., & Fisher, C. (2009). "Really? She blicked the baby?": Two-year-olds learn combinatorial facts about verbs by listening. *Psychological Science*, 20, 619–626. https://doi.org/10.1111/j.1467-9280.2009.02341.x

Yuan, S., Fisher, C., & Snedeker, J. (2012). Counting the nouns: Simple structural cues to verb meaning. *Child Development*, 83, 1382–1399. https://doi.org/10.1111/j.1467-8624.2012.01783.x

CHAPTER 8

Analogical structure mapping and the formation of abstract constructions

A novel construction learning study

Ben Ambridge[i,iv], Micah B. Goldwater[ii] and Elena V. M. Lieven[iii,iv]
[i]University of Liverpool / [ii]University of Sydney / [iii]University of Manchester / [iv]ESRC International Centre for Language and Communicative Development (LuCiD)

This chapter reports an experimental study in which children aged 4;9–6;1 were taught two novel constructions with meanings of *enabling/helping* and *preventing/stopping* with OSV and VOS word order. The aim was to test Tomasello's (2003) proposal that children form abstract constructions by performing analogical structure mapping across lexically-specific slot-and-frame patterns (e.g., *He's [X]ing it*) and/or particular sentences that instantiate them. All surface cues were minimized, such that successful learning of the construction required children to align the relational structure (i.e., *helper-action-helpee* or *preventer-action-preventee*, as appropriate) of the training sentences. A forced-choice pointing test revealed that children did not successfully learn either the constructions' global meanings or argument linking patterns (though a control group of adults succeeded at the latter task).

Keywords: syntax, construction learning, constructivist approach, structure mapping, analogy, relational structure

1. Introduction

Traditionally, accounts of syntax acquisition have assumed an important innate – and specifically linguistic – component. For example, children have been argued to acquire word-order (e.g., SUBJECT VERB OBJECT for English) by setting innate parameters (e.g., SV/VS and VO/OV; e.g., Sakas & Fodor, 2012). Partly in response to the difficulties encountered by such an approach (e.g., Boeckx & Leivada, 2013), many researchers have proposed alternatives, under which children acquire syntax gradually on the basis of the input (e.g., Clark, 1976; Peters, 1983; Pine & Lieven, 1993; Tomasello, 2003; MacWhinney, 2014; Ambridge & Lieven, 2015).

https://doi.org/10.1075/tilar.24.08amb

One of the most influential of these so-called constructivist accounts, the usage-based theory (see e.g. Tomasello, 2003), posits a three-stage process. Children start out by acquiring rote-learned, unanalyzed holophrases (e.g., *He's+kicking+it; He's+eating+it; He's+drinking+it*) from the input. Next, as soon as (in principle) two suitable holophrases have been acquired, children begin gradually to schematize across them to form partially-productive, semi-abstract *schemas* or *slot-and-frame patterns* (e.g., *He's ACTIONing it*). Finally, children analogize across these patterns (and/or concrete sentences that instantiate them) to arrive at fully-productive adult-like abstract constructions that reflect both semantic generalizations (e.g., *[AGENT] [ACTION] [PATIENT])*[1] and how the semantics is syntactically realized (e.g., *[SUBJECT-NOUN] [VERB] [OBJECT-NOUN]*).

On our reading of the literature, the existence of (a) rote learned holophrases and (b) partially-productive slot-and-frame patterns is well established. With regard to the former, four recent studies have demonstrated that young children's utterances are more fluent and/or accurate when they are able to make use of a multi-word string that is of high frequency in the input (and hence potentially stored as a holophrase), even when controlling for the frequency of the relevant words and bigram sequences (Bannard & Matthews, 2008; Matthews & Bannard, 2010; Arnon & Snider, 2010; Arnon & Clark, 2011). With regard to the latter, many studies (see Ambridge & Lieven, 2015: 491–492 for a summary) have found that children are able to correctly produce particular sentence types only when they can potentially use a well-learned template such as *He's ACTIONing it* (e.g., Dodson & Tomasello, 1998; Akhtar, 1999; Childers & Tomasello, 2001), *It got ACTIONed by it* (e.g., Savage, Lieven, Theakston, & Tomasello, 2003) or *What's he ACTIONing*? (e.g., Ambridge, Rowland, Theakston, & Tomasello, 2006; Rowland, 2007; Ambridge & Rowland, 2009).

The situation is very different, however, for the third stage in the constructivist account of acquisition. In fact, we are not aware of a single empirical study which directly tests the idea that children acquire abstract constructions by analogizing across slot-and-frame patterns, or concrete sentences instantiating them (though see Goldberg, Casenhiser, & White, 2007; Onnis, Waterfall, & Edelman, 2008 for

1. To be clear, we are not committing to any particular formal account of the transitive construction semantics. Given the vast variety of meanings that the transitive construction can express, there have been many proposals that have either tried to find a single highly abstract generalized meaning such as *[MORE PROMINENT EVENT PARTICIPANT] [EVENT] [LESS PROMINENT EVENT PARTICIPANT]*, propose a cluster of several distinct more concrete meanings, or even propose that the transitive is so abstract it has no meaning, unlike other more semantically constrained constructions such as the resultative (see Jackendoff, 2002, among others for thorough discussion). For now, we will simply assume that construction representations reflect the semantic and syntactic generalizations of their exemplar utterances.

related studies that use some degree of surface overlap). A version of this claim is presented by Tomasello (2003: 164) who bases the claim on the research of Gentner and colleagues, which show that children and adults are capable of *structure mapping* (albeit they do not show that children learn via structure mapping; Gentner & Markman, 1995, 1997; Gentner & Medina, 1998; Kotovsky & Gentner, 1996). For example, in the study published in Richland, Morrison and Holyoak (2006), children were shown a picture of a cat chasing a mouse above a picture of a boy chasing a girl, and were asked "what is like the cat in the bottom picture?". Richland et al. found that 3 to 4-year-old children would reliably pick the boy (matching the agent), and 6 to 7-year-old children would reliably do so even when an identical cat was present in the bottom picture. That is, rather than choosing the best literal match – the identical cat – children align the relational structure of the two pictures (i.e., *chaser-chasee*) and hence perceive a deep structural similarity between the item playing the same role (*chaser*) in both. Even for this younger age group, Christie and Gentner (2010) showed that 3 to 4-year-old children, when given some scaffolding to compare across examples, can see past the superficial object differences and generalize a novel relational structure to a novel pair of objects.

Based on these findings, and similar findings into the domain of language acquisition Tomasello (2003: 166) suggests "it is thus possible that abstract linguistic constructions are created by a structural alignment across different item-based constructions, or the utterances emanating from them", offering the following "somewhat whimsical linguistic example" (p. 165).

Consider the following two sequences of letters:

> U R X
> I M A B
> Let us try to make an analogy. It is not easy because there is not much in common between the two strings; they even have different numbers of items. But what if we now translate them into English in a fanciful context. The situation is that we are role-playing with a child in a pretend game, and one role is to be a creature named X and another role is to be one of several honeybees. The above sequences now translate into "You are X" and "I am a bee". We can now see that these are both predicate nominative constructions, and they are analogous: *you* corresponds to *I* as the one to be anointed, *are* corresponds to *am* as the identifying relation, and *X* corresponds to the two-word phrase *a bee* as the new identity taken on. Such correspondences can only be made once we know the functions of the items and structures involved. (Tomasello, 2003: 165)

Note that, exactly as with Richland and colleagues' *chaser-chasee* example, the pairs that are to be analogized across need not share any surface similarity. Rather, what is crucial is that they share *relational similarity*; a similar relationship (chaser-chasee or person-identity taken on) holds between the members of each pair.

Ambridge and Lieven (2015) discuss how such an account could potentially explain the acquisition of basic transitive word order. Suppose, for example, that a child has formed the slot-and-frame patterns schemas *I'm [ACTION]ing it* and *[KISSER] kissed [KISEE]*. Although these patterns share no surface similarity, they do share some kind of AGENT-ACTION relation (*I-[ACTION]* and *[KISSER]-kissed*) and ACTION-PATIENT relation (*[ACTION]it* and *kissed [KISSEE]*). Tomasello's (2003) claim is that this type of relational overlap (presumably between many such schemas, rather than just two) is sufficient for children to analogize across them, and hence move towards a fully-abstract [AGENT] [ACTION] [PATIENT] sentence-level argument-structure construction.

The aim of the present study is to begin to test this claim. Although several of Gentner and colleagues' studies have demonstrated analogical structure mapping in a nonlinguistic context, we are unware of any previous experimental attempts to test Tomasello's (2003) claim that this mechanism is operational in the domain of syntax acquisition (but see Goldwater et al., 2011 for computational simulations of such a process, and Goldwater, Tomlinson, Echols & Love, 2011 and Goldwater & Echols, *in prep* for evidence for structure mapping as a mechanism in children's syntactic production of familiar constructions).

Methodologically speaking, the starting point for the present study was the novel construction learning studies of Goldberg and colleagues (e.g., Boyd & Goldberg, 2012; Boyd, Gottschalk & Goldberg, 2009; Casenhiser & Goldberg, 2005; Wonnacott, Boyd, Thomson, & Goldberg, 2012), in which learners were taught a new *appearance* or *approaching* construction (e.g., *The bird the flower moopoed*). In this study, children watched short videos of events of appearance or approaching, (e.g., a bird appears on a flower), while hearing sentences describing the event using the novel word order. After a series of learning trials, children would be presented with test trials that presented two novel videos, one showing appearance, and one showing an event describable with a familiar construction, such as the transitive. On half of the test trials, children heard a sentence with a novel construction, and half with a familiar construction. The children on average pointed to the appearance events when the novel construction was heard more than when a familiar construction was heard, consistent with having generalized the construction semantics and word order beyond the initial training set.

Because these previous studies were not concerned with investigating analogical structure mapping *per se*, they did not attempt to prevent learners from using surface rather than relational similarity. Thus a particular sentence position was always filled by an *appearer* or an *approacher*, and hence shared across trials a type of surface-level similarity (i.e., it always appeared, or always moved). In some studies, instantiations of the construction also shared surface similarity in the form of lexical overlap of the definite articles and the novel construction marker *–o* (e.g., *The [THEME] the [LOCATION] [VERB]-o-s/ed*). Thus in order to acquire the construction, learners did not necessarily need to align the underlying *appearer-location* (or *approacher-location*) structure of the training trials (although, of course, they may well have done so). In contrast, the present study – in addition to eliminating lexical overlap – required learners to align across *preventer-action-preventee* events and *helper-action-helpee* events. This is because being a *preventer or helper* requires both an action to prevent/help and a *preventee/helpee*. In contrast, the *appearer* and *approacher* roles in these previous studies could be understood without reference to the *location* role.

It is worth noting, that to more strictly test the role of structural alignment specifically, we moved away from how many natural language constructions work, thus potentially making the task more difficult for the learner. Consider ditransitive sentences such as *John threw/kicked/gave/handed the ball to Bob*. They all convey a meaning of transfer of an object from one possessor to another. The verb expresses the means of the transfer. Transfer in these sentences is a "first-order" relation in that it takes objects as its arguments. On the other hand, a "second-order" relation takes first-order relations as its arguments. The *help* and *prevent* relations of the current artificial constructions are second-order in this way, because what is helped or prevented is a whole event unto itself with its own first-order relation, as in an example animated event from our study, "The king helped the queen climb the ladder." Here, the verb *climb* names the first-order relation that is the argument of the second-order *help* relation that is the meaning of the artificial construction. This is unlike the dative construction, wherein the verb fleshes out the specific means (e.g., *throwing*) of the first-order relation that the construction denotes in a more abstract way (*transfer*). Goldberg and colleague's *appearance* semantics is more akin to natural construction semantics in this way.

A second modification from the prior work is that participants in the present studies were taught two novel constructions, as opposed to only one in the studies of Goldberg and colleagues. Although this, of course, makes the task more difficult for learners, it ensures that they cannot succeed on construction-meaning test trials by using a process of elimination. In these previous studies, learners heard either the novel construction (e.g., *The frog the apple zoopos*) or – for example – a transitive construction (e.g., *The frog zoopos the apple*) and had to choose between, for example, the frog appearing on the apple or the frog pushing the apple. Thus, for these test trials, it is possible that learners could have succeeded by mapping the transitive construction to the transitive actions and – by process of elimination – the novel construction to the non-transitive action. In the present study, learners were given a forced choice between two novel constructions.

A third modification was that in place of the subject-first word order used in some of these previous studies (which is somewhat similar to English), we used two word orders designed to be as dissimilar as possible to English: OSV and VOS. A fourth and final modification involved the use of familiar verbs, as opposed to novel verbs in the studies of Goldberg and colleagues. Because, in the present study, the construction meaning (helping/preventing) was independent of the meaning of the verb (e.g., *leave, fall, hop, sit*) it was felt that the use of novel verbs would unnecessarily increase the burden on learners.

To summarise, the aim of the present study was to investigate Tomasello's (2003) claim that learners can acquire abstract syntactic constructions on the basis of underlying relational similarity between sentences that share no surface similarity. To this end, we taught children aged 4–6 two novel constructions: OSV and VOS. The different exemplars of each construction had no surface lexical material in common, but shared higher-order relational structure. For one construction, the SUBJECT prevented the OBJECT from performing the action denoted by the VERB (i.e., PREVENTER+ACTION and ACTION+PREVENTEE relations). For the other, the SUBJECT helped or helped the OBJECT to perform the action denoted by the VERB (i.e., HELPER+ACTION and ACTION+HELPEE). Subsequent forced-choice tests examined whether the participants had learned (a) the meaning of each of the two constructions (e.g., Naigles, 1990) and (b) the linking patterns of each (e.g., Gertner, Fisher & Eisengart, 2006).

Unlike an account that suggests children can learn construction meaning from higher-order relational structure alone, alternative accounts of construction generalization that rely on the distribution of lexical overlap across sentences (e.g., Twomey, Chang, & Ambridge, 2014) would not predict much learning in the task. Following the results, we will describe the implications for several alternative explanations in more depth.

2. Method

2.1 Participants

Participants were 24 children (13 male) aged 4;9–6;1 (*M* = 5;7; *SD* = 0;5), recruited from schools and nurseries in Greater Manchester (UK). The study was approved by the University of Manchester ethics committee, and parents gave written consent. This age range was picked to be consistent with work reviewed above that showed either successful artificial construction learning (by Goldberg and colleagues) or structure-mapping in language production (Goldwater et al., 2011).

2.2 Materials and procedure

2.2.1 *Training*

Each child was taught two novel constructions (with English words); one with OSV word order, the other with VOS word order (chosen to be as dissimilar as possible to English SVO order). One construction was with paired with the meaning of *helping/enabling*, the other with *stopping/preventing* (counterbalanced across children). Thus, according to the structure mapping theory, children should learn each construction by aligning, as appropriate, the *helper-action-helplee* or *preventer-action-preventee* structure of the training sentences. In order to ensure that children had to rely on this deep structural cue, as opposed to surface similarity, we took care to ensure that the precise means of helping/enabling or stopping/preventing, as illustrated in the accompanying animations, varied in every case (see Table 1). Thus the stopping/helping could be physical, by giving an order, removing/supplying some necessary object etc. Although all four roles (*enabler, enablee, preventer, preventee*) were always animate, we ensured that each was filled by both humans and animals, and that none had any consistent direction of motion between different animations. The animations, which used characters chosen to be familiar to most British children, were created in Anime Studio Pro and displayed using Processing (Reas & Fry, 2007). Each animation was displayed in the centre of the screen, accompanied by an animated robot head, below, who spoke the sentence. In order to minimize confusion, blocked presentation was used, such that children saw either 12 stopping/preventing videos then 12 helping/enabling videos or vice versa, as shown in Table 1. This entire training set was shown once per day for 3 consecutive days, for a total of 36 training trials per construction.

Table 1. Training sentences and animations

Stopping/preventing training sentences		
OSV (CB Group A)	**VOS (CB Group B)**	**Animation**
Homer Marge leave	leave Homer Marge	Homer tries to exit screen left, but Marge drops a brick wall in his path
Lisa Bart come	come Lisa Bart	Lisa tries to enter a building, but Bart (already inside) closes the door
Wendy Bob run	run Wendy Bob	Wendy runs towards Bob (The Builder), who stops her by giving the "halt" signal
Grandpa Duck hop	hop Grandpa Duck	Grandpa hops towards the duck, who kicks him over
Grandma Sheep laugh	laugh Grandma Sheep	Grandma is laughing, but stops when sheep "Ssh"s her

(*continued*)

Table 1. (*continued*)

Stopping/preventing training sentences		
OSV (CB Group A)	**VOS (CB Group B)**	**Animation**
Dora Frog talk	talk Dora Frog	Dora says "Hi, I'm Dora the Ex…", but frog sticks a band aid over her mouth
Wolf Dennis sit	sit Wolf Dennis	The wolf walks towards the chair but Dennis (The Menace) pulls it off screen right
Lion Tracy drink	drink Lion Tracy	The lion reaches for a cup of water, but Tracy (Beaker) pours the water onto the floor
Dog Ash eat	eat Dog Ash	Ash (a Pokemon character) reaches for a burger, but the dog pushes it off the table and into the bin
Tiger Fox fall	fall Tiger Fox	The tiger runs towards the edge of a cliff, but the fox runs after him and drags him back
Bunny Elephant cry	cry Bunny Elephant	The bunny is crying, but the elephant dabs his tears, and he stops
Cat Pig sing	sing Cat Pig	The cat sings, but Piglet (from Winne The Poo)kicks the hi-fi system off the screen and the cat stops

Helping/enabling training sentences		
VOS (CB Group A)	**OSV (CB Group B)**	**Animation**
leave Bart Wendy	Bart Wendy leave	Bart walks right, Wendy kicks down a wall that blocks his path, and he exits screen right
come Bob Homer	Bob Homer come	Homer (already inside) opens the door to a building, and Bob enters
run Lisa Marge	Lisa Marge run	Marge runs behind Lisa, pushing her along, and causing her to run
hop Dora Sheep	Dora Sheep hop	The sheep (walking on two legs) holds Dora's hand as she hops along
laugh Grandpa Frog	Grandpa Frog laugh	The frog tickles Grandpa, and he laughs
talk Grandma Duck	Grandma Duck talk	The dog removes a band aid from over Grandma's mouth, and she begins to speak
sit Dog Tracy	Dog Tracy sit	Tracy drags in a chair, and the dog sits down
drink Wolf Ash	Wolf Ash drink	Ash pours some water into a cup and the wolf drinks it
eat Lion Dennis	Lion Dennis eat	Dennis brings some fruit to the lion, and he eats it
fall Cat Elephant	Cat Elephant fall	The cat strolls towards the edge of a cliff and the elephant runs up and pushes her off
cry Pig Tiger	Pig Tiger cry	The tiger hits piglet in the face and he cries
sing Bunny Fox	Bunny Fox sing	The fox presses "play" on the hi-fi system, and the bunny starts to sing

2.2.2 *Test*

Immediately after the final training session, children completed (in random order) eight forced-choice pointing test trials (in one of four different versions counterbalanced for the side of the target, and for which particular character filled each role). Four trials (modeled on Naigles, 1990) assessed children's learning of the meanings of the constructions (two trials each with the helping and stopping constructions as target, counterbalanced for side). These trials showed one helping video, and one preventing video, each with the same characters playing the corresponding roles (e.g., with Homer as preventer/helper of Marge). The goal was to point to the event that matched the construction meaning. Four trials (modeled on Gertner, Fisher & Eisengart, 2006) assessed their learning of the constructions' semantics-syntax linking patterns (again, two trials each with the helping and stopping constructions as target, counterbalanced for side). These trials presented two events of the same type (either two preventing, or two helping), but which character played which role differed across the two videos (e.g., Homer was the helper and Marge was the helpee, or vice versa). The goal was to point to the event that matched the word order rule of the construction. Table 2 shows the test trials for children in the first counterbalance condition. Note that the test trials did not reuse any of the same characters or verbs as the training trials. For each trial, children saw a preview of, first, the left-hand video and, second, the right-hand video, before the two played simultaneously, accompanied by the audio sentence. The videos ended on a freeze frame that made it clear which was which (e.g., the man with his foot on the chicken vs the chicken running free) in order to give children as long as they needed to process the sentence and make their choice.

Table 2. Test sentences and animations (correct choice shaded grey)

Test Type	Sentence	Animation (Left screen)	Animation (Right screen)	Meaning (Group A)	Target screen (Group A)
Construction meaning	Chicken man play	The chicken runs around squawking, but the man puts his foot on it and stops it	The man picks up the squawking chicken and dances around with it	Stop/ Prevent	L
	Sleep panda woman	The panda is asleep in bed, but the woman shouts "wake up", and he does	The woman sings a lullaby to the panda, who then falls asleep, "zzzz"	Help/ Enable	R

(*continued*)

Table 1. (*continued*)

Test Type	Sentence	Animation (Left screen)	Animation (Right screen)	Meaning (Group A)	Target screen (Group A)
	Fish prince swim	The prince throws the fish into a bowl, and he starts swimming	The fish is swimming in a bowl, but the prince kicks it over, spilling the water and fish	Stop/ Prevent	R
	Jump horse Princess	The princess rides a horse over a jump	The horse approaches a jump, but the princess – on the other side – raises her hands, "stop"	Help/ Enable	L
Semantics-syntax linking	Boy girl draw	The boy is drawing on a piece of paper, but the girl takes it away	The girl is drawing on a piece of paper, but the boy takes it away	Stop/ Prevent	L
	Rabbit bear fly	The bear is flying around in a plane, but the (giant) rabbit punches it out of the sky	The rabbit is flying around in a plane, but the (giant) bear punches it out of the sky	Stop/ Prevent	R
	Climb king queen	The king, atop the castle, throws down a ladder, and the queen climbs it	The queen, atop the castle, throws down a ladder, and the king climbs it	Help/ Enable	R
	Dance cow zebra	The zebra plays the guitar while the cow dances	The cow plays the guitar while the zebra dances	Help/ Enable	L

3. Results

All children made an unambiguous point on every trial, with no missing data. The data were analyzed using mixed effects models (using the *lme4* package for *R* (R Core Team, 2014). Binomial models were used to compare the proportion of correct (1) vs incorrect (0) choices to chance (i.e., by comparing the intercept to 0.5). All models included random intercepts for participant. Random intercepts for trial were included in the *semantics-syntax linking* analysis, but caused convergence-failure in the *construction-meaning* analysis (random slopes were not possible, since the models contained no predictors, only the intercept).

For the *construction-meaning* trials ($N = 4$), children made a mean of 57.29% correct points, which did not differ significantly from chance (*intercept* = 0.29, $SE = 0.21$, $z = 1.42$, $p = 0.16$). One factor in children's poor performance may be a small (though not significant) bias towards the helping/enabling action (perhaps a pro-social effect); a supplementary analysis found that children chose the helping/enabling action on 68% of trials (*intercept* = 1.74, $SE = 0.09$, $Z = 1.88$, $p = 0.06$, n.s.). Note that, because they are making a forced-choice between these two alternative meanings, it is not meaningful to compare children's performance on the stopping and helping target trials (children appear to show better performance for the latter, but this is only because of their overall bias towards the helping videos).

For the *semantics-syntax* linking trials ($N = 4$) children made a mean of 58.33% correct points, which again did not differ significantly from chance (*intercept* = 0.34, $SE = 0.21$, $z = 1.625$, $p = 0.10$). Children's performance on the linking trials did not differ significantly between the stopping trials (56.25% correct) and helping trials (60.42% correct); adding trial type as predictor did not improve model coverage ($Chi_1 = 0.17$, $p = 0.68$, n.s.).

3.1 Manipulation check

Given the failure of the children to learn the constructions, it seemed important to verify that the constructions were, in principle, learnable; i.e., that we succeeded in our aim of teaching stopping/preventing and helping/enabling constructions with identifiable and distinct links between semantic and syntactic roles. We therefore repeated the study with eight adult volunteers at Northwestern University. For the *semantics-syntax* linking trials, adults made a mean of 91% correct points, which easily beat chance (*intercept* = 2.27, $SE = 0.61$, $p < 0.001$). Although they did not significantly beat chance on the *construction-meaning* trials (81% correct points, *intercept* = 5.38, $SE = 6.85$, $z = 0.78$, $p = 0.43$, n.s.), this probably reflects nothing more than a lack of statistical power, as their performance on the linking trials demonstrates that these adults clearly learned the semantics of the two constructions. We also included two "fill in the blank" questions: "How did the language work?" and "Did you use a strategy (if so, what)?". In their answers, all eight adults mentioned either stopping/preventing versus helping/enabling or some clearly related alternative formulation that mapped on to the manipulation (e.g., action complete/incomplete; action positive/negative). Thus, while the constructions are, in principle, learnable, we cannot rule out the possibility that they are learnable only by means of an explicit strategy.

4. Discussion

The aim of the present study was to investigate Tomasello's (2003) claim that learners can acquire abstract syntactic constructions on the basis of underlying relational similarity between sentences that share no surface similarity. Children aged 4–6 were taught two novel constructions, OSV and VOS, with the different instantiations of each exhibiting no surface similarity. For one construction (OSV for half of the children; VOS for the remainder), the SUBJECT prevented the OBJECT from performing the action denoted by the VERB (i.e., PREVENTER+ACTION and ACTION+PREVENTEE relations). For the other, the SUBJECT helped the OBJECT to perform the action denoted by the VERB (i.e., HELPER+ACTION and ACTION+HELPEE relations). Subsequent forced-choice tests found no evidence that children had learned either the global meaning of each construction (e.g., preventing/helping) or the particular linking patterns of each.

It is never easy to interpret (or publish) a null result because there is no straightforward way to choose between two competing possibilities: (a) the null result is real, meaning in this case that children cannot in fact learn syntactic constructions by performing structure mapping or (b) the null result is a consequence of some flaw in the design of the study, and the hypothesized effect – in this case children's ability to learn syntax via structure mapping – is real (and would be detected by a better-designed study).

Let us take the least interesting possibility first. Although previous studies (and – to some extent – our adult manipulation check) demonstrate that something like the present paradigm is, in principle, capable of demonstrating construction learning, it may not be optimally suited to doing so. Perhaps, for example, the study was underpowered, in that we simply did not have enough participants to detect a relatively small effect. For example, Wonnacott et al.'s (2012) study with 5-year-olds used 42 and 35 participants in Experiments 1 and 2 respectively, while ours used just 24 (and is potentially harder, in that children had to learn two constructions). Casenhiser and Goldberg's (2005) study with 5–7 year olds tested 51 and

48 children in Experiments 1 and 2 respectively (and, again, is probably somewhat easier than ours). Although Boyd and Goldberg (2012) tested just 18 children in each of two age groups (5- and 7-year olds), they also failed to find an effect for the younger group (comparable in age to the participants of the present study). Perhaps, then, the present paradigm would also reveal evidence of learning if a larger sample were tested (we would be delighted to share our stimuli with any colleagues who would like to investigate this possibility).

It may also be the case that we did not include enough training trials (we used 36 per construction). This is somewhere in between the regimes of Wonnacott et al. (2015) (48 trials) and Casenhiser and Goldberg (2005) and Boyd and Goldberg (2012) (16 trials in each case), though it is again important to remember that children in these studies were trained on only a single construction, and that the younger group in the latter study failed to learn it. However, it is also possible that there were too *many* trials. There is evidence from research on structure mapping in children's category learning that more examples may have some negative effects on generalization (by overloading children), and that as few as six can be optimal for discovering abstract structural commonalities (Thibaut & Witt, 2015).

Another potential problem is that, like all construction-learning studies, the experimental scenario was too artificial to tap into children's underlying abilities at analogical reasoning. Dunbar (2001: 313) discusses an *analogical paradox*, whereby "analogy is so easy in naturalistic settings, yet so difficult in the psychological laboratory". Dunbar's solution to the paradox is that the apparently irrelevant contextual cues that are stripped out in laboratory studies are actually crucial in allowing learners both to notice and encode structural regularities in the first place and to retrieve them in suitable situations. Perhaps, then, a more naturalistic version of our study, conducted in familiar surroundings with familiar, real-world protagonists, would be more likely to yield evidence of analogical structure mapping in syntax acquisition.

As discussed from the beginning, the final potential problem is that the types of relations to be learned in this study, higher-order relations, were more difficult than the first-order relations that are typical of basic argument structure constructions. It is worth noting that even adults, in a similar experimental context to the current one, sometimes do not spontaneously see second-order relational connections across events when the first-order relations are distinct (Goldwater, Bainbridge, & Murphy, 2016). (The editors of this volume raised one more potential design flaw: "for some animations, the goal must be inferred, and for others this is not the case. Consider the wolf heading to the chair; one must infer the intention to sit. In other cases, the action is in motion (e.g., the cat singing) when it is halted".)

Indeed, there is some evidence from preferential-looking/pointing studies that at least one particular first-order relation, AGENT of an action, and its tendency to appear clause-initially in English, is acquired by 2;0, or even younger (Arunachalam

& Waxman, 2010; Bavin & Growcott, 2000; Gertner, Fisher & Eisengart, 2006; Gertner & Fisher, 2012; Kidd, Bavin, & Rhodes, 2001; Naigles & Kako, 1993; Noble, Rowland, & Pine, 2011; Noble, Iqbal, Lieven, & Theakston, 2016; Pozzan, Gleitman & Trueswell, 2016; Yuan & Fisher, 2009). Thus it may be that even young infants would show analogical structure mapping for a novel syntactic construction with simpler relations. On the other hand, there is some evidence that children of this age struggle to process sentences with inanimate agents. Chan, Lieven and Tomasello (2009) found that, when word-order and animacy cues were placed in conflict (e.g., *The ball tams the chicken*), Cantonese-, German-, and English-speaking children aged 2;6 did not show correct comprehension (though they did so by 3;6). This suggests that two-year-olds may have not yet acquired a true ***relational*** category – AGENT ***of an action*** – but rather a fuzzy, more probabilistic notion of the types of entities that tend to be mentioned first in the clause (human, animate, moves first, moves more etc.).

This brings us to the second of the possibilities set out above; that the kind of relational structure mapping procedure set out by Tomasello (2003) is not in fact sufficient for the acquisition of abstract constructions. While it would, of course, be premature to accept such a conclusion without first running several more studies of this type, we suggest that this possibility is one that is worth taking seriously. Why? There are (at least) two kinds of alternative processes that can help children transition from lexically-based to fully abstract linguistic constructions. The first kind of alternative process is a form of structural alignment-based learning rooted in the need for overlapping surface similarity for a prolonged period of time. For example, in the real world, very young children hear and use a relatively narrow range of AGENTS and PATIENTS (e.g., *You, Me, Mummy, Daddy*), and it may be that hearing exactly the same word (e.g., *Mummy*) used in a variety of different types of AGENT roles somehow anchors this category. This seems quite likely, given evidence that even adults struggle to learn arbitrary linguistic categories that lack surface similarity (e.g., Brooks, Braine, Catalano, Brody, & Sudhalter, 1993). Indeed, direct tests of children's analogical reasoning ability, demonstrate that manipulations of cognitive load have large effects on their ability to solve analogies (e.g., Richland et al., 2006; Thibaut, French, & Vezneva, 2010).

Further, there is direct evidence for this more graded form of analogical learning. The findings of Chan et al. (2009) suggest that children may start out with non-relational categories that are formed largely on the basis of surface features (e.g., humanness and animacy for an early AGENT category). The surface similarities captured by this early category would then aid children's acquisition of a truly relational AGENT (of an ACTION) category. There is strong evidence that children can more easily discover relational commonalities when the relational commonalities are correlated with superficial similarities, which then allows children to recognize when this common structure is present in more disparate exemplars (e.g., Gentner & Loewenstein, 2002).

Additional support for this more graded analogical learning (as discussed in Ambridge & Lieven, 2015) starts from the observation that children do not necessarily have to analogize between entirely dissimilar instances (e.g., very different types of AGENT-ACTION relations) in a single leap. Rather, the process of analogy could proceed via a series of baby steps. For example, the KICKER-KICKEE relation is very similar to the HITTER-HITTEE relation, which in turn is similar to the TOUCHER-TOUCHEE relation, the KISSER-KISSEE relation, and so on, right up to AGENT-ACTION relations that are very different to the KICKER-KICKEE relation with which we began. This is similar to the pattern of how 4-year-olds categorized arrays of objects in Kotovsky and Gentner (1996). The children could see the connection between a set of objects monotonically increasing in *size* from left to right with a set of objects monotonically increasing in *luminance* from left to right (showing evidence of understanding a general "monotonically increasing" relation) only after they had already made the connection between two sets of objects, each monotonically increasing along the same dimension (either size or luminance). There is some evidence of this *progressive alignment* in artificial construction learning (Goldberg, Casenhiser, & White, 2007), and in children's production of familiar constructions (Goldwater & Echols, in prep).

A third reason why a pure structure abstraction process is never truly engaged (also discussed in Ambridge & Lieven, 2015) is that children might not be learning purely abstract constructions, such as a single AGENT-ACTION-PATIENT (or SUBJECT-VERB-OBJECT) construction at all. Perhaps speakers have a family of distinct constructions that share similar word order and case marking, but need never analogize across them. For example, English speaking adults may have at least six different transitive constructions.

Contact (non-causative)	[AGENT] [ACTION] [PATIENT]	John hit Bill
Causative	[CAUSER] [ACTION] [CHANGE]	John broke the plate
Experiencer-Theme	[EXPERIENCER] [EXPERIENCE] [THEME]	John heard Bill
Theme-Experiencer	[THEME] [EXPERIENCE] [EXPERIENCER]	John scared Bill
"Weigh" Construction	[THING] [MEASURE/COST/WEIGH] [AMOUNT]	John weighed 100lbs
"Contain" Construction	[CONTAINER] [CONTAIN] [CONTENTS]	The tent sleeps four people

Given that usage-based accounts such as that of Tomasello (2003) posit rampant redundancy (perhaps even the storage of every exemplar), there is certainly no *a priori* reason to reject the notion of multiple transitive constructions.

Aligned with these first three possible learning procedures, computational models of analogical learning simulate the processes of progressive alignment and the learning of related, but differentiated, relational structures in a unified framework, similar to the variants of the transitive (Doumas, Hummel, & Sandhofer, 2008; Goldwater et al., 2011; Kuehne, Forbus, Gentner, & Quinn, 2000). Future studies using some of the modifications suggested above should explore these possibilities.

Last, we suggest that the second kind of alternative explanation for our results is the real possibility that children do not learn constructions via structural alignment, either purely abstractly or even with these more graded forms that may never produce pure abstraction. More specifically, this final proposal is highly consistent with Tomasello's (2003) overall approach, as it emphasizes gradual generalization at a lexical level. Recall that Tomasello (2003) invoked structure mapping to solve the problem of how learners abstract across two instantiations of a construction that share no surface lexical material (e.g., *I ate the cake*; *Sue kissed Bill*). Indeed, when looking at just a single sentence pair, this problem does seem to require some very high-level, abstract solution, such as analogical structure mapping. But in fact, structurally related sentences that are, on the surface, very different, can generally be linked by handful of sentences that differ by a single constituent:

I ate the cake
I like the cake
I like Bill
I kissed Bill
Sue kissed Bill

So if, as seems likely, children are abstracting across a large number of stored sentences, rather than just in a pairwise fashion, this abstraction could in principle proceed largely on the basis of surface lexical similarities. Indeed, when processing an incoming sentence, children likely activate all somewhat-related exemplars stored in memory (we assume here that abstraction does not happen offline, but as part of real time language processing and productions). Similarly, children may be able to link slot-and-frame patterns on the basis that a single lexical item (here, *kick*) serves as a slot-filler in one pattern and (part of) the frame in another:

I'm [X]ing it [X = *kick, hit, eat, push*] + Kick [Y] [Y = *the ball, the cat, the man*] → I'm [X]ing [Y]

As we have already stressed, lower-level abstraction on the basis of surface similarities – both these lexical similarities and semantic similarities such as animacy, humanness etc. – need not necessarily replace structure mapping. It might. But equally, we have discussed evidence that lower-level similarities can facilitate more abstract analogical reasoning. Figuring out the relative contributions of lower and higher level abstraction mechanisms is an important goal for future research, and a question that is probably best tackled using computational modeling (see Chang, Dell and Bock, 2006; Twomey, Chang and Ambridge, 2014, for models that acquire something like abstract syntactic constructions on the basis of lexical-distributional overlap).

For now, the present study has provided no evidence for the claim that children acquire abstract syntactic constructions using analogical structure mapping (though only the most preliminary evidence against this claim). However, our endeavours will not have been wasted if we have succeeded in persuading researchers that this hypothesis is one that can – indeed, must – be investigated empirically, perhaps using some improved version of the paradigm that we have established here.

Acknowledgement

Ben Ambridge and Elena Lieven are Professors in the ESRC International Centre for Language and Communicative Development (LuCiD) at the University of Liverpool and the University of Manchester respectively. The support of the Economic and Social Research Council [ES/L008955/1] is gratefully acknowledged.

References

Akhtar, N. (1999). Acquiring basic word order: Evidence for data-driven learning of syntactic structure. *Journal of Child Language*, 26, 339–356. https://doi.org/10.1017/S030500099900375X

Ambridge, B., & Lieven, E. V. M. (2015). A Constructivist account of child language acquisition. In B. MacWhinney & W. O'Grady (Eds.), *Handbook of language emergence* (pp. 478–510). Hoboken, NJ: Wiley Blackwell.

Ambridge, B., & Rowland, C. F. (2009). Predicting children's errors with negative questions: Testing a schema-combination account. *Cognitive Linguistics*, 20, 225–266. https://doi.org/10.1515/COGL.2009.014

Ambridge, B., Rowland, C. F., Theakston, A., & Tomasello, M. (2006). Comparing different accounts of non-inversion errors in children's non-subject wh-questions: 'What experimental data can tell us?' *Journal of Child Language*, 30, 519–557. https://doi.org/10.1017/S0305000906007513

Arnon, I., & Clark, E. V. (2011). When 'on your feet' is better than 'feet': Children's word production is facilitated in familiar sentence-frames. *Language Learning and Development*, 7, 107–129. https://doi.org/10.1080/15475441.2010.505489

Arnon, I., & Snider, N. (2010). More than words: Frequency effects for multi-word phrases. *Journal of Memory and Language*, 62, 67–82. https://doi.org/10.1016/j.jml.2009.09.005

Arunachalam, S., & Waxman, S. R. (2010). Meaning from syntax: Evidence from 2-year-olds. *Cognition*, 114, 442–446. https://doi.org/10.1016/j.cognition.2009.10.015

Bannard, C., & Matthews, D. (2008). Stored word sequences in language learning: The effect of familiarity on children's repetition of four-word combinations. *Psychological Science*, 19, 241–248. https://doi.org/10.1111/j.1467-9280.2008.02075.x

Bavin, E. L., & Growcott, C. (2000). Infants of 24–30 months understand verb frames. In M. Perkins & S. Howard (Eds.), *New directions in language development and disorders* (pp. 169–177). Dordrecht: Kluwer. https://doi.org/10.1007/978-1-4615-4157-8_16

Boeckx, C., & Leivada, E. (2013). Entangled parametric hierarchies: Problems for an overspecified Universal Grammar. *PloS one*, 8, e72357. https://doi.org/10.1371/journal.pone.0072357

Boyd, J. K., & Goldberg, A. E. (2012). Young children fail to fully generalize a novel argument structure construction when exposed to the same input as older learners. *Journal of Child language*, 39, 457–481. https://doi.org/10.1017/S0305000911000079

Boyd, J. K., Gottschalk, E. A., & Goldberg, A. E. (2009). Linking rule acquisition in novel phrasal constructions. *Language Learning*, 59, 64–89. https://doi.org/10.1111/j.1467-9922.2009.00536.x

Brooks, P. J., Braine, M. D. S., Catalano, L., Brody, R. E., & Sudhalter V. (1993) Acquisition of gender-like noun subclasses in an artificial language: The Contribution of phonological markers to learning. *Journal of Memory and Language*, 32, 76–95. https://doi.org/10.1006/jmla.1993.1005

Casenhiser, D., & Goldberg, A. E. (2005). Fast mapping between a phrasal form and meaning. *Developmental Science*, 8, 500–508. https://doi.org/10.1111/j.1467-7687.2005.00441.x

Chan, A., Lieven, E., & Tomasello, M. (2009). Children's understanding of the agent-patient relations in the transitive construction: Cross-linguistic comparisons between Cantonese, German, and English. *Cognitive Linguistics*, 20, 267–300. https://doi.org/10.1515/COGL.2009.015

Chang, F., Dell, G. S., & Bock, K. (2006). Becoming syntactic. *Psychological Review*, 113, 234–272. https://doi.org/10.1037/0033-295X.113.2.234

Childers, J. B., & Tomasello, M. (2001). The role of pronouns in young children's acquisition of the English transitive construction. *Developmental Psychology*, 37, 739–748. https://doi.org/10.1037/0012-1649.37.6.739

Christie, S., & Gentner, D. (2010). Where hypotheses come from: Learning new relations by structural alignment. *Journal of Cognition and Development*, 11, 356–373. https://doi.org/10.1080/15248371003700015

Clark, R. (1976). Report on methods of longitudinal data-collection. *Journal of Child Language*, 3, 457–459. https://doi.org/10.1017/S0305000900007352

Dodson, K., & Tomasello, M. (1998). Acquiring the transitive construction in English: The role of animacy and pronouns. *Journal of Child Language*, 25, 555–574. https://doi.org/10.1017/S0305000998003535

Doumas, L. A., Hummel, J. E., & Sandhofer, C. M. (2008). A theory of the discovery and predication of relational concepts. *Psychological Review*, 115, 1–43. https://doi.org/10.1037/0033-295X.115.1.1

Dunbar, K. (2001). The analogical paradox: Why analogy is so easy in naturalistic settings yet so difficult in the psychological laboratory. In D. Gentner, K. Holyoak, & B. N. Kokinov (Eds.), *The analogical mind: Perspectives from cognitive science* (pp. 313–334). Cambridge, MA: Bradford.

Gentner, D., & Loewenstein, J. (2002). Relational language and relational thought. In E. Amsel & J. P. Byrnes (Eds.), *Language, literacy, and cognitive development: The development and consequences of symbolic communication* (pp. 87–120). Mahwah, NJ: Lawrence Erlbaum Associates.

Gentner, D., & Markman, A. B. (1995). Similarity is like analogy: Structural alignment in comparison. In C. Cacciari (Ed.), *Similarity in language, thought and perception* (pp. 111–147). Brussels: Brepols.

Gentner, D., & Markman, A. B. (1997). Structure mapping in analogy and similarity. *American Psychologist*, 52, 45–56. https://doi.org/10.1037/0003-066X.52.1.45
Gentner, D., & Medina, J. (1998). Similarity and the development of rules. *Cognition*, 65, 263–297. https://doi.org/10.1016/S0010-0277(98)00002-X
Gertner, Y., & Fisher, C. (2012). Predicted errors in children's early sentence comprehension. *Cognition*, 124, 85–94. https://doi.org/10.1016/j.cognition.2012.03.010
Gertner, Y., Fisher, C., & Eisengart, J. (2006). Learning words and rules: Abstract knowledge of word order in early sentence comprehension. *Psychological Science*, 17, 684–691. https://doi.org/10.1111/j.1467-9280.2006.01767.x
Goldberg, A. E., Casenhiser, D., & White, T. R. (2007). Constructions as categories of language. *New Ideas in Psychology*, 25, 70–86. https://doi.org/10.1016/j.newideapsych.2007.02.004
Goldwater, M. B., Bainbridge, R. & Murphy, G (2016). Learning role-governed and thematic categories. *Acta Psychologica*, 164, 112–126. https://doi.org/10.1016/j.actpsy.2015.10.011
Goldwater, M. B., Tomlinson, M. T., Echols, C. H., & Love, B. C. (2011). Structural priming as structure-Mapping: Children use analogies from previous utterances to guide sentence production. *Cognitive Science*, 35, 156–170. https://doi.org/10.1111/j.1551-6709.2010.01150.x
Goldwater, M. B., & Echols, C. H. (in preparation). The progressive alignment of grammatical constructions: evidence from structural priming.
Goldwater, M. B., Friedman, S. E., Gentner, D. G., Forbus, K. F., Taylor, J. L M. (2011). An analogical learning model of the development of thematic roles & structural priming. Poster presented at the *36th Annual Boston University Conference on Language Development*.
Kidd, E., Bavin, E. L., & Rhodes, B. (2001). Two-year-olds' knowledge of verbs and argument structures. In M. Almgren, A. Barreña, M. J. Ezeizabarrena, I. Idiazabal, & B. MacWhinney (Eds.), *Research on child language acquisition: Proceedings of the 8th Conference of the International Association for the Study of Child Language* (pp. 1368–1382). Somerville, MA: Cascadilla Press.
Kotovsky, L., & Gentner, D. (1996). Comparison and categorization in the development of relational similarity. *Child Development*, 67, 2797–2822. https://doi.org/10.2307/1131753
Kuehne, S., Forbus, K., Gentner, D., & Quinn, B. (2000, August). SEQL: Category learning as progressive abstraction using structure mapping. In *Proceedings of the 22nd Annual Meeting of the Cognitive Science Society* (Vol. 4, pp. 286–291).
MacWhinney, B. (2014). Item-based patterns in early syntactic development. In T. Herbst, H. Schmid, & S. Faulhaber (Eds.), *Constructions, collocations, patterns* (pp. 33–70). Berlin: De Gruyter Mouton.
Matthews, D., & Bannard, C. (2010). Children's production of unfamiliar word sequences is predicted by positional variability and latent classes in a large sample of child-directed speech. *Cognitive Science*, 34, 465–488. https://doi.org/10.1111/j.1551-6709.2009.01091.x
Naigles, L. (1990). Children use syntax to learn verb meanings. *Journal of Child Language*, 17, 357–374. https://doi.org/10.1017/S0305000900013817
Naigles, L. G., & Kako, E. T. (1993). First contact in verb acquisition: Defining a role for syntax first contact in verb acquisition. *Child Development*, 64, 1665–1687. https://doi.org/10.2307/1131462
Noble, C., Iqbal, F., Lieven, E., & Theakston, A. (2016). Converging and competing cues in the acquisition of syntactic structures: The conjoined agent intransitive. *Journal of Child Language*, 43, 811–842.
Noble, C. H., Rowland, C. F., & Pine, J. M. (2011). Comprehension of argument structure and semantic roles: Evidence from English-learning children and the forced-choice pointing paradigm. *Cognitive Science*, 35, 963–982. https://doi.org/10.1111/j.1551-6709.2011.01175.x

Onnis, L., Waterfall, H. R., & Edelman, S. (2008). Learn locally, act globally: Learning language from variation set cues. *Cognition*, 109, 423–430. https://doi.org/10.1016/j.cognition.2008.10.004
Peters, A. M. (1983). *The units of language acquisition*. Cambridge: Cambridge University Press.
Pine, J. M., & Lieven, E. V. M. (1993). Reanalyzing rote-learned phrases–Individual-differences in the transition to multi-word speech. *Journal of Child Language*, 20, 551–571.
Pozzan, L., Gleitman, L. R., & Trueswell, J. C. (2016) Semantic ambiguity and syntactic bootstrapping: The case of conjoined-subject intransitive sentences. *Language Learning and Development*, 12, 14–41. https://doi.org/10.1080/15475441.2015.1018420
R Core Team. (2014). *R: A language and environment for statistical computing R*. Vienna, Austria: Foundation for Statistical Computing. <http://www.R-project.org>
Reas, C., & Fry, B. (2007). *Processing: A programming handbook for visual designers and artists*. Cambridge, MA: The MIT Press.
Richland, L. E., Morrison, R. G., & Holyoak, K. J. (2006). Children's development of analogical reasoning: Insights from scene analogy problems. *Journal of Experimental Child Psychology*, 94(3), 249–273. https://doi.org/10.1016/j.jecp.2006.02.002
Rowland, C. F. (2007). Explaining errors in children's questions. *Cognition*, 104(1), 106–134. https://doi.org/10.1016/j.cognition.2006.05.011
Sakas, W. G., & Fodor, J. D. (2012). Disambiguating syntactic triggers. *Language Acquisition*, 19(2), 83–143. https://doi.org/10.1080/10489223.2012.660553
Savage, C., Lieven, E., Theakston, A., & Tomasello, M. (2003). Testing the abstractness of children's linguistic representations: lexical and structural priming of syntactic constructions in young children. *Developmental Science*, 6(5), 557–567. https://doi.org/10.1111/1467-7687.00312
Thibaut, J. P., French, R., & Vezneva, M. (2010). The development of analogy making in children: Cognitive load and executive functions. *Journal of Experimental Child Psychology*, 106, 1–19. https://doi.org/10.1016/j.jecp.2010.01.001
Thibaut, J. P., & Witt, A. (2015). Young children's learning of relational categories: Multiple comparisons and their cognitive constraints. *Frontiers in Psychology*, 6, 643. https://doi.org/10.3389/fpsyg.2015.00643
Tomasello, M. (2003). *Constructing a language: A usage-based theory of language acquisition*. Cambridge, MA: Harvard University Press.
Tomasello, M. (2003). *Constructing a language*. Cambridge, MA: Harvard University Press.
Twomey, K. E., Chang, F., & Ambridge, B. (2014). Do as I say, not as I do: A lexical distributional account of English locative verb class acquisition. *Cognitive Psychology*, 73, 41–71. https://doi.org/10.1016/j.cogpsych.2014.05.001
Wonnacott, E., Boyd, J. K., Thomson, J., & Goldberg, A. E. (2012). Input effects on the acquisition of a novel phrasal construction in 5 year olds. *Journal of Memory and Language*, 66, 458–478. https://doi.org/10.1016/j.jml.2011.11.004
Yuan, S., & Fisher, C. (2009). "Really? She blicked the baby?": Two-year-olds learn combinatorial facts about verbs by listening. *Psychological Science*, 20, 619–626. https://doi.org/10.1111/j.1467-9280.2009.02341.x

CHAPTER 9

The labeling problem in syntactic bootstrapping

Main clause syntax in the acquisition of propositional attitude verbs

Aaron Steven White[i], Valentine Hacquard[ii] and Jeffrey Lidz[ii]
[i]University of Rochester / [ii]University of Maryland

In English, the distinction between belief verbs, such as *think*, and desire verbs, such as *want*, is tracked by tense found in the subordinate clauses of those verbs. This suggests that subordinate clause tense might be a useful cue for learning the meanings of these verbs via *syntactic bootstrapping*. However, the correlation between tense and the belief v. desire distinction is not cross-linguistically robust; yet the acquisition profile of these verbs is similar cross-linguistically. Our proposal in this chapter is that, instead of using concrete cues like subordinate clause tense, learners may utilize more abstract syntactic cues that must be tuned to the syntactic distinctions present in a particular language. We present computational modeling evidence supporting the viability of this proposal.

Keywords: syntactic bootstrapping, word learning, verb learning, propositional attitude verbs, main clause syntax, labeling problem

1. Introduction

Syntactic bootstrapping encompasses a family of approaches to verb learning wherein learners use the syntactic contexts a verb is found in to infer its meaning (L. Gleitman, 1990; Landau & Gleitman, 1985). Any such approach must solve two problems. First, it must specify how learners *cluster* verbs – i.e. figure out that some set of verbs shares some meaning component – according to their syntactic distributions. For instance, an English learner might cluster verbs based on whether they embed tensed subordinate clauses (1) or whether they take a noun phrase (2).

(1) a. John {thinks, believes, knows} that Mary is happy.
b. *John {wants, needs, orders} that Mary is happy.

https://doi.org/10.1075/tilar.24.09whi

(2) a. John {believes, knows, wants, needs} Mary.
b. *John {thinks, orders} Mary.

For different parts of the lexicon, different clusterings may be better or worse. Among propositional attitude verbs – like *think, believe, know, want, need*, and *order* – clustering based on whether the verb takes a subordinate clause yields intuitively better clusters than clustering based on whether it takes a noun phrase (at least when these structures are considered in isolation). That is, CLUSTERs 1 and 2 are intuitively more coherent than CLUSTERs 3 and 4 (see White, Hacquard, and Lidz 2018 for empirical corroboration of this intuition).

(3) a. CLUSTER 1: think, believe, know
b. CLUSTER 2: want, need, order
c. CLUSTER 3: believe, know, want, need
d. CLUSTER 4: think, order

We refer to the problem of choosing how to cluster verbs based on syntactic context as the *clustering problem*, and we call the learning mechanism that solves this problem – i.e. outputs clusters like those in (3) – the *clustering mechanism*.

The second problem a syntactic bootstrapping approach must solve involves the method by which learners *label* the clusters output by a clustering mechanism – i.e. figure out what meaning component a particular cluster of verbs corresponds to. For instance, a common way of labeling CLUSTERS 1 and 2 is to say that all verbs in CLUSTER 1 have a BELIEF component and all verbs in CLUSTER 2 have a DESIRE component.

(4) a. CLUSTER 1 ↔ BELIEF
b. CLUSTER 2 ↔ DESIRE

We refer to this second problem, which is in many ways more difficult than the clustering problem, as the *labeling problem*, and we call the learning mechanism that solves this problem – i.e. labels the clusters output by the clustering mechanism – the *labeling mechanism*.

In this chapter, we present evidence from the domain of propositional attitude verbs that previous labeling mechanisms are unsatisfactory both empirically and explanatorily, and we propose a novel labeling mechanism. Propositional attitude verbs are a useful case study for making this point because they are a parade case of verbs that fall prey to the *observability problem* – one cannot see propositional attitudes such as thinking or wanting – and thus likely require learners to rely heavily on syntactic evidence for their acquisition (Gillette, Gleitman, Gleitman, & Lederer, 1999; L. Gleitman, 1990; L. R. Gleitman, Cassidy, Nappa, Papafragou, & Trueswell, 2005; Papafragou, Cassidy, & Gleitman, 2007; Snedeker & Gleitman, 2004).

We focus in particular on the distinction among propositional attitude verbs between belief verbs, like *think*, and desire verbs, like *want*, because properties of this distinction make the empirical and explanatory inadequacies of prior approaches particularly apparent. But our proposal has implications well beyond explaining this distinction – indeed, implications for the fundamental architecture of the learning mechanism itself.

In Section 2, we discuss the two main approaches to solving the labeling problem that have been instantiated in the literature – the *top-down approach* and the *bottom-up approach* – and show that both of these approaches are inadequate: the top-down approach makes incorrect predictions and the bottom-up approach makes essentially no predictions at all. In Section 3, we propose a modification to the top-down approach that makes correct predictions. In Section 4, we present a learning algorithm that implements our proposal. In Section 5, we present a proof-of-concept experiment, which shows that our algorithm finds a correct labeling when run on syntactic distributions found in child-directed speech. In Section 6, we discuss what our proposal entails for the theory of verb learning, and in Section 7, we conclude.

2. Approaches to the labeling problem

Current approaches to the labeling problem fall into two broad categories: the *top-down approach* and the *bottom-up approach*. In the top-down approach – the traditional one laid out in L. Gleitman 1990; Landau and Gleitman 1985 – labeling is part-and-parcel with clustering. The learner has some innate mappings – *projection rules* – from semantic features to syntactic features (Carter, 1976; Chomsky, 1981; Grimshaw, 1990; Gruber, 1965; Hale & Keyser, 2002; Levin, 1993; Pinker, 1984, 1989), and upon noticing that a particular verb occurs with a particular syntactic feature, the learner "reverses" those projection rules to get from that syntactic context to that word's corresponding semantic components (Kako, 1997; Lidz, Gleitman, & Gleitman, 2004; White, 2015).

Continuing the examples above, the fact that a verb takes a tensed subordinate clause correlates (in English) with that verb being a belief verb (5a), like *think, believe*, or *know* (Anand & Hacquard, 2013; Bolinger, 1968; Farkas, 1985; Heim, 1992; Stalnaker, 1984; Villalta, 2000, 2008, among others). This is corroborated by the fact that desire verbs, like want, prefer, and order, (arguably) do not have a belief component (though see Heim 1992) and do not take tensed subordinate clauses (5b).

(5) a. John {thinks, believes, knows} that Mary is happy.
b. *John {wants, needs, orders} that Mary is happy.

Assuming this correlation to be cross-linguistically robust, a syntactic bootstrapping account might then posit that learners have some innate projection rule (6a) that they can reverse to get from the fact that *think, believe*, and *know* occur with finite complements (6b) to the fact that *think, believe*, and *know* have a meaning that involves belief (6c) (de Villiers, 2005; De Villiers & De Villiers, 2000; De Villiers & Pyers, 2002).

(6) **Top-down approach**
a. *Knowledge*: BELIEF → S[+TENSE]
b. *Data*: {think, believe, know} S[+TENSE]
c. *Inference*: BELIEF ← {think, believe, know}

The top-down approach makes strong predictions about learners' inferences: labeling is an automatic consequence of noticing a distributional fact – in this case, that some verb takes a tensed subordinate clause.

One difficulty that arises with the top-down approach is that it is not robust to cross-linguistic variation. For instance, suppose that the projection rule in (6a) were innate. One would expect either (i) that all languages show a correlation between a verb's having a belief component and its taking tensed clauses; or (ii) that, if a language does allow non-belief verbs – e.g., desire verbs like *want* – to take tensed clauses, learners might go through a stage where they incorrectly believe that those verbs actually have a BELIEF component. Neither of these possibilities are realized: (i) there are languages, such as German (7) and Spanish (8), where both belief and desire verbs take tensed subordinate clauses; and (ii) in these languages – or at the very least, in German – children do not mistake one type of verb for the other (Perner, Sprung, Zauner, & Haider, 2003).

(7)

a.	Ich	glaube,	dass	Peter	nach	Hause	geht.
	I	think	that	Peter	to	home	goes.
b.	Ich	will,	dass	Peter	nach	Hause	geht.
	I	want	that	Peter	to	home	goes.

(8)

a.	Creo	que	Peter	va	a	la	casa.
	think.1S.PRES	that	Peter	go.PRES.IND	to	the	house.
b.	Quiero	que	Peter	vaya	a	la	casa.
	want. 1S.PRES	that	Peter	go.PRES.SBJ	to	the	house.

The bottom-up approach remedies this issue at the cost of making weaker distributional and developmental predictions. In the bottom-up approach (Alishahi & Stevenson, 2008; Barak, Fazly, & Stevenson, 2012, 2013, 2014a, 2014b), learners cluster verbs (9c-i) based on syntactic context (9b), but the clustering mechanism itself does not provide the labels for these clusters. Rather, the learner must notice some correlation between the unlabeled clusters and the sorts of conceptualizations that are triggered by external stimuli when that cluster is instantiated (9c-ii).[1] Then, given the cluster each verb falls into (9c-i) and the labeling of that cluster (9c-ii), the learner can make the inference that those verbs have that label (9c-iv).[2]

(9) **Bottom-up approach**
  a. *Knowledge*: [Ø]
  b. *Data*: (BELIEF, {think, believe, know} S[+TENSE])
  c. *Inferences*
    i. CLUSTER 1 ← {think, believe, know}
    ii. CLUSTER 1 ↔ BELIEF
    iii. CLUSTER 1 ↔ S[+TENSE]
    iv. BELIEF ← {think, believe, know}
    v. BELIEF → S[+TENSE]

1. The bottom-up approach is similar in form to semantic bootstrapping, in which learners are presumed to have access to the semantics relevant to a particular learning instance (Grimshaw, 1981, 1994; Pinker, 1984, 1989, 1994): in this case, the fact that the conceptual content BELIEF "cooccurs" with the linguistic content {think, believe, know} S[+TENSE]. It differs, however, in the sense that semantic bootstrapping is a theory of how children come to learn the syntax of their language, whereas the bottom-up approach assumes access to the syntax as a prerequisite. Further, the traditional version of the semantic bootstrapping is more like the top-down approach in assuming learners have innate projection rules (cf. Connor, Fisher, and Roth 2013).

2. There is of course knowledge that learners are required to have for either approach to work that we are not listing here. For example, both the top-down and bottom-up approaches require (i) that the relevant syntactic structures can be parsed by learners at the relevant developmental stage, and (ii) that the relevant conceptual material is accessible to them at that stage. This second requirement may or may not be met at certain points in development. See Baillargeon, Scott, and He 2010; Onishi and Baillargeon 2005 for evidence that this conceptual material is accessible from a very young age.

The bottom-up approach is thus robust to cross-linguistic variability, since a learning mechanism that implements it can learn arbitrary projection rules – e.g. from BELIEF to tense (9c-v) in English – by noticing a correlation between the cluster and the syntax (9c-iii).[3]

This robustness is also the source of its major problem. To make the inferential step in (9c-ii), the learner must have access to the pairing of a conceptualization – e.g. BELIEF – with a word – e.g. *think*. But there is mounting evidence that, even in contexts that are constructed so as to heavily bias toward activating abstract concepts like BELIEF and DESIRE, propositional attitude meanings are not considered as candidates the majority of the time (Papafragou et al., 2007); and in more naturalistic contexts, they are almost never considered (Gillette et al., 1999; L. R. Gleitman et al., 2005; Snedeker & Gleitman, 2004).

If we take this problem of observability seriously, we need a way to resolve the labeling problem, and a bottom-up approach just cannot work. For this reason, we pursue a solution that modifies the top-down approach so as to be robust to cross-linguistic variability. We present the outlines of our proposal in Section 3. We implement this proposal in Section 4, and we conduct an experiment using this implementation in Section 5.

3. Our proposal

To reiterate, a major challenge for the standard top-down approach is that the particular syntactic features associated with belief vs. desire verbs differ cross-linguistically. In English, this meaning distinction is tracked by whether the complement clause is tensed or not. But as we saw, in German and Spanish, both belief and desire verbs take tensed complements.

---

3. This is similar to the explanation for how the label itself is associated with the cluster. Indeed, some bottom-up models, such as Alishahi and Stevenson's (2008), explicitly treat the association between the cluster and the concept as of the same type as the association between the cluster and the syntactic feature – i.e. the projection rule (see also Barak et al., 2012, 2013, 2014a, 2014b). This is because they treat both the concept and the syntax as observed features of the verb, which can both be used in forming the cluster in the first place.

In fact, for the purposes of learning a word's meaning, the syntactic features are to some extent superfluous for models that employ the bottom-up approach, since the semantics themselves are observed and can thus contribute to forming a cluster with a particular label. In this sense, the bottom-up approach is essentially a cross-situational word learning model (Medina, Snedeker, Trueswell, & Gleitman, 2011; Smith & Yu, 2008; Trueswell, Medina, Hafri, & Gleitman, 2013; Yu & Smith, 2007, 2012) with additional context features (cf. Frank, Goodman, & Tenenbaum, 2009).

Interestingly, the belief v. desire distinction is still tracked by the syntax of subordinate clauses, albeit via different means. In Spanish (and other Romance languages), it is tracked by the mood of the subordinate clause: belief verbs tend to take subordinate clauses with indicative mood and desire verbs tend to take subordinate clauses with subjunctive mood, exemplified in (8) from Section 2. And in German, the distinction is tracked by whether the complement has verb second (V2) syntax (Scheffler, 2008; Truckenbrodt, 2006): belief verbs tend to allow subordinate clauses with V2 syntax and desire verbs tend not to, exemplified in (10).

(10)	a.	Ich	glaube,	Peter	geht	nach	Hause.
		I	think	Peter	goes	to	home.
	b.	*Ich	will,	Peter	geht	nach	Hause.
		I	want	Peter	goes	to	home.

We argue that these syntactic features converge at an abstract level and that this convergence could help resolve the cross-linguistic challenge for a top-down approach (see Hacquard, 2014; Hacquard & Lidz, 2017).

In particular, belief verbs take subordinate clauses with syntactic hallmarks of declarative main clauses in their respective languages. For instance, in English, the subordinate clauses of belief verbs tend to be tensed, just like declarative main clauses in English; in Romance, the subordinate clauses of belief verbs tend to have the indicative mood, just like declarative main clauses in Romance; and in German, the subordinate clauses of belief verbs can have verb second word order, just like German declarative main clauses. Analogously, desire verbs tend to take subordinate clauses that show hallmarks of imperative main clauses. (This latter claim is somewhat more tentative; though interestingly, we find that assuming its truth significantly improves the performance of the model we describe below – at least on English.)

We propose to exploit this cross-linguistic convergence by invoking what we term *abstract projection rules*. An abstract projection rule is a generalization of the traditional notion of a projection rule discussed in Section 2. Instead of a particular semantic component – e.g. BELIEF – mapping onto a particular syntactic feature value – e.g. S[+TENSE] – in an abstract projection rule, a particular semantic component maps onto a set of unvalued syntactic features. Over the course of learning, learners must learn a valuation of the syntactic features that appear in this abstract projection rule before that rule can be used in syntactic bootstrapping.

The *featural anchor* for an abstract projection rule is a class of syntactic structures that (i) determine how the syntactic features in the abstract projection rule are valued and (ii) are identifiable prior to verb learning. That is, the featural anchor is, in essence, a valuation of the syntactic features listed in the abstract projection rule that is easy to identify.

Based on the correlation mentioned above, we suggest that the featural anchor for the BELIEF projection rule is the declarative main clause and the featural anchor for the DESIRE projection rule is the imperative main clause – i.e. BELIEF tends to project onto whatever syntactic features are instantiated by a language's declarative main clauses and DESIRE tends to project onto whatever syntactic features are instantiated by a language's imperative main clauses. For instance, instead of the projection rule (11), which we saw fails for languages like German and Spanish, a learner might instead have rules of the form in (12).

(11) a. BELIEF → S[+TENSE]
b. DESIRE → S[-TENSE]

(12) a. BELIEF → DECLARATIVE MAIN CLAUSE
b. DESIRE → IMPERATIVE MAIN CLAUSE

Learners must then find the valuations for DECLARATIVE MAIN CLAUSE and IMPERATIVE MAIN CLAUSE specific to their language, at which point they can use (12) as they would projection rules like those in (11).

But why should main clause syntax matter? Is there a principled link between this abstract syntax and the attitude verbs' underlying semantics? We have argued that there is and that the connection comes from the association of particular clause types with different speech acts (see Hacquard, 2014; Hacquard & Lidz, 2017).

Cross-linguistically, languages devote particular clause types to different speech acts: declaratives are typically associated with assertions, imperatives with commands, and interrogatives with questions (Sadock & Zwicky, 1985). Attitude reports are often used to perform indirect speech acts (Searle, 1975), and different attitudes easily lend themselves to different indirect speech acts because of the meaning they express.

For instance, an assertion is an expression of a judgment of truth: if one asserts (13), one commits oneself to the truth of *it is raining*.

(13) It is raining.

And because belief verbs report judgments of truth (Anand & Hacquard, 2013; Bolinger, 1968; Farkas, 1985; Heim, 1992; Stalnaker, 1984; Villalta, 2000, 2008, a.o.), they easily lend themselves to indirect assertions (Anand & Hacquard, 2013; Hooper, 1975; Lewis, 2013; Simons, 2007; Urmson, 1952). A speaker can indirectly assert the content of the complement clause by implicitly endorsing the reported judgment of truth.

(14) **A**: Why are you putting on a rain jacket?
B: John {thinks, said} that it's raining.

Analogously, a command is an expression of a desire: (15a) expresses my desire that you leave.

(15) (You,) leave!

Desire verbs report desire, and hence easily lend themselves to indirect commands. A speaker can indirectly demand that the state of affairs expressed by the complement clause be brought about by implicitly endorsing the reported desire – compare the pragmatic effects of (15) to those of (16).

(16) I want you to leave

To summarize, attitude verbs seem to split into two main classes: those that express a judgment of truth (belief and speech verbs) and those that express a preference (desire and command verbs). This semantic split is reflected both in the syntax, and in the pragmatics.

In the pragmatics, it is reflected by the type of indirect speech acts that these verbs are routinely used for: indirect assertions for belief verbs, indirect commands for desire verbs. Syntactically, this split seems to be tracked by syntactic features of their complement clauses: belief verbs take complements that resemble declarative main clauses – the syntax typically used for assertions – and desire verbs take complements that resemble imperative main clauses – the syntax typically used for commands.

We suggest that learners may exploit these parallels between speech act and clause type, inferring that a verb that takes a complement with "assertive" syntax expresses a judgment of truth, while a verb that takes a complement with "imperative" syntax expresses a preference. In Section 4, we demonstrate how to implement this idea in a learning model, and we apply this implementation to child-directed speech data in Section 5.

4. Implementing our proposal

In this section, we define a probabilistic model that implements the proposal from Section 3. We do this in two steps; first, we describe a base model that clusters verbs based on their syntactic distributions; second, we show how to augment this model with the abstract projection rule and featural anchor proposed in Section 3.

4.1 Base model

Our base model has two components. The first component describes the relationship between a verb's semantic features and its acceptability in syntactic contexts with particular feature valuations. We refer to this component as the *competence model*. The second component describes the relationship between a verb's acceptability in a particular syntactic context and the syntactic contexts it actually occurs in.

4.1.1 *The competence model*

We base our competence model on White and Rawlins's (2016) model of semantic selection – itself based on White's (2015, Chapter 3) model of syntactic bootstrapping. The competence model has two components: a representation of a verb's semantic components and a representation of projection rules. Both components are represented as probabilities: s_{vk} is the probability that verb v has semantic component k, and p_{kf} is the probability that a semantic component k projects onto syntactic feature f.

Following White and Rawlins, we define the probability d_{vf} that a verb v is acceptable with a particular syntactic feature f in terms of s_{vk} and p_{kf}.

(17) $d_{vf} \equiv 1 - \prod_k 1 - s_{vk}p_{kf}$ where k ranges over semantic components

The definition in (17) follows from the assumption that s_{vk} and p_{kf} are independent; in this case, $s_{vk}p_{kf}$ is the joint probability that verb v has semantic component k and that semantic component k projects onto syntactic feature f. Thus, $1 - s_{vk}p_{kf}$ is the probability that either verb v does not have semantic component k or semantic component k does not project onto syntactic feature f. Again assuming independence, $\Pi_k\ 1 - s_{vk}p_{kf}$ gives the probability that, for all semantic components k, either verb v does not have semantic component k or semantic component k does not project onto syntactic feature f. Finally, (17) itself gives the probability that there is some semantic component k such that verb v has semantic component k and that semantic component k projects onto syntactic feature f, and thus that a verb v is acceptable with syntactic feature f (see White and Rawlins 2016 for an explicit derivation).

4.1.2 *The performance model*

To complete the base model, we need some way of linking the competence model to the observed data. As it stands, the model only specifies the probability d_{vf} that a particular verb v is acceptable with a particular syntactic feature f. This is importantly distinct from the probability of actually seeing verb v with syntactic feature f.

There are various ways to model the latter probability. In the current case, we define a probability o_{vk} that verb v instantiates semantic component k on any particular observation of a verb. Then, we define the probability d_{vf} of seeing verb v with syntactic feature f on any particular observation of the verb.

(18) $\hat{d}_{vf} \equiv 1 - \prod_k 1 - o_{vk} s_{vk} p_{kf}$

The definition in (18) follows from reasoning analogous to that given for (17) in Section 4.1.1.

We assume that each datapoint i consists of a verb v_i and a syntactic feature combination x_i, which represents a sequence of F binary syntactic features as a bit vector of length F, where $x_{if} = 0$ means that the *fth* syntactic feature has a – value and $x_{if} = 1$ means that the *fth* syntactic feature has a + value. We then define the likelihood of the *ith* feature valuation $\mathbf{x}_i$ given **S, O, P** and the i^{th} verb v_i as in (19), assuming that each feature valuation x_{if} arises via an independent Bernoulli distribution.

(19) $\mathbb{P}(\mathrm{x}_i \mid v_i, \mathbf{S}, \mathbf{O}, \mathbf{P}) = \prod_f \mathrm{Bernoulli}(x_{if}; \hat{d}_{vf}) = \prod_f \hat{d}_{v_i f}^{x_i f}(1-\hat{d}_{v_i f})^{1-x_{if}}$

Assuming that each $\mathbf{x}_i$ is conditionally independent (given v_i, S, O, P) of all the other datapoints, the log-likelihood of the entire dataset is given by (20).

(20) $\mathcal{L}(\mathbf{X} \mid \mathrm{v}, \mathbf{S}, \mathbf{O}, \mathbf{P}) = \Sigma_i \Sigma_f x_{if} \log \hat{d}_{vif} + (1-x_{if}) \log(1-\hat{d}_{v_i f})$

Our objective is to find values for **S, O**, and **P** that maximize (20), relative to some set of constraints. We need at least two somewhat uninteresting constraints. The first is that, because they are probabilities, o_{vk}, s_{vk}, and p_{kf} should be between 0 and 1 (for all *v*, *k*, and *f*).

The second constraint pertains to the fact that, as we have stated the optimization problem, **S** and **O** are not identifiable: for arbitrary v and f, if $s_{\mathrm{vk}} = x \neq y = o_{vk}$, there is an equivalent model, with respect to (19), wherein $s_{vk} = y \neq x = o_{vk}$. To remedy this issue, we place an independent sparse Beta(0.5, 0.5) prior on s_{vk} and an independent dense Beta(2, 2) prior on o_{vk} for all *v*, *k*. This encourages the model to associate values closer to 0 or 1 with s_{vk} and values closer to 0.5 with o_{vk}. Thus, if the model finds that a value near 0 or 1 is necessary, it favors placing it in **S** rather than **O**. Beyond breaking the symmetry of **S** and **O** with respect to the objective, this also means that the model is encouraged to make confident guesses about which verbs have which semantic components.

With these somewhat uninteresting – but nonetheless necessary – constraints set up, we now turn to a more interesting constraint: the one that we use to implement our proposal on top of the base model discussed up until now.

4.2 Implementing abstract projection rules and featural anchors

To implement our proposal, we need some way of representing abstract projection rules and featural anchors for those rules. Recall that we proposed the abstract projection rules in (21), where the right hand side of this rule indicates the class of structures – i.e. the featural anchors – one must observe to fix the syntactic feature valuation for that rule.

(21) a. BELIEF → DECLARATIVE MAIN CLAUSE
b. DESIRE → IMPERATIVE MAIN CLAUSE

To allow our model to fix these features, we pretend that declarative and imperative main clauses are themselves embedded under abstract verbs ASSERT and REQUEST, respectively. So every time our model receives a sentence to process it will also receive either a datapoint like (ASSERT, **x**) or one like (REQUEST, **x**), where **x** encodes the syntactic features of the main clause of that sentence. For instance, in English, the main clause features observed with ASSERT will tend to be that the clause has a subject and tense but no complementizer, and the main clause features observed with REQUEST will tend to be that the clause has neither a subject nor tense nor a complementizer. (We ignore interrogative main clauses for the purpose of our experiment, but they can be treated in an analogous way.)

Then, we initialize the model in such a way that ASSERT and REQUEST have only a single semantic component each (and no other semantic components) with probability 1. We stipulate that the semantic component ASSERT has with probability 1 is the BELIEF component – i.e. $s_{\text{ASSERT, BELIEF}} = 1$ and $s_{\text{ASSERT, K}} = 0$ if $k \neq$ BELIEF – and that the semantic component that REQUEST has with probability 1 is the DESIRE component – i.e. $s_{\text{REQUEST, DESIRE}} = 1$ and $s_{\text{REQUEST, K}} = 0$ if $k \neq$ DESIRE.

We then disallow the model from raising the probability of any other semantic component for ASSERT or REQUEST over the course of learning (as it does for observed verbs, such as *think*). This ensures that the projection rules p_{BELIEF} and p_{DESIRE} have a strong pressure to have high probability for syntactic features that are observed in declarative main clauses and imperative main clauses, respectively, since this ensures that $\text{d}_{\text{ASSERT}} = \text{p}_{\text{BELIEF}}$ and $\text{d}_{\text{REQUEST}} = \text{p}_{\text{DESIRE}}$ regardless of how many sentences have been observed.

Crucially, note that we are *not* hard-coding what the projection rules p_{BELIEF} and p_{DESIRE} look like. These are randomly initialized and change over the course of learning (see Section 4.3 for details). Rather, we set the model up in such a way that where is a strong pressure to have a p_{BELIEF} and a p_{DESIRE} that give high probability for whatever syntactic features are observed with ASSERT and REQUEST, which will differ across languages.

Relatedly, there is no *a priori* guarantee that this implementation will work even for English, which is why it is useful to test it on real data. While the sorts of clauses that belief and desire verbs take tend to match the feature valuations of declarative and imperative clauses, respectively, they do not do so perfectly: belief verbs like *think* and *know* can take complementizers, which are not found in main clauses, and desire verbs like *want* and *order* take infinitives with subjects, where imperatives tend not to have subjects are do not contain the infinitival *to*. Our implementation can in principle handle such partial matches between clauses, but it is an empirical question whether it can do so on real data – one which we address in our experiment in Section 5.

4.3 Learning algorithm

Many different kinds of learners can be defined to respect this model's assumptions to varying degrees. Here, we define an incremental learner, which observes pairings of verbs and syntactic features one at a time and makes inferences after each observation. This learner is implemented using a form of stochastic gradient descent with adaptive gradient (Duchi, Hazan, & Singer, 2011). We do not delve into the specifics of this algorithm, though we do give a high-level description of what it is doing.

The learner begins with randomly initialized matrices **S**, **O**, and **P** with positive values near 0 (except for $\mathbf{s}_{\text{ASSERT}}$ and $\mathbf{s}_{\text{REQUEST}}$, which are constrained as described in Section 4.2). Upon receiving a particular subset of datapoints corresponding to the verbs found in a particular sentence, the learner calculates how likely that datapoint is given the current model using the Bernoulli likelihood function given in (19). The learner then attempts to change the semantic representation for the verb $\mathbf{s}_v$ and the projection rules **P** so that they give a higher likelihood to the data. The adaptive gradient piece of the learner ensures that the changes to the verb's semantic representation are not very extreme if the verb has been seen many times before but are potentially extreme if the verb is very infrequent.

5. Experiment

We now apply the learning algorithm described in Section 4.3 to a dataset of English child-directed speech. We begin by describing the dataset and then present the results. All data and code for this experiment are available at github.com/aaronstevenwhite/MainClauseModel.

5.1 Data

We utilize the subcategorization frame data extracted by (White, Resnik, Hacquard, & Lidz, under revision) from the Gleason corpus (Gleason, 1980) in CHILDES (MacWhinney, 2014a, 2014b). Gleason is a useful corpus in our case for a couple reasons. First, it contains transcripts for 24 children in the age range that children are acquiring propositional attitude verbs: 2;1 to 5;2 (de Villiers, 2005; De Villiers & De Villiers, 2000; De Villiers & Pyers, 2002). Second, it contains three types of transcripts that are at least somewhat representative of common situations children find themselves in on a daily basis: play contexts, one with the mother and one with the father, and meal contexts, with both the mother and the father.

For every occurrence of a verb in Gleason, White et al. use the MOR and POST morphological analyses (Parisse, 2000) and the MEGRASP dependency parses (Sagae, Davis, Lavie, MacWhinney, & Wintner, 2007) that ship with some CHILDES corpora to extract the syntactic features listed in (22).

(22) a. [+/− DIRECT OBJECT]
b. [+/− PREPOSITIONAL PHRASE]
c. [+/− EMBEDDED CLAUSE]
d. [+/− EMBEDDED SUBJECT]
e. [+/− COMPLEMENTIZER]
f. [+/− TENSE]
g. [+/− INFINITIVAL]

This means that each observation is constituted by a verb paired with a string of seven boolean values – one for each valuation of the feature.

For instance, (23) is an example of a sentence found in the dinner transcript for Bobby from the Gleason corpus. This sentence has the valuation in (23a), which would be fed to the model as (23b).

(23) Do you want me to teach you too?
a. [+ DO, − PP, + EMB CLAUSE, − EMB SUBJ, − COMP, − EMB TENSE, + EMB INFINITIVAL]
b. [1, 0, 1, 0, 0, 0, 1]

(This extraction method necessarily makes certain decisions about the syntactic structure of the sentence – e.g. that *me* in (23) is a direct object and not an embedded subject. Another reasonable annotation would be that *me* is both a direct object and an embedded subject or that it is only an embedded subject. Such decisions are unavoidable and where there is one to make White et al. follow the dependency labels available from parse itself as closely as possible.)

Only 22 of the 24 children (11 females) have transcripts for both the dinner session and the play session, and so following previous analyses of these data (Ely, Gleason, MacGibbon, & Zaretsky, 2001), we use only data from these children. For each of these 22 children, we combine the data extracted from the dinner and play contexts. Across children, the mean number of sentences in each dataset is 632.7 (median: 625, IQR: [544, 729]), and the mean number of verbs in each dataset is 1601.2 (median: 1555.5, IQR: [1358.5, 1824.75]).

5.2 Fitting

We apply the algorithm described in Section 4.3 to each of the 22 datasets, randomly selecting a sentence to reveal at each time step until the model has seen 20,000 sentences total. Given that the average number of sentences in the transcript is 632.7 and assuming this number is a reasonable lower bound on the number of sentences a child hears in a day, this simulates approximately a month's worth of input (as an upper bound).

We repeat this procedure 10 times for each dataset. All reported results are based on averages over these 10 runs. For each run, we set the total number of semantic components to eight, based on the fact that over 95% of the feature valuation tokens found across the 22 datasets come from the eight most frequent feature valuation types. Two of these semantic components are, by necessity, reserved for the BELIEF and DESIRE components, which are the only components we report on here.

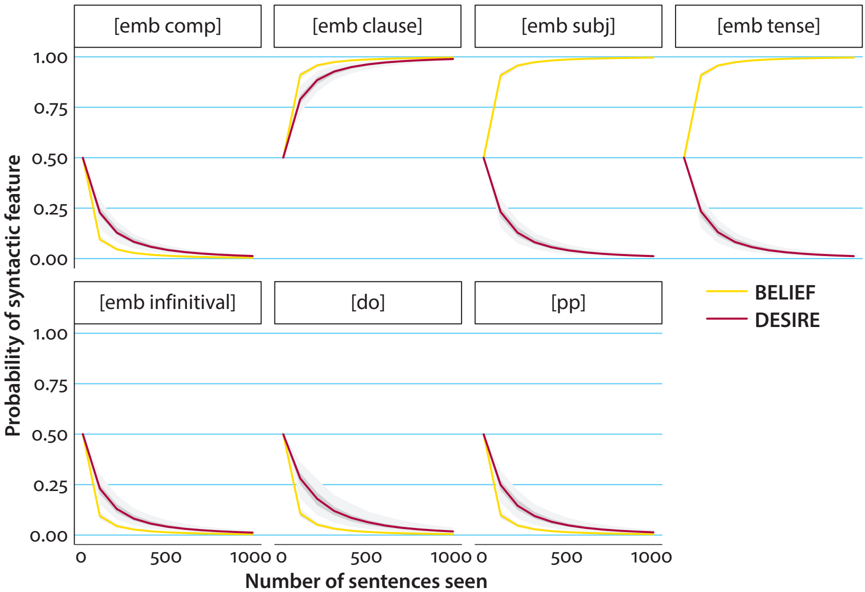


Figure 1. Median probability that BELIEF and DESIRE semantic components project onto different syntactic features as a function of the number of total sentences seen.

5.3 Results

Figure 1 shows the median probability that BELIEF and DESIRE semantic components project onto different syntactic features as a function of the number of total sentences seen. The dark shading shows the interquartile range over the 22 datasets, and the light shading show the minimum and maximum over the 22 datasets.

We see that the model robustly learns that, in English, BELIEF projects onto tensed subordinate clauses with a subject but no complementizer and that DESIRE projects onto untensed subordinate clauses without a subject or complementizer. Further, this learning happens extremely quickly: by 1000 sentences – roughly, two days worth of input – the model has converged to the aforementioned feature probabilities. This is almost certainly a product of the fact that main clauses are necessarily extremely common.

Figure 2 shows the median probability of BELIEF and DESIRE semantic components for the ten most frequent propositional attitude verbs (in order of frequency in Gleason) as a function of the number of total sentences seen (whether or not those sentences contained the verb in question or not). The dark shading shows the interquartile range over the 22 datasets, and the light shading show the minimum and maximum over the 22 datasets. (Note that the scale on the *x*-axis of Figure 2 is an order of magnitude larger than that of Figure 1.)

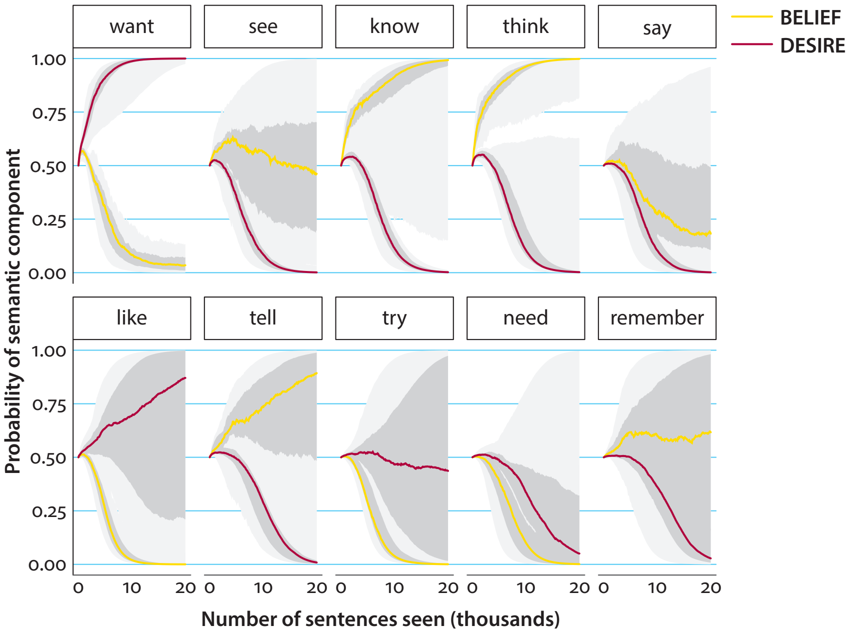


Figure 2. Median probability of BELIEF and DESIRE semantic components for ten most frequent attitude verbs as a function of the number of total sentences seen.

We see that the model robustly learns that *think* has a BELIEF component and that *want* has a DESIRE component. It also robustly learns that *want* does not have a BELIEF component and, except in the most extreme case, that *think* does not have a DESIRE component (and even then, it assigns a 50% probability to *think* having a DESIRE component).

Further, the algorithm converges to the solutions for *think* and *want* relatively quickly. We see that, by about 7,500 sentences, the algorithm has learned that *want* has a DESIRE component and that, by about 10,000 sentence, it has learned that *want* does not have a BELIEF component. The time to convergence is similar for *think*, though it takes slightly longer. This difference is likely a function of the fact that *think* is approximately 25% less frequent than *want*.

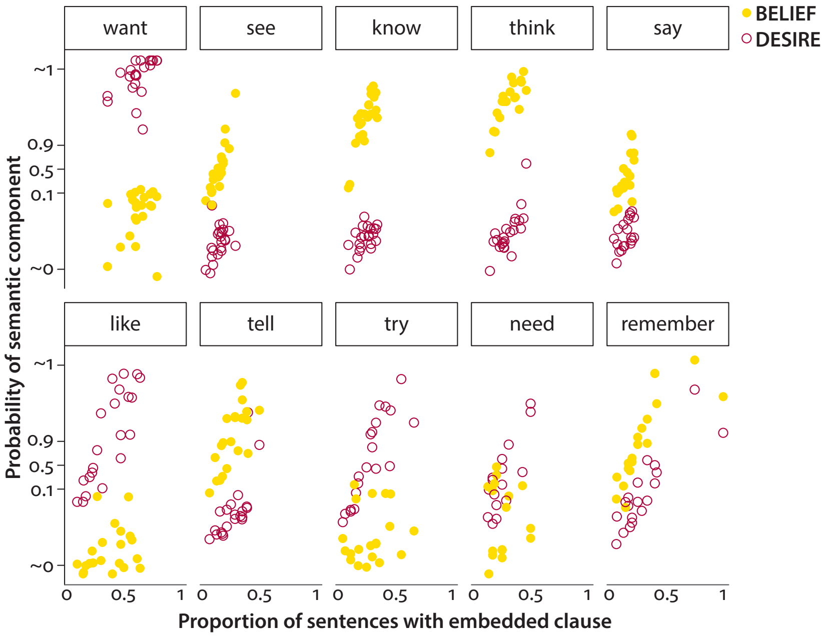


Figure 3. Mean probability of BELIEF and DESIRE semantic components after seeing the final sentence for ten most frequent attitude verbs as a function of the proportion of times that verb is found with an embedded clause in a particular transcript.

Turning now to the other verbs in Figure 2, we see that our algorithm does well in labeling *know* and *tell* with a BELIEF component on a majority of the datasets. It shows much higher variability across datasets with *remember* and *see*, and its performance with *say* is poor except in some extreme cases.

A similar variability is found for verbs like *like* and *need*, which one might expect to be labeled with a desire meaning. On a (slim) majority of the transcripts, the model assigns high probability to *like* having a DESIRE component, but it does not do this for a significant proportion. And though *need* tends to be assigned higher probability for DESIRE than for BELIEF, for most transcripts this probability is low. (It is somewhat unclear whether *try* should get a desire meaning or not, but insofar as it should, it is something of an intermediate case between *like* and *need*.)

What appears to be driving this variability is that some of the 22 children's transcripts contain very few occurrences of some verbs with embedded clauses. For instance, while *know, tell*, and *like* occur with embedded clauses on average 24%, 27%, 37% of the time, repectively, *see* and *say* occur with embedded clauses on average 15% and 14% of the time, repectively. This is corroborated when looking at how the proportion of embedded clauses found in a transcript affects the final state of the model.

(*Remember* and *need* break this trend to some extent, since they occur with embedded clauses on average 30% and 26% of the time, respectively. Upon inspection of the dataset and the dependency parses on which it is based, this appears to be driven by a combination of poor parses for a signficant portion of sentences containing *remember* and *need*.)

Figure 3 plots the mean probability of the BELIEF and DESIRE semantic components after seeing the final sentence for the ten most frequent propositional attitude verbs (in order of frequency in Gleason) as a function of the proportion of times that verb is found with an embedded clause in a particular transcript. (The y-axis is logit-scaled.) Here, we see that the relative frequency of embedded clauses for a particular a verb in a transcript is strongly related to the probability that the model assigns to the verb having either a BELIEF component or a DESIRE component (Spearman's $\rho = 0.84$).

5.4 Discussion

We have shown that our model works well for labeling core cases of belief and desire verbs, such as *think, know, want*, and *tell*, but that it shows variability for other verbs, such as *say, like*, and *need*. We gave evidence that this variability is related to the relative frequency of clausal embedding for a particular verb.

This raises two questions. First, to what extent are the empirical relative frequencies found in each dataset indicative of each child's experience? Each transcript is only a small sample of children's experience, and so the relative frequencies plotted in Figure 3 may well not accurately reflect this experience. This matters, especially in cases where the average relative frequency of subordinate clauses for a particular verb is near 0 – e.g. for *say* – since it may well be that children's actual experience includes many more instances of the verb with a subordinate clause.

(A potential confound in these transcripts – particularly for *say* – is that the play sessions include book reading, which involves many cases of quotation – e.g. *say 'hi'*. Depending on the amount of book reading input a particular child receives, this may warp the distribution of communicative verbs like *say*.)

Second, to what extent is this pattern really a fact about relative frequency of a particular syntactic feature paired with a verb and not overall frequency of the verb? For instance, even if a verb shows up only rarely with a subordinate clause, does this necessarily mean that the child will not learn that the verb has, e.g., a belief or desire component? One possibility is that it just takes a substantially longer time than we have simulated here to learn that such verbs have a belief or desire component.

6. General discussion

The main substantive addition we make to the theory of syntactic bootstrapping in this chapter is the notion of a featural anchor, which is itself a class of structures known prior to the selection of a projection rule. One question this addition raises is to what extent learners can easily discover the featural anchor itself. A deflationary response to our proposal might argue that we have merely pushed the job of verb learning back to discovering which classes of structures constitute featural anchors: isn't the job of figuring out which syntactic features are indicative of a particular anchor just as hard as learning a verb itself?

Yes and no. Yes, because it is true that, at the end of the day, one must identify some structure as a declarative or imperative main clause, and this must presumably be done by perceiving that, in using a particular structure, an utterer of a declarative main clause intends the utterance to be taken as part of a particular conversational move – such as an assertion (Austin, 1975; Stalnaker, 1978). That is, the learner must be able to identify the illocutionary force intended for the utterance.

No, because illocutionary force is a concept that is presumably prerequisite to learning a language in a first place. Indeed, children appear to be adept at recognizing an utterance's illocutionary force quite early (Spekman & Roth, 1985). This is to say that, though language is clearly not *for* communication, the data a learner uses to learn their language tends to come wrapped in communicative acts, which the learner presumably has no problem perceiving as such.

And no, because the mapping between illocutionary force and syntactic structure is relatively stable within a language: assertions, at least as conveyed by clauses, are conveyed by clauses with the same syntactic features – in English, [+TENSE, -COMP, …]. And insofar as an assertion is not conveyed by a clause – such as when it is conveyed by a polarity particle or a fragment – we submit that, if a learner has enough syntactic knowledge to represent a clause as a set of features, they have enough syntactic knowledge to represent that the valuation of those features is dependent on the fact that that clause is a syntactic object in the first place.

This question of recognizing a syntactic class of complements, such as a clause, is related to the issue we saw our algorithm having with nominal and propositional anaphor complements, like *so*. These complements cannot be valued for the same sorts of features that a clause can be – indeed, they do not appear to be valued as such at any level of syntactic representation (Hankamer & Sag, 1976) – but as it stands our model views them as simply unvalued. But rather than view them as unvalued for these features, it seems that the distinction in syntactic class must be baked into the model itself. That is, the syntactic features the model pays attention to in making an inference from a particular piece of data must be dependent on the syntactic class involved in that datum. In the abstract, we need to incorporate some decision tree-like representation into the model. One way this might be implemented is by employing a likelihood function that incorporates a hurdle model. (See White, Rawlins, and Van Durme 2017 for a recent use of hurdle models in a related domain.)

Beyond providing a case for a decision tree-like structure for syntactic feature valuation, the case of propositional anaphors may also suggest that the mechanisms used by syntactic bootstrapping to infer a verb's meaning may need to incorporate a notion of semantic type over and above that given by syntactic type. (See White and Rawlins 2016 for evidence that semantic type signatures can be extracted from syntactic distribution.)

Another question that arises is to what extent there are other abstract projection rules and featural anchors. There seems to us to be at least one further candidate: factive and interrogative verbs like *know* and *wonder*.

Dudley (2017) shows that a large part of children's experience with *know* is in the frame *do you know Q*? We would like to suggest that, here again, the child may be able exploit the syntactic parallels between direct and indirect speech acts: *know* and *wonder* are used to ask indirect questions.

(24) a. Do you know where the keys are?
b. I was wondering where you put my keys.

Our suggestion is that, analogous to what we claim for belief and desire verbs, children might infer from this parallel that the meaning of *know* must be one that relates the subject to the answer of that question. In future work, we aim to investigate this possibility.

7. Conclusion

In this chapter, we proposed a novel solution to the labeling problem in syntactic bootstrapping that augments the standard top-down approach to syntactic bootstrapping with the concepts of an *abstract projection rule* and a concomitant *featural anchor*. We motivated this proposal by noting that neither the top-down nor the bottom-up approaches solve the labeling problem for belief and desire verbs: the top-down approach is brittle in the face of cross-linguistic variability, while the bottom-up approach makes unrealistic assumptions about the data learners have access to.

We showed that our proposed solution can deal with the labeling problem given theoretically justified featural anchors for particular labels, and using belief and desire predicates as a case study, we implemented a computational model that incorporates the labeling mechanism we propose. We presented a proof-of-concept fit of this model to data derived from child-directed speech and showed that our model works well for labeling core cases of belief and desire verbs, such as *think, know, want*, and *tell*.

References

Alishahi, A., & Stevenson, S. (2008). A computational model of early argument structure acquisition. *Cognitive Science*, 32(5), 789–834. https://doi.org/10.1080/03640210801929287

Anand, P., & Hacquard, V. (2013). Epistemics and attitudes. *Semantics and Pragmatics*, 6(8), 1–59.

Austin, J. L. (1975). *How to do things with words*. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198245537.001.0001

Baillargeon, R., Scott, R. M., & He, Z. (2010). False-belief understanding in infants. *Trends in Cognitive Sciences*, 14(3), 110–118. https://doi.org/10.1016/j.tics.2009.12.006

Barak, L., Fazly, A., & Stevenson, S. (2012). Modeling the acquisition of mental state verbs. In Proceedings of the 3rd Workshop on Cognitive Modeling and Computational Linguistics (pp. 1–10). Montreal, Canada: Association for Computational Linguistics.

Barak, L., Fazly, A., & Stevenson, S. (2013). Acquisition of desires before beliefs: A computational investigation. In Proceedings of the 17th Conference on Computational Natural Language Learning (pp. 231–240). Sofia, Bulgaria: Association for Computational Linguistics.

Barak, L., Fazly, A., & Stevenson, S. (2014a). Gradual acquisition of mental state meaning: A computational investigation. In *Proceedings of the 36th Annual Conference of the Cognitive Science Society* (pp. 1886–1891). Québec City, Canada: Cognitive Science Society.

Barak, L., Fazly, A., & Stevenson, S. (2014b). Learning verb classes in an incremental model. In Proceedings of the 2014 ACL Workshop on Cognitive Modeling and Computational Linguistics (pp. 37–45). Baltimore, MD: Association for Computational Linguistics.

Bolinger, D. (1968). Postposed main phrases: An English rule for the Romance subjunctive. *Canadian Journal of Linguistics*, 14(1), 3–30.

Carter, R. (1976). Some linking regularities. In *On linking: Papers by Richard Carter*. Cambridge, MA: MIT Center for Cognitive Science.

Chomsky, N. (1981). *Lectures on government and binding: The Pisa lectures*. Dordrecht: Foris.

Connor, M., Fisher, C., & Roth, D. (2013). Starting from scratch in semantic role labeling: Early indirect supervision. In *Cognitive Aspects of Computational Language Acquisition* (pp. 257–296). Berlin: Springer. https://doi.org/10.1007/978-3-642-31863-4_10

de Villiers, J. G. (2005). Can language acquisition give children a point of view? In J. W. Astington & J. A. Baird (Eds.), *Why language matters for theory of mind* (pp.186–219). Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195159912.003.0010

De Villiers, J. G., & De Villiers, P. A. (2000). Linguistic determinism and the understanding of false belief. In Peter Mitchell & Kevin J. Riggs (Eds.), *Children's reasoning and the mind* (pp. 191–228). Hove: Psychology Press.

De Villiers, J. G., & Pyers, J. E. (2002). Complements to cognition: A longitudinal study of the relationship between complex syntax and false-belief-understanding. *Cognitive Development*, 17(1), 1037–1060. https://doi.org/10.1016/S0885-2014(02)00073-4

Duchi, J., Hazan, E., & Singer, Y. (2011). Adaptive subgradient methods for online learning and stochastic optimization. *The Journal of Machine Learning Research*, 12, 2121–2159.

Dudley, R. (2017). The role of input in discovering presupposition triggers: Figuring out what everybody already knew (Unpublished doctoral dissertation). University of Maryland, College Park, MD.

Ely, R., Gleason, J. B., MacGibbon, A., & Zaretsky, E. (2001). Attention to language: Lessons learned at the dinner table. *Social Development*, 10(3), 355–373. https://doi.org/10.1111/1467-9507.00170

Farkas, D. (1985). *Intensional descriptions and the Romance subjunctive mood*. New York, NY: Garland.

Frank, M. C., Goodman, N. D., & Tenenbaum, J. B. (2009). Using speakers' referential intentions to model early cross-situational word learning. *Psychological Science*, 20(5), 578–585. https://doi.org/10.1111/j.1467-9280.2009.02335.x

Gillette, J., Gleitman, H., Gleitman, L., & Lederer, A. (1999). Human simulations of vocabulary learning. *Cognition*, 73(2), 135–176. https://doi.org/10.1016/S0010-0277(99)00036-0

Gleason, J. B. (1980). The acquisition of social speech routines and politeness formulas. In H. Giles, W. P. Robinson, & P. N. Smith (Eds.), *Language: Social psychological perspectives* (pp. 21–27). Oxford: Pergamon Press.

Gleitman, L. (1990). The structural sources of verb meanings. *Language Acquisition*, 1(1), 3–55. https://doi.org/10.1207/s15327817la0101_2

Gleitman, L. R., Cassidy, K., Nappa, R., Papafragou, A., & Trueswell, J. C. (2005). Hard words. *Language Learning and Development*, 1(1), 23–64. https://doi.org/10.1207/s15473341lld0101_4

Grimshaw, J. (1981). Form, function and the language acquisition device. In C. L. Baker & J. J. McCarthy (Eds.), *The logical problem of language acquisition* (pp. 165–182). Cambridge, MA: The MIT Press.

Grimshaw, J. (1990). *Argument structure*. Cambridge, MA: The MIT Press.

Grimshaw, J. (1994). Lexical reconciliation. *Lingua*, 92, 411–430. https://doi.org/10.1016/0024-3841(94)90348-4

Gruber, J. S. (1965). Studies in lexical relations (Unpublished doctoral dissertation). MIT.

Hacquard, V. (2014). Bootstrapping attitudes. *Semantics and Linguistic Theory*, 24, 330–352. https://doi.org/10.3765/salt.v24i0.2434

Hacquard, V., & Lidz, J. (to appear). Children's attitude problems: Bootstrapping verb meaning from syntax and pragmatics. *Mind and Language*.

Hale, K., & Keyser, S. J. (2002). *Prolegomena to a theory of argument structure*. Cambridge, MA: The MIT Press.

Hankamer, J., & Sag, I. (1976). Deep and surface anaphora. *Linguistic Inquiry*, 7(3), 391–428.

Heim, I. (1992). Presupposition projection and the semantics of attitude verbs. *Journal of Semantics*, 9(3), 183–221. https://doi.org/10.1093/jos/9.3.183

Hooper, J. B. (1975). On assertive predicates. In J. P. Kimball (Ed.), *Syntax and semantics* (Vol. 4, pp. 91–124). New York, NY: Academy Press.

Kako, E. (1997). Subcategorization semantics and the naturalness of verb-frame pairings. *University of Pennsylvania Working Papers in Linguistics*, 4(2), 11.

Landau, B., & Gleitman, L. R. (1985). *Language and experience: Evidence from the blind child*. Cambridge, MA: Harvard University Press.

Levin, B. (1993). *English verb classes and alternations: A preliminary investigation*. Chicago, IL: University of Chicago Press.

Lewis, S. (2013). Pragmatic enrichment in language processing and development (Unpublished doctoral dissertation). University of Maryland, College Park, MD.

Lidz, J., Gleitman, H., & Gleitman, L. (2004). Kidz in the 'hood: Syntactic bootstrapping and the mental lexicon. In D. G. Hall & S. R. Waxman (Eds.), *Weaving a lexicon* (pp. 603–636). Cambridge, MA: The MIT Press.

MacWhinney, B. (2014a). *The CHILDES project: Tools for analyzing talk* (Vol. 1). New York, NY: Psychology Press.

MacWhinney, B. (2014b). *The CHILDES project: Tools for analyzing talk* (Vol. 2). New York, NY: Psychology Press.

Medina, T. N., Snedeker, J., Trueswell, J. C., & Gleitman, L. R. (2011). How words can and cannot be learned by observation. *Proceedings of the National Academy of Sciences*, 108(22), 9014–9019. https://doi.org/10.1073/pnas.1105040108

Onishi, K. H., & Baillargeon, R. (2005). Do 15-month-old infants understand false beliefs? *Science*, 308(5719), 255–258. https://doi.org/10.1126/science.1107621

Papafragou, A., Cassidy, K., & Gleitman, L. (2007). When we think about thinking: The acquisition of belief verbs. *Cognition*, 105(1), 125–165. https://doi.org/10.1016/j.cognition.2006.09.008

Parisse, C. (2000). Automatic disambiguation of morphosyntax in spoken language corpora. *Behavior Research Methods, Instruments, & Computers*, 32(3), 468–481. https://doi.org/10.3758/BF03200818

Perner, J., Sprung, M., Zauner, P., & Haider, H. (2003). Want that is understood well before say that, think that, and false belief: A test of de Villiers's linguistic determinism on German–speaking children. *Child Development*, 74(1), 179–188. https://doi.org/10.1111/1467-8624.t01-1-00529

Pinker, S. (1984). *Language learnability and language development*. Cambridge, MA: Harvard University Press.

Pinker, S. (1989). *Learnability and cognition: The acquisition of argument structure*. Cambridge, MA: The MIT Press.

Pinker, S. (1994). How could a child use verb syntax to learn verb semantics? *Lingua* 92, 377–410. https://doi.org/10.1016/0024-3841(94)90347-6

Sadock, J. M., & Zwicky, A. M. (1985). Speech act distinctions in syntax. In T. Shopen (Ed.), *Language typology and syntactic description* (Vol. 1, pp. 155–196). Cambridge: Cambridge University Press.

Sagae, K., Davis, E., Lavie, A., MacWhinney, B., & Wintner, S. (2007). High-accuracy annotation and parsing of CHILDES transcripts. Proceedings of the Workshop on Cognitive Aspects of Computational Language Acquisition (pp. 25–32). Prague: Association for Computational Linguistics.

Scheffler, T. (2008). Semantic operators in different dimensions (Unpublished doctoral dissertation). University of Pennsylvania, PA.

Searle, J. (1975). Indirect speech acts. In P. Cole & J. L. Morgan (Eds.), *Speech acts* (Vol. 3, pp. 59–82). New York, NY: Academic Press.

Simons, M. (2007). Observations on embedding verbs, evidentiality, and presupposition. *Lingua*, 117(6), 1034–1056. https://doi.org/10.1016/j.lingua.2006.05.006
Smith, L., & Yu, C. (2008). Infants rapidly learn word-referent mappings via cross-situational statistics. *Cognition*, 106(3), 1558–1568. https://doi.org/10.1016/j.cognition.2007.06.010
Snedeker, J., & Gleitman, L. (2004). Why it is hard to label our concepts. In D. G. Hall & S. R. Waxman (Eds.), *Weaving a lexicon* (pp. 257–294). Cambridge, MA: The MIT Press.
Spekman, N. J., & Roth, F. P. (1985). Preschool children's comprehension and production of directive forms. *Journal of Psycholinguistic Research*, 14(3), 331–349. https://doi.org/10.1007/BF01068090
Stalnaker, R. (1978). Assertion. *Syntax and Semantics*, 9, 315–332.
Stalnaker, R. (1984). *Inquiry*. Cambridge: Cambridge University Press.
Truckenbrodt, H. (2006). On the semantic motivation of syntactic verb movement to C in German. *Theoretical Linguistics*, 32(3), 257–306. https://doi.org/10.1515/TL.2006.018
Trueswell, J. C., Medina, T. N., Hafri, A., & Gleitman, L. R. (2013). Propose but verify: Fast mapping meets cross-situational word learning. *Cognitive Psychology*, 66(1), 126–156. https://doi.org/10.1016/j.cogpsych.2012.10.001
Urmson, J. O. (1952). Parenthetical verbs. *Mind*, 61(244), 480–496. https://doi.org/10.1093/mind/LXI.244.480
Villalta, E. (2000). Spanish subjunctive clauses require ordered alternatives. In Brendan Jackson and Tanya Matthews (Eds.), *10th Annual Proceedings of Semantics and Linguistic Theory* (pp. 239–256). https://doi.org/10.3765/salt.v10i0.3123
Villalta, E. (2008). Mood and gradability: An investigation of the subjunctive mood in Spanish. *Linguistics and Philosophy*, 31(4), 467–522. https://doi.org/10.1007/s10988-008-9046-x
White, A. S. (2015). Information and incrementality in syntactic bootstrapping (Unpublished doctoral dissertation). University of Maryland, College Park, MD.
White, A. S., Hacquard, V. & Lidz, J. (2018). Semantic information and the syntax of propositional attitude verbs. *Cognitive Science*, 42: 416–456. https://doi.org/10.1111/cogs.12512
White, A. S., & Rawlins, K. (2016). A computational model of S-selection. In Mary Moroney, Carol-Rose Little, Jacob Collard, and Dan Burgdorf (Eds.), *26th Annual Proceedings of Semantics and Linguistic Theory* (pp. 641–663). https://doi.org/10.3765/salt.v26i0.3819
White, A. S., Rawlins, K., & Van Durme, B. (2017). The semantic proto-role linking model. In *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics* (Vol. 2, pp. 92–98). Valencia, Spain: Association for Computational Linguistics.
White, A. S., Resnik, P., Hacquard, V., & Lidz, J. (under revision). The contextual modulation of semantic information.
Yu, C., & Smith, L. B. (2007). Rapid word learning under uncertainty via cross-situational statistics. *Psychological Science*, 18(5), 414–420. https://doi.org/10.1111/j.1467-9280.2007.01915.x
Yu, C., & Smith, L. B. (2012). Modeling cross-situational word-referent learning: Prior questions. *Psychological Review*, 119(1), 21. https://doi.org/10.1037/a0026182

Author note

We are grateful to audiences at Semantics and Linguistic Theory 24, the University of Maryland, and UC Santa Cruz as well as Alexander Williams, Shevaun Lewis, Kaitlyn Harrigan, Naho Orita, Rachel Dudley, Morgan Moyer, and Tom Roberts for useful comments. This work was supported by NSF BCS DDRIG grant 1456013, NSF BCS grant 1124338, and NSF DGE IGERT grant 0801465.

CHAPTER 10

Perspectives on truth

The case of language and false belief reasoning

Jill de Villiers
Smith College

Many theorists take language – vocabulary, mental verbs, syntax, counterfactuals, discourse – to be a significant help in the development of explicit Theory of Mind. Does conversation, with all its point-of-view indicators, betray another's perspective? By comparing how different linguistic markers behave across clausal environments, I demonstrate that they fall into distinct classes, only one of which – tense – patterns with the truth of the clause in terms of perspective. Sentences with embedded finite complements thus have a special role in representing the truth or falsity of others' beliefs. Children who master embedded sentential complements can then more readily reason about others' false beliefs.

Keywords: perspective, point of view, deixis, complements, theory of mind, direct speech, syntax, finiteness

1. Introduction

There are several alternative theoretical positions for the relationship of language development to Theory of Mind development. Many have found a role for the child's exposure to relevant vocabulary, particularly of mental state terms such as *think* and *know* (Dunn & Brophy, 2005). Others find a role for general syntactic development, for children to follow conversations that reveal the cultural theory about the mind (Ruffman, Slade, Rowlandson, Rumsey, & Garnham, 2003). Some point to conversation itself as a major source of understanding others' mental states (Nelson, 2005). For example, Harris (2005) argues that conversation exposes children to the shifting perspectives of the parties involved, enriching their developing Theory of Mind. Linguistic carriers of point of view abound: these include evidentials, spatial and temporal deixis, pronoun shifts, mental state verbs with complements, and even the different words people use to designate entities depending on their knowledge of their properties. For example, the same object may be called "Spot", "that beastly dog" or "the prize winning bulldog", depending on the speaker's knowledge.

https://doi.org/10.1075/tilar.24.10vil

Among these markers, only tensed sentential complements have been clearly linked as directly contributing to the cognitive achievement that occurs around age 4 in judging and making decisions on others' false beliefs (de Villiers, 2007; de Villiers & de Villiers, 2009). Comparing several different clause types: matrix clauses, adjuncts, non-finite and finite complements, and direct and indirect speech reports, I assess how the different point-of-view markers behave in each with respect to whose perspective they carry: is it the speaker's or the subject's? From that analysis it emerges that the carriers of perspective fall into three distinct types. I argue that there is no difference between direct and indirect speech reports in their importance for encoding false belief, but a big difference between finite and non-finite speech reports. I discuss proposals by which tense carries truth.

Finally some experimental data are presented in support of the view that encoding and attributing speech reports linguistically is more critical than observation of conversation, even when that conversation implicitly carries differing perspectives.

2. Points of view

Consider the array of linguistic phenomena in items a-f. I have inadvertently worked on each of these over the course of 45 years of studying language acquisition.

a. Personal pronouns: I, you, he/she, we, they etc.
b. Spatial deixis: here/there/yonder
c. Demonstrative deixis: this/that
d. Personal taste adjectives: damned, yucky, delicious, wonderful etc.
e. Opinion adverbs: unfortunately, surprisingly, sadly, happily, etc.
f. Designators: my best friend, the president, the mayor, the Pulitzer prize winner etc.
g. Evidentials: e.g. in Tibetan, a verb marker of how the *speaker* knows the truth of a proposition he expresses

Each phenomenon raises interesting problems for acquisition.

2.1 Personal pronouns

The child's control of personal pronouns is present almost from the start of multi-word utterances in English (Clark & Sengul, 1978; Oshima-Takane, 1999). Three sets of findings suggest that acquisition is not instantaneous, however. First, there is a well-attested stage at which reversals of I/you are common, and very young children say things like "Pick you up!" when reaching to be picked up (Dale

& Crain-Thoreson, 1993). These may be routinized formulae before genuine pronouns, and reversals are more common even later in children with autism (Fay, 1979). Second, although production seems good, children struggle with using pronouns as the only clue to meaning of where something is hidden ("It's under your box!" It's under her box!") for a year or two longer (Girouard, Ricard, & Decarie, 1997; Thomas, 2010). The meaning of "you" remains ambiguous in dyadic conversation, because it could mean "the other person" rather than the addressee. The best circumstance in which to fix the meaning of "you" is actually to be in a triadic conversation, where the person being addressed can be identified as the referent for "you" (Oshima-Takane, 1988; Oshima-Takane, Takane, & Shulz, 1999). Third, the behavior of pronouns inside embedded clauses is a source of confusion and uncertainty (Tanz, 1980/2009; de Villiers, Nordmeyer, & Kravitz, 2010), not unexpectedly given the variability in the world's languages in how pronouns behave in certain environments. For instance, in Navajo and Arabic, first person pronouns can occur as the subject of embedded clauses yet coreferent with the third person matrix subject (1), as if the clause was direct speech (2) (Speas, 2004):

(1) Martha said that I bought a house.

(2) Martha said "I bought a house".

In English we would use the 3rd person pronoun to agree as in (3).

(3) Martha said that she bought a house.

Children mistakenly allow the Navajo-style reading of coreference for a first person pronoun as in (1), at least in a truth value judgment task, in the later preschool years (O'Connor, Burgin, de Villiers, Speas, & Roeper (2007).

2.2 Spatial deixis

The spatial terms *here* and *there* behave like the personal pronouns *I* and *you* in that their reference is linked to who is speaking. But their use is relative to the context: I could say "Here it is!" if I find a pen right beside me and you are across the desk from me, or I could say "Here in the US" if I am talking to you on Skype half way across the globe (Fillmore, 1975). The terms also have abstract metaphorical reference, "*Here's* the problem with such a proposal", in which the location is in discourse, that is, mental rather than physical. Finally, "there" is also used as an existential. A child must filter out these cases to make sense of the primary locational use (Roeper, 2007).

One important thing to notice: *here* and *there* are not like *I* and *you* in simply switching between speakers. If I am sitting next to you, *here* is *here* and *there* is *there* for both of us. The rule is not simply to switch meanings depending on the speaker. Many languages have more than two distinctions, as in Latin hic / iste / ille (near speaker / near hearer / away from both). Like the archaic English *yonder*, this choice adds genuine point of view considerations, in that the deictics require attention not just to one's own perspective but whether the listener shares it. "Yonder" means away from *both* of us, even if we are apart.

In this category certain direction-shifting verbs occur as well , such as *come* versus *go*, *bring* versus *take*, but they also have neutral meanings, and the directional distinction can be lost in some English dialects (Fillmore, 1997).

2.3 Demonstrative deixis

The same set of considerations apply to *this* and *that* as demonstratives, but a) a large part – perhaps a majority – of their uses is related to discourse focus rather than location e.g. "That's a good idea!" b) *that* has other common uses e.g. as a complementizer, to cloud code-cracking (for child data, see de Villiers & de Villiers, 1974).

2.4 Personal taste adjectives

Expressive adjectives take the point of view of the speaker, at least most of the time. For example, in a sentence such as (4):

(4) She brought her damn dog to the party.

"she" does not hold the negative attitude towards the dog, but rather, the speaker does. A tricky case is the word *beloved*, which has to be anchored to the subject of the sentence in (5):

(5) She brought her beloved dog to the party.

It is even possible to embed the adjectives:

(6) She brought her damn beloved dog to the party.

wherein the speaker's attitude is overlaid on that of the matrix subject. Enticingly, the possibility of embedding is not so clear in (7) when the order of adjectives is reversed:

(7) ?She brought her beloved damn dog to the party. (Roeper, p.c.)

How is the referent for these points of view fixed for the ordinary personal taste adjectives words like *delicious*, *yucky*, *wonderful* and so forth? A vast semantics literature is blossoming about these forms, but a fundamental idea is that a salient "judge" in the context – the speaker, the hearer, the sentence subject – can be chosen as the point of view for the adjective of personal taste (Potts, 2007; Lasersohn, 2005) But as we will see, syntax plays a constraining role in this assignment (Stephenson, 2007; Pearson, 2015)

2.5 Opinion adverbs

Richards (1976) discusses speaker-oriented adverbs like *luckily* that reflect the attitude of the speaker rather than the subject of a sentence. For example, one might say:

(8) Luckily, my roommate did not get accepted into Yale.

where it is typically not true that the roommate benefited from his rejection (see also Gu & Roeper, 2011). Adverbs appear fairly late in child language and these types have not been fully investigated to date.

2.6 Designators

Noun phrases themselves have a point of view not often recognized. A straightforward nominal like "the dog" is usually neutral, but when one uses a DP that is descriptive in regards to some non-perceptible attribute, that description is determined by the knowledge and beliefs of the speaker. If I refer to the same dog as the "Westminster prize winner", that is a particular description that someone else might not know. Role nouns have this quality: the President, the Mayor, the baker, the dentist. Normally any co-referring noun can be substituted in an ordinary sentence like (9) and preserve truth (Frege, 1948). In the Greek myth, the Queen of Thebes is Oedipus's mother, though he does not know that.

(9) Oedipus married the Queen of Thebes → Oedipus married his mother.

But in embedded sentences such as (10), substitution of co-referring terms does not necessarily preserve truth (11) because the word used depends on the knowledge not of the speaker but of the matrix subject.

(10) Oedipus knew he married the Queen of Thebes.

(11) Oedipus knew he married his mother.

A sizeable and contentious literature covers when children understand how to use the right noun phrases, namely understand the conditions of *referential opacity* (Russell, 1987; Apperly & Robinson, 1998; Kamawar & Olson, 1999). Most studies find that children succeed at understanding the conditions on substitution at a later age than they succeed at mastery of truth conditions in complements, and later than success on false belief tasks.

Depending on what is already established in discourse, a speaker could also choose to use pronouns instead of nouns, definite instead of indefinite articles, and so forth. In fact, young children have a notoriously hard time with devices that require keeping track of a listener's existing knowledge in a discourse (Karmiloff-Smith, 1981; Berman & Slobin, 1994). The forms involve additional considerations than just switches in perspective. (For further analysis of the complexity of determiners see Klein, 1998; van Hout, Harrigan & de Villiers, 2010; for pronouns, see e.g. Hendriks & Spenader 2005/6).

2.7 Evidentials

Evidential morphology indicates the means by which a speaker knows something. Evidentials have attracted a lot of attention partly because they seem "exotic", though almost a quarter of the world's languages are estimated to have linguistic evidential markings (Aikhenvald, 2004). In languages with evidentials these markings can be as obligatory as tense is for English: the speaker has to commit to how they know the truth of it as they articulate a proposition. Furthermore, these are not "hedges" on truth, that is, the speaker is committed to the truth of the proposition (de Villiers, Garfield, Gernet Girard, Roeper, & Speas, 2009). Evidentials are speaker-centered in ordinary sentences, that is, they are based on the information the speaker has available to her in the situation. However, in some language like Tibetan, a "reflection principle" requires consideration of the point of view of the listener when asking a question of that listener. In Tibetan, you anticipate the evidential that your listener will use in reply (Garrett, 2001; de Villiers & Garfield, 2017).

3. Conversation and perspective

All of these elements (and there are undoubtedly more) are indicators of the speaker's perspective: on identity, location, discourse focus, source of evidence, linked to existing sentiment, opinion, or knowledge. In order to be a competent speaker, the child must understand these devices and use them from his or her own perspective. Is that enough to appreciate and understand other minds?

In order to answer questions about what others believe, know, or feel – that is, explicit Theory of Mind – the child has to be able to take the perspective of the other and answer *as if* s/he were that other person. That is, the child must NOT take the ordinary tack of considering one's own perspective, but rather take a contrasting one. In the Theory of Mind literature, considerable attention has been given to the possibility of Simulation: that a child must step into the shoes of the Other and answer like that person (Gordon, 1986). But other theorists contend that there must be a more explicit, not implicit, representation of the Other's perspective to answer appropriately (Carruthers, 1996; Hutto, 2007). Some even propose that language provides that representation (de Villiers & de Villiers, 2009; Collins, 2000; Segal, 1998).

Do the devices described above not already supply that? I conjecture here that they do so only when seen metalinguistically, from above. The young child who responds appropriately to pronouns, deixis, evidentials and so forth is not yet representing anything about the Other. It is only when the child can set the forms in contrast as belonging to another person that they serve the right function for representing false beliefs. That becomes clear when one looks at how (un)successful young children are at making judgments with respect to appropriate use by a separate dyad, a task that is considerably harder than the child's own production (Clark & Sengul, 1978; O'Connor et al., 2007; de Villiers et al., 2010). In this and so many other domains, it is as if there is an implicit understanding that is sufficient to drive production and even its mirror in comprehension. However explicit decisions in a judgment task, or a task involving a computation beyond what is said, lag behind and require a higher level of representation.

The use of complement structures that attribute a propositional content to an individual may offer the right kind of representation. I make that case next.

4. Embedded complements

4.1 Propositional attitudes

The acquisition of false belief reasoning has been linked to the acquisition of sentential complements (Collins, 2000; de Villiers & de Villiers, 2009; Segal, 1998). These arguments are theoretical, in that the propositional attitudes have a set of properties that would seem to require a representation as rich as they are. For example, propositional attitudes like belief have an indicator of the holder of that attitude: Phil believes. Furthermore, a propositional attitude has content (its crucial property of intentionality: Brentano, 1874):

(12) Phil believes *p*.

where p is a proposition. Furthermore, that p may contain reference to nonexisting elements such as:

(13) Phil believes there is a unicorn.

And *p* might be false:

(14) Phil believes the sun is square

In addition, the contents of the proposition, the entities in it, are under a certain description that is tightly tied to the holder of the belief, e.g.:

(15) Phil believes *his aunt's dog* is a nuisance.

Finally, they can be recursive:

(16) Phil believes his aunt thinks the dog is a nuisance.

Therefore, propositional attitudes have the properties of potential *intentional non-existence*, *falsity*, *opacity*, and *recursion*. Whatever medium of representation they are also needs to have those properties, and images and words – not to mention neural networks or embodiment – do not seem to fit the bill. This led Fodor (1975), among other philosophers, to propose the existence of a Language of Thought. Yet other philosophers resist the positing of yet another layer of cognitive representation. Some argue that *natural* language has the right properties for the representation of propositional attitudes and the reasoning that proceeds from it (Segal, 1998; Collins, 2000; Hutto, 2009).

4.2 Infant theory of mind

If this argument holds, then one might expect that a child has to reach a certain level of linguistic competence with complementation in order to succeed at reasoning about other's beliefs, and there are a number of supportive findings reviewed in Section 4.3 below. But before that review, we first discuss results troublesome for this story from several studies that find apparent false belief understanding in much younger children who do not know even rudimentary syntax (Onishi & Baillargeon, 2005; Southgate, Senju, & Csibra, 2007; Baillargeon, Scott & He, 2010; Southgate, 2013).

The theorists in this area of infant Theory of Mind (ToM) fall into two major camps, characterized by Scott, Richman and Baillargeon (2015) as "mentalists" versus "minimalists". The mentalists believe that the infant research proves that very young children, well before language is established, can represent the beliefs of others. In contrast the minimalists argue that one cannot draw such a firm conclusion because other simpler explanations might suffice. Almost all of these studies use the infant's looking time or direction of looking as the index of understanding, and the worry is always that what the infants are responding to may not be a person's false beliefs but instead an intention (Fenici, 2014) or a registration of an object (Butterfill & Apperly, 2013).

In their synthesis, Apperly and Butterfill (2009) argued that the character of the infant's understanding of mind might have a fundamental signature – some limitation – that differentiates it from that of 4-year-olds and adults. By analogy, they point out that infants can do elementary arithmetic, but the signature that betrays the nature of their understanding is that they are only successful with numbers less than four (Carey, 2004). Low and Watts (2013) contend that such a signature in the Theory of Mind domain might be that infants cannot represent the *contents* of false beliefs, perhaps just appreciate someone's intention to act towards a location. Butterfill and Apperly concur (2013) that it is identity that is key, not just location, perhaps because object identity means the person is *seeing as*. Philosophers of science (Hanson, 1958) made interesting distinctions between plain *seeing*, versus *seeing* an object *as* something, versus *seeing that* something is the case. Interestingly, only the latter is the level at which propositions can be asserted.

That is why recent papers on infant's appreciation of deceptive appearance are important in challenging this synthesis. These recent empirical findings promise to stretch our understanding even further, as they show that infants can recognize the other's "registration" of an object as having a deceptive appearance (Scott et al., 2015).

Southgate (2013) makes the very interesting claim that infants may succeed in reading the other's intent – perhaps even their belief state – precisely because they do not consider the alternative point of view, namely their own! In other words, infants may be especially attentive to the attention that others pay towards an object and its location, and follow it keenly, ignoring what they themselves know. As this diminishes over time, they begin giving attention primarily to their own knowledge access. Only at around age four can children juxtapose the two representations and make a reasoned choice between them, allowing them to pass the classic false belief tasks. The time between these achievements entails many developments: in attention, in social understanding, in executive function, and of course in language: in vocabulary, pragmatics, semantics and syntax.

Space limitations, the fast moving empirical and theoretical literature on the infant ToM results, and the focus of this volume, require me to lay the issue aside, and make the assumption that the 4-year-old who passes a false belief reasoning task is doing something explicit with beliefs, and whatever that something is, it seems to entail the use of language. It is to that issue that we turn next.

4.3 When do complements emerge?

Rudimentary embedded forms such as small clauses, lacking tense or even a verb, emerge early, when the child is around 2. Non-finite forms come in early, sometimes around the same time with certain common verbs like *wanna* or *hafta* in invariant form, hinting that these might serve as auxiliaries rather than matrix verbs (Brown, 1973). The point at which the nature of the embedding becomes clearer is when the sentence has a second subject, such as *he wants her to do it* is, but those do not typically come in until age 3–4 years. Tensed complements also emerge at 3 or 4, and the complementizer (e.g. *that*) is frequently absent (Bloom, Rispoli, Gartner, & Hafitz, 1989). For a fuller account see de Villiers & Roeper (2016).

Semantically, it has been claimed that the first tensed complements may not express the full range of meanings in adult English. For example they are often first person, "I think.", and they do not express false beliefs, but opinions. For those reasons some writers argue that some of the earliest complements are idiomatic, thus not truly flexible, embedded, forms (e.g. Diessel & Tomasello, 2001). A thorough analysis of the English CHILDES data by Bartsch and Wellman (1995) suggested that the tensed complements under mental verbs emerge gradually, only achieving full status as forms that can carry the false beliefs of others by around 3;5 years.

It is difficult to elicit complements from young children, so in my own work comprehension has been the method to test embedding. This began serendipitously in a series of studies done with Tom Roeper, looking at the development of the syntax and semantics of *wh*-questions (e.g. de Villiers, Roeper, & Vainikka, 1990; Roeper & de Villiers, 1991). We designed some sentences to see if the child was appropriately interpreting long distance questions such as (17):

(17) How did he say he rode the horse?

We provided stories to contrast the long distance reading (how he rode the horse) with the short distance reading (how he said it), but the character in our early stories never mis-spoke, that is, he always correctly reported on his riding. Juan Uriagereka was the person who suggested in conversation that we make the lower clause false in a sentence such as:

(18) What did the mother say she bought?

in order to test whether the children were appropriately applying scope to the *wh*-word, namely, that both verbs need to be taken into account in answering. To our surprise, young children (below age four) were very prone to answer simply what the mother bought (de Villiers, 1999). The full account is developed in several papers (de Villiers & Pyers, 2002; de Villiers, 2005; Roeper & de Villiers, 2011), and for yet more varied opinions based on the child's problems with pragmatics, see Lewis, Hacquard & Lidz (2013) and Van Cleave & Gauker (2010).

Is this mistake because the children do not yet have a Theory of Mind? That is, perhaps they cannot yet understand lies and mistakes because they do not understand the point of view on truth of another individual, so they 'fix' it to their truth. We undertook several major studies to test this, and arrived at a most disturbing conclusion. The result was the other way around, namely, it is *after* children master these structures that they can pass classic false belief tasks (de Villiers & Pyers, 2002). The important parallel findings with language delayed deaf children (Schick, Hoffmeister, de Villiers & de Villiers, 2007; Pyers & Senghas, 2009) and children with autism (Tager Flusberg & Joseph, 2005) are often neglected in reviews. A meta-analysis of language and Theory of Mind concluded that the contribution of complement mastery to false belief understanding is a fairly robust finding, though studies were scarce (Milligan, Astington & Dack, 2007). Two training studies in English (Lohmann & Tomasello, 2003; Hale & Tager-Flusberg, 2003) showed that teaching complements of the right sort can pay off in passing false belief tasks, though teaching other complex structures such as relative clauses does not. There is supportive evidence from other languages (Aksu Koç, Avca, Aydin, Sefer, & Yasa, 2005; Lohmann & Tomasello, 2003; Perner, Zauner, & Sprung, 2005).

However, there are some failures to replicate the result that complements play a decisive role. In some studies, other indices of general syntax outweigh the contribution of a complement task as a predictor (Milligan et al., 2007; Ruffman et al., 2003; Cheung et al., 2004). Unfortunately many of the existing studies lack sufficient power to detect the contributions of different prerequisite skills. In a recent large longitudinal study, we have confirmed that complement mastery is a major *independent* predictor of passing false belief tasks (de Villiers, de Villiers, Lindley, & Chen, 2015), though as is found in other studies, vocabulary and general syntax also have roles to play.

Yet tensed complements – not infinitival complements – play the crucial role. Critically, it is *realis* complements, for which there is a truth value (de Villiers, 2005; de Villiers & de Villiers, 2009). The next section will begin to address that.

5. Point of view across clauses

It turns out that all perspective-taking elements are not alike in how they interact with syntactic embedding. Most, but not all, of the point-of-view elements take a speaker's perspective in ordinary matrix sentences, but then can switch to the matrix subject's perspective once embedded. A succession of tables is presented that shows how the PoV phenomena fall into at least three types when one considers how each one behaves in different types of clauses.

The rows in Table 1 constitute the type of sentence structures that are in question: simple matrix sentences / adjoined clauses, non-finite complements, finite

Table 1. Types of clause and perspective of Truth and PoV indices inside them

Type of clause	Example	Truth	Tense
Matrix or adjunct clause	*He threw out the food in* ***her*** *fridge over* ***there yesterday***	speaker	speaker
Non finite complement	She said *to throw out the food in* ***her*** *fridge over* ***there yesterday***	indeterminate	indeterminate
Finite complement	She said *he threw out the food in* ***her*** *fridge over* ***there yesterday***	subject	subject
Direct speech: nonfinite	She said *"Throw out the food in* ***my*** *fridge over* ***here today"***	indeterminate	indeterminate
Direct speech: finite	She said *"He is throwing out the food in* ***my*** *fridge over* ***here today****"*	subject	subject

complements, and direct speech under a verb. Then a representative sentence containing the PoV elements is given. The next column represents the perspective on truth of the clause: in whose world is it true? The speaker, the matrix subject, or is it indeterminate? The last columns represent the types of PoV phenomena under consideration, namely, a group from those considered earlier, such as tense, pronouns, spatial deixis, reference, and so forth. (Evidentials are excluded because their syntactic behavior, especially in regard to sentential complements, is still unclear). In each case the question is asked, whose perspective is represented by the element (boldface) in the particular clause (the one italicized) under consideration? From whose point of view is this taken, as indicated by the particular PoV elements? Each clause type is considered.

Take the sentence in (19) containing several different PoV indices. It is unwieldy, but necessarily so, to be able to show the contrasts in how the various elements behave:

(19) He threw out the food in her fridge over there yesterday.

(19) is then modified to reflect the various syntactic contexts, in the first column.

5.1 Truth and tense

First consider whose *truth* is represented by the clause: is it the matrix subject, the speaker, or is it indeterminate? The whole sentence in (19) expresses the truth asserted by the speaker. In the non-finite complement of sentence (20):

(20) She said *to throw out the food in her fridge over there yesterday*.

there can be no determination of truth: it refers to an irrealis event, we cannot determine whether the event in the clause – the throwing out – happened or not. In the tensed complement (21), as argued in Section 4.1, the truth is from the matrix subject's perspective:

(21) She said *he threw out the food in her fridge over there yesterday*

From our perspective, or the speaker's, (21) could report a lie, or a mistake, and so the event in the complement might never have happened. In the Direct speech case in Table 2 the truth parallels that for the indirect speech: indeterminate for the irrealis (imperative) case, and subject-oriented for the directly quoted tensed clause.

Now consider the point of view attached to the Tense on the verb in Table 1. Whose time perspective is it from? Clearly, in the matrix or adjunct clause, the tense is from the speaker's perspective. That is, the throwing out of food happened in the past with respect to the speaker, not the matrix subject. In the untensed complement, the tense is not determinable. When did the event happen, or did it even happen? It is certainly not connected to the speaker's time of utterance. But in the tensed complement variant in (21) the tense is the matrix subject's perspective. The *saying* event is the speaker's tense, but the *throwing* event is prior to that. If it was coincident, it would more likely be in the past progressive as in (22):

(22) She said he was throwing out the food in her fridge over there yesterday.

(22) shows the phenomenon of sequence of tense, wherein the two tenses agree when there is an embedded clause. Sequence of tense occurs in some but not all languages (Hollebrandse, 2000).

In Direct speech, this patterns similarly with the indirect complements, that is, indeterminate for the non-finite (imperative), and subject PoV for the tensed version.

Inspecting the two rows in Table 1, it is evident that Tense and Truth pattern identically. In addition, no difference occurs between direct and indirect speech reports. Importantly, the tensed variety explicitly provides a different perspective, namely that of the matrix subject.

5.2 Deixis

The second group of phenomena contains pronouns, spatial and temporal deixis. Whose PoV is "he" and "your"? Whose PoV are "there" and "yesterday"? (See Table 2). In matrix clauses and direct complements of either tensed variety, the speaker's PoV is imposed throughout. For indirect speech there is a switch, in which the subject's perspective is imposed (i.e. the subject who is the actual speaker). This group all pattern alike, and do not align with Truth and Tense.

Table 2. Types of clause and perspective of deictic PoV indices inside them

Type of clause	Example	Pronouns	Spatial deixis	Temporal deixis
Matrix or adjunct clause	*He threw out the food in **her** fridge over **there yesterday***	speaker	speaker	speaker
Non finite complement	She said *to throw out the food in **her** fridge over **there yesterday***	speaker	speaker	speaker
Finite complement	She said *he threw out the food in **her** fridge over **there yesterday***	speaker	speaker	speaker
Direct speech: nonfinite	She said *"Throw out the food in **my** fridge over **here today**"*	subject	subject	subject
Direct speech: finite	She said *"He is throwing out the food in **my** fridge over **here today**"*	subject	subject	subject

5.3 Reference and description

The third group requires fresh examples to avoid too much unnecessary clutter in the sentence. Sticking with the invidious fridge cleaner, we now focus attention on the objects he attacked to examine the phenomena of Personal Taste adjectives and Referential Opacity, or by what name a thing shall be called.

Table 3. Types of clause and perspective of personal taste adjectives indices inside them

Type of clause	Example	Personal Taste adjective
Matrix or adjunct clause	*He threw out **the yucky food***	speaker
Non finite complement	She said *to throw out **the yucky food***	indeterminate
Finite complement	She said *he threw out **the yucky food***	indeterminate
Direct speech: nonfinite	She said "*throw out **the yucky food***"	subject
Direct Speech: finite	She said "*he's throwing out **the yucky food***"	subject

Take (23) as the event for personal taste adjectives in Table 3:

(23) He threw out the yucky food.

According to recent analyses, "yucky" must be pragmatically linked to a salient judge in the context (Lasersohn, 2005). In this bare context, the most natural interpretation is a speaker interpretation: the food is yucky from the speaker's angle.

In the non-finite complement, however, there seem to be many alternatives available for a "judge", so it is marked indeterminate. The same is true for the tensed complement. Was it the speaker's, the subject's, or could it even be the embedded subject's, "his" view of the food? When we reach the direct speech cases, however, it becomes clear: the matrix subject considers the food yucky. I can also get a reading in which the instruction is:

(24) Throw out the food that you consider yucky.

but I suspect some domestic battles might result.

Finally, we reach the case of referential opacity. In this case, there must be two designations of the same object. Imagine that there is a bowl of onion dip in the fridge, but unbeknownst[1] to the fridge clearer, the onion dip is an award-winning appetizer of the speaker's, perhaps about to be transported to a cocktail party the next night. These scenarios are depicted in Table 4, and here attention must be paid to whether substitution of the coreferential term would preserve truth.

Table 4. Types of clause and transparency of referential substitution permitted in them

Type of clause	Example	Substitute designation for referent	Referential substitution allowed by:
Matrix or adjunct clause	*He threw out* ***the onion dip***	*He threw out* ***the award winning appetizer***	speaker
Non finite complement	She said *to throw out* ***the onion dip***	She said *to throw out* ***the award winning appetizer***	indeterminate
Finite complement	She said *he threw out* ***the onion dip***	She said *he threw out* ***the award wining appetizer***	indeterminate
Direct speech: nonfinite	She said "*throw out* ***the onion dip***"	She said "*throw out* ***the award winning appetizer***"	subject
Direct Speech: finite	She said "*he's throwing out* ***the onion dip***"	She said "*he's throwing out* ***the award winning appetizer***"	subject

In ordinary sentences, referential substitution works fine. The PoV on the noun phrase is the speaker's, as the speaker knows the referent *both* as onion dip and the award-winning appetizer. The difficulty comes with the embedded complements, where the classic problem arises of *de dicto* and *de re* (Quine, 1956). On a *de re* reading, the matrix subject may have said, "Throw out the onion dip", but in reporting it in indirect speech, the speaker can legitimately say (25):

1. This is the handiest word ever in Theory of Mind work

(25) I can't believe what happened! She said to throw out the award-winning appetizer!

That is, the speaker can substitute the words and still talk about the object from the speaker's PoV. On another reading, the *de dicto*, the speaker might be challenged in court if he claims that the woman called the object an award-winning appetizer, when in fact she just said:

(26) Throw out the onion dip!

Hence, PoV is ambiguous, or indeterminate in such a case, as we do not know whose PoV is intended. When we turn to direct speech, the PoV reverts to that of the matrix subject, the actual utterer of the words. In accurately reporting the speech, I must use the words the other speaker used. The last two cases then pattern together across environments, and differently in their behavior from the pronoun/deixis set, and unlike Tense and Truth.

5.4 Summary

In summation, three distinct PoV types occur in their behavior in sentence contexts (see summary Table 5). Hollebrandse (2000) made the proposal that there might be a PoV operator in the CP of a sentence that coordinates agreement across pronouns, spatial and temporal deixis. However that coordination is not so simple across different clauses.

Table 5. Three different classes of behavior of PoV types across clauses

Type of clause	Matrix or adjunct clause	Non finite complement	Finite complement	Direct speech: nonfinite	Direct Speech: finite
Truth	speaker	indeterminate	subject	indeterminate	subject
Tense	speaker	indeterminate	subject	indeterminate	subject
Pronouns	speaker	speaker	speaker	subject	subject
Spatial Deixis	speaker	speaker	speaker	subject	subject
Temporal Deixis	speaker	speaker	speaker	subject	subject
Personal Taste	speaker	indeterminate	indeterminate	subject	subject
NP designation	speaker	indeterminate	indeterminate	subject	subject

6. Implications

What are the implications for the larger story of how children could acquire perspective from conversational contexts? Embedded forms – whether direct or indirect – provide crucial information about Truth, and they do so primarily via Tense markers.

Verbs are subcategorized not only for whether they take a complement at all, but also for whether it is finite, non-finite or subjunctive. Finiteness is the domain in which truth or assertion seems to operate. Klein (1998) also discusses the notion of finiteness, or abstractly, FIN, as having two distinct meaning components. One is to mark tense, specifically whether the topic time precedes, contains, or follows the time of utterance; and the other, that an assertion is being made. Sentences without finiteness make no assertions (for rich elaboration, see Klein, 1998).

Hinzen (2013) also discusses naturalization of the concept of Truth. He argues persuasively that truth is a property that emerges internally, from the syntax of natural language, not from reference or external considerations. In particular, he argues that anything less than a clause cannot have a truth value, that is cannot be evaluated as true or false: noun phrases, small clauses, and infinitive complements all lack the necessary structure.

The obvious question arises about how specific this argument is to languages like English or German. What about languages like Mandarin, which lacks tense morphology? The debate over the existence of Tense in Mandarin is a vast topic. Two alternatives exist to accommodate such languages within the current story. One is to agree with those that claim that Mandarin has a Tense node, in line with claims about universal language, but lacks overt tense morphology (Sybesma, 2007). Sybesma argues that in many respects Mandarin behaves like Dutch, but lacks overt tense agreement. The second is to argue with those who claim that Aspect plays the role in Mandarin that Tense does in English (Lin, 2010). Importantly, Mandarin speakers can make the linguistic distinction between complements that are realis or irrealis (see also Lin, 2011).

Tom Roeper and I, elaborating on Klein, proposed that Tense (covertly) moves to the edge of the clause, carrying with it the point of view of the subject to be represented (Roeper & de Villiers, 2011). If the construction has no Tense, such as an infinitive, then it will also carry no point-of-view shift. We predicted significant differences in how children treat tensed and non-finite/irrealis clauses, and only the former should be linked to mastery of false beliefs. This is exactly what we found, when we tested children on scenarios containing discrepancies between what was said and what happened (de Villiers, Harrington, Gadilauskas, & Roeper, 2012). Children had relatively little difficulty answering questions posed with non-finite complements, such as that in (27):

(27) What did the boy say to buy?

versus finite complements (28):

(28) What did the boy say his Dad bought?

In both cases, the event of *buying* was other than the one ordered/described by the boy. The disparity between finite and non-finite complements is confirmed in a much larger sample (*N* = 674) of children aged 3 through 5 years who were given two of each kind as part of a new language assessment (Golinkoff, de Villiers, Hirsh-Pasek, Iglesias & Wilson, 2017). Children are much more competent with the non-finite than the finite complement, even when the scenarios are well matched, and the verb (say) is the same.

An interesting case arises with modals. The contrast was tested in de Villiers (2005) between (29–31):

(29) Mom thinks that Bella was playing on the computer.

(30) Mom thinks that Bella should play on the computer.

(31) Mom wants Bella to play on the computer.

These occurred in scenarios where Mom could not see Bella, and Bella was doing something else, say painting. Responses to the modal case in (29) were found to pattern exactly like the infinitival case with *want* in (31). Three- to five-year-olds found it easy to judge them true, unlike the case in (30) of *think that*. The case of modals in intentional contexts still requires further theoretical and empirical exploration.

7. Is linguistic encoding required?

As a final point, consider again whether direct speech or indirect speech encoding matters. Table 5 would suggest not, in that both sentence types are linked to truth in the same way. However, here we are talking about fully encoded direct and indirect speech complements. What about witnessing speech acts directly? Is that sufficient for children to learn that others have beliefs different than their own, or do children learn by hearing the description of the speaker and the speech act encoded?

Consider these contrasts in how it might be presented, in (32–34). The quotation marks indicate what the child hears, and the remaining parts of the scenarios are not expressed, but witnessed directly.

(32) Mom: (witnessed saying) “Dad is outside”
Child sees Dad upstairs.
“What did Mom say?”

In (32), the person speaking is not linguistically encoded. The contradiction is also not marked overtly in language, therefore the participant must encode the discrepancy between what Mom says about Dad, and what the child sees about his location.

(33) "Mom said 'Dad is outside.'"
Child sees Dad upstairs.
"What did Mom say?"

In (33), who the speaker is gets directly encoded, and a direct quotation is used rather than just a speech act being witnessed. However, the participant must still encode the discrepancy between Mom's statement about Dad and what the child sees.

(34) "Mom said Dad is outside.
But look, Dad is upstairs.
What did Mom say?"

In (34), everything is overtly encoded: Mom is identified as the speaker, and the content of her speech is formed as indirect speech rather than a quotation. In addition, the discrepant fact that Dad is in a different location is expressed in speech, rather than leaving the inference to the participant to encode.

The type (34) is what we have traditionally used in the complement comprehension task (de Villiers & Pyers, 2002). But in designing a new language assessment (Golinkoff et al., 2017), we had the opportunity to use animated events with cartoon speakers, and therefore to drop the narration. To check that this would be equivalent, we piloted type (34) (events only with speakers) and (33) (linguistic encoding and indirect report) on a sample of children ($N = 56$) aged 3–4 to see if the difference in scenarios mattered (See Figure 1).

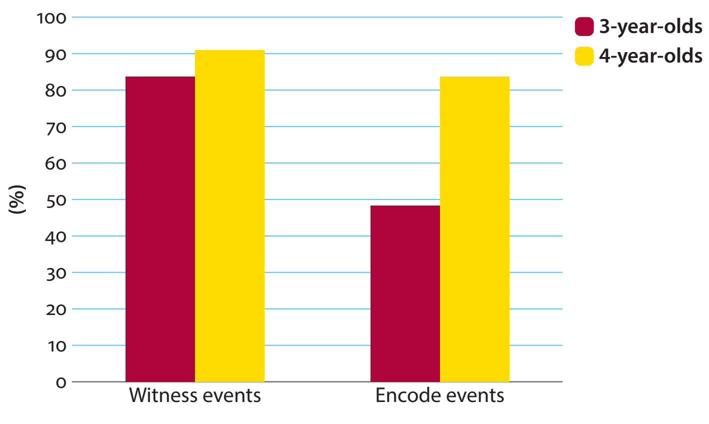


Figure 1.

To our surprise, young children had much *less* difficulty answering the question if they had NOT heard it encoded in an indirect report (type 35). It is not clear that they even noticed any discrepancy; they just answered what the Mom said. It is my strong suspicion that success on this version of the task would pattern very poorly with False Belief tasks, because the children did not engage in any comparison of two representations. Children may succeed on scenario type (32) in the same way that infants succeed in the eye-gaze studies (Southgate, 2013), because *attributing* an intention or an utterance to someone does not necessarily engage with **truth** without an additional step of comparison of representations.

8. Conclusion

Writers such as Harris would argue that discrepancies in conversation reveal point of view/perspective and hence lead to development of Theory of Mind (Harris, 2005). I have made two arguments against the sufficiency of this claim. First, I demonstrate that not all linguistic elements that mark point-of-view indices are alike, by showing how each behaves in different clauses. There are several interesting subtypes that pattern together, and the tense or finiteness of a clause seem especially significant in its connection to truth, or assertion. In addition I argue that observation of speech is not enough, and suggest that *linguistically encoded speech reports* are needed to help the child along towards the contrast in representations necessary for explicit, rather than implicit, false belief reasoning. That is why mastering tensed complements is implicated as a predictor of that cognitive achievement.

References

Aikhenvald, A. Y. (2004). *Evidentiality*. Oxford: Oxford University Press.

Aksu-Koc, A., Avci, G., Aydin, C., Sefer, N., & Yasa, Y. (2005). The relation between mental verbs and ToM performance: Evidence from Turkish children. Paper presented at the IASCL Convention, Berlin, July.

Apperly, I., & Butterfill, S. (2009). Do humans have two systems to track beliefs and belief-like states? *Psychological Review*, 20, 521–536.

Apperly, I., & Robinson, E. J. (1998). Children's mental representation of referential relations. *Cognition*, 67, 287–309. https://doi.org/10.1016/S0010-0277(98)00030-4

Baillargeon, R., Scott, R., & He, Z. (2010). False-belief understanding in infants. *Trends in Cognitive Science*, 14, 110–118. https://doi.org/10.1016/j.tics.2009.12.006

Bartsch, K., & Wellman, H. (1995). *Children talk about the mind*. Oxford: Oxford University Press.

Berman, R., & Slobin, D. I. (1994). *Relating events in narrative: A cross-linguistic developmental study*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Bloom, L., Rispoli, M., Gartner, B., & Hafitz, J. (1989). Acquisition of complementation. *Journal of Child Language*, 16, 101–120. https://doi.org/10.1017/S0305000900013465

Brentano, F. (1874). *Psychologie vom empirischen Standpunkt*. Leipzig: Duncker & Humblot. Online: <http://www.archive.org/details/psychologievome02brengoog>

Brown, R. (1973). *A first language: The early stages*. Cambridge, MA: Harvard University Press. https://doi.org/10.4159/harvard.9780674732469

Butterfill, S. A., & Apperly, I. A. (2013). How to construct a minimal Theory of Mind. *Mind & Language*, 28, 606–637. https://doi.org/10.1111/mila.12036

Carey, S. (2004). Bootstrapping and the origin of concepts. *Daedalus*, 133, 59–68. https://doi.org/10.1162/001152604772746701

Carruthers, P. (1996). Simulation and self-knowledge: A defence of theory-theory. In P. Carruthers & P. K. Smith (Eds.), *Theories of theories of mind*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511597985.004

Cheung, H., Hsuan-Chih, C., Creed, C. Ng, L., Ping Wang, S. & Mo, L. (2004). Relative roles of general and complementation language in theory-of-mind development: Evidence from Cantonese and English. *Child Development*, 75, 1155–1170. https://doi.org/10.1111/j.1467-8624.2004.00731.x

Clark, E. V., & Sengul, C. J. (1978). Strategies in the acquisition of deixis. *Journal of Child Language*, 5, 457–475. https://doi.org/10.1017/S0305000900002099

Collins, J. (2000). Theory of Mind, Logical Form and eliminativism. *Philosophical Psychology*, 13, 465–490. https://doi.org/10.1080/09515080020007616

Dale, P., & Crain-Thoreson, C. (1993). Pronoun reversals: Who, when, and why? *Journal of Child Language*, 20, 573–589. https://doi.org/10.1017/S0305000900008485

de Villiers, J. G. (1999). On acquiring the structural representations for false complements. In B. Hollebrandse (Ed.), *New perspectives on language acquisition* (pp. 191–236). Amherst, MA: GLSA.

de Villiers, J. G. (2005). Can language acquisition give children a point of view? In J. W. Astington & J. Baird (Eds.), *Why language matters for Theory of Mind*. (pp. 189–219). Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195159912.003.0010

de Villiers, J. G. (2007) The interface of language and Theory of Mind. *Lingua*, 117, 1858–1878. https://doi.org/10.1016/j.lingua.2006.11.006

de Villiers, P. A., & de Villiers, J. G. (1974). On this, that and the other: Non-egocentrism in very young children. *Journal of Experimental Child Psychology*, 18, 438–447. https://doi.org/10.1016/0022-0965(74)90122-2

de Villiers, J. G., & de Villiers, P. A. (2009). Complements enable representation of the contents of false beliefs: Evolution of a theory of Theory of Mind. In S. Foster-Cohen (Ed.), *Language acquisition* (pp. 169–195). Houndmills: Palgrave Macmillan.

de Villiers, P. A., de Villiers, J. G., Lindley, E., & Chen, M. (2015). Language, inhibitory control, and explicit false belief understanding: A longitudinal structural equation model. Poster presentation at the Society for Research in Child Development Biennial Convention, Philadelphia, PA.

de Villiers, J. G., & Garfield, J. (2017). Evidentiality, questions and the reflection principle in Tibetan: What do children learn when they learn about evidentiality? In F. N. Ketrez, A. C. Küntay, Ş. Özçalışkan, & A. Özyürek (Eds.), *Social environment and cognition in language development: Studies in honor of Ayhan Aksu-Koç* (pp. 113–132). Amsterdam: John Benjamins. https://doi.org/10.1075/tilar.21.08vil de Villiers, J. G., Garfield, J., Gernet Girard, H., Roeper, T., & Speas, P. (2009). Evidentials in Tibetan: Acquisition, semantics and cognitive development. In S. Fitneva & Matsui, T. (Eds.), *Evidentiality: A window into language and development*. Special issue of New Directions for Adolescent and Child Development 125: 29–48.

de Villiers, J. G., Roeper, T., Harrington, E., & Gadilauskas, E. (2012). Tense and Truth in children's question answering. In A. Biller, E. Chung, & A. Kimball (Eds.), *Proceedings of the 36th annual Boston University Conference on Language Development* (pp. 152–163). Somerville, MA: Cascadilla Press.

de Villiers, J. G., Nordmeyer, A., & Kravitz, M. (2010). Maintaining a point of view across multiple pronoun switches. Poster presented at Generative Approaches to Language Acquisition–North America, Toronto.

de Villiers, J. G., & Pyers, J. (2002). Complements to cognition: A longitudinal study of the relationship between complex syntax and false belief understanding. *Cognitive Development*, 17, 1037–1060. https://doi.org/10.1016/S0885-2014(02)00073-4

de Villiers, J. G., & Roeper, T. (2016). The acquisition of complements. In J. Lidz, W. Snyder, & J. Pater (Eds.), *Oxford handbook of developmental linguistics* (pp. 279–309). Oxford: Oxford University press.

de Villiers, J. G., Roeper, T., & Vainikka, A. (1990). The acquisition of long distance rules. In L. Frazier & J. G. de Villiers (Eds.), *Language processing and acquisition* (pp. 257–297). Dordrecht: Kluwer. https://doi.org/10.1007/978-94-011-3808-6_10

Diessel, H., & Tomasello. M. (2001) The acquisition of finite complement clauses in English: A corpus-based analysis. *Cognitive Linguistics*, 12, 1–45. https://doi.org/10.1515/cogl.12.2.97

Dunn, J., & Brophy, M. (2005). Communication, relationships, and individual differences in children–s understanding of mind. In J. W. Astington & J. A. Baird (Eds.), *Why language matters for Theory of Mind* (pp. 50–69). Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195159912.003.0003

Fay, W. (1979). Personal pronouns and the autistic child. *Journal of Autism and Developmental Disorders*, 9, 247–260. https://doi.org/10.1007/BF01531739

Fenici, M. (2014) A simple explanation of apparent early mindreading: Infants' sensitivity to goals and gaze direction. *Phenomenology in Cognitive Science*, 14, 497–515. https://doi.org/10.1007/s11097-014-9345-3

Fillmore, C. (1975). *Santa Cruz lectures on deixis, 1971*. Bloomington, IN: Indiana University Linguistics Club.

Fillmore, C. J. (1971/1997). Coming and going. *Lectures on deixis*. Distributed by the Indiana Linguistics Club. Available online: <http://www-personal.umich.edu/~jlawler/FillmoreDeixisLectures.pdf>

Fodor, J. A. (1975). *The language of thought*. Cambridge, MA: Harvard University Press.

Frege, G. (1948). Sense and reference. *The Philosophical Review*, 57, 209–230. https://doi.org/10.2307/2181485

Garrett, E. (2001). Evidentiality and assertion in Tibetan (Unpublished doctoral dissertation). University of California, Los Angeles, CA.

Girouard, P., Ricard, M., & Decarie, T. G. (1997). The acquisition of personal pronouns in French-speaking and English-speaking children. *Journal of Child Language*, 24, 311–326. https://doi.org/10.1017/S030500099700305X

Golinkoff, R. M., de Villiers, J., Hirsh-Pasek, K., Iglesias, A., & Wilson, M. S. (2017). User's manual for the Quick Interactive Language Screener™ (QUILS™): A measure of vocabulary, syntax, and language acquisition skills in young children. Baltimore: Brookes Publishing Co.

Gordon, R. M. (1986). Folk psychology as simulation. *Mind and Language*, 1, 158–171. https://doi.org/10.1111/j.1468-0017.1986.tb00324.x

Gu, C. C., & Roeper, T. (2011) Acquisition arguments for a general point of view as an alternative to relativism and subjectivity (Unpublished manuscript). University of Massachusetts, Amherst.

Hale, C. M., & Tager-Flusberg, H. (2003). The influence of language on Theory of Mind: A training study. *Developmental Science*, 6, 346–359. https://doi.org/10.1111/1467-7687.00289

Hanson, N. R. (1958). *Patterns of discovery: An inquiry into the conceptual foundations of science*. Cambridge: Cambridge University Press.

Harris, P. (2005). Conversation, pretense, and Theory of Mind. In J. W. Astington & J. A. Baird (Eds.), *Why language matters for Theory of Mind* (pp. 70–83). Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195159912.003.0004

Hendriks, P., & Spenader, J. (2005/2006). When production precedes comprehension: An optimization approach to the acquisition of pronouns. *Language Acquisition*, 13, 319–348. https://doi.org/10.1207/s15327817la1304_3

Hinzen, W. (2013). Naturalizing pluralism about truth. In C. Wright & N. Pedersen (Eds.), *Truth and pluralism* (pp. 213–237). Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195387469.003.0011

Hollebrandse, B. (2000). The acquisition of sequence of tense (Unpublished doctoral dissertation). University of Massachusetts, Amherst.

van Hout, A., Harrigan, K., & de Villiers, J. G. (2010). Asymmetries in the acquisition of definite and indefinite noun phrases. In P. Hendriks & C. Koster (Eds.), *Asymmetries in child language*. Special issue of Lingua, 120, 1973–1990.

Hutto, D. D. (2007). *Folk psychological narratives: The social basis of understanding reasons*. Cambridge, MA: The MIT Press.

Kamawar, D., & Olson, D. R. (1999). Children's representational theory of language: The problem of opaque contexts. *Cognitive Development*, 14, 531–548. https://doi.org/10.1016/S0885-2014(99)00018-0

Karmiloff-Smith, A. (1981). *A functional approach to child language: A study of determiners and reference*. Cambridge: Cambridge University Press.

Klein, W. (1998). Assertion and finiteness. In N. Dittmar & Z. Penner (Eds.), *Issues in the theory of language acquisition: Essays in honor of Jürgen Weissenborn* (pp. 225–245). Bern: Peter Lang.

Lasersohn, P. (2005). Context dependence, disagreement, and predicates of personal taste. *Linguistics and Philosophy*, 28, 643–686. https://doi.org/10.1007/s10988-005-0596-x

Lewis, S., Hacquard, V., & Lidz, J. (2013). The semantics and pragmatics of belief reports in preschoolers. In A. Chereches (Ed.) *Proceedings of Semantics and Linguistic Theory 22* (pp. 247–267). Ithaca, NY: CLC Publications.

Lin, J. W. (2010). A tenseless analysis of mandarin Chinese revisited: A response to Sybesma 2007. *Linguistic Inquiry*, 41(2), 305–329. https://doi.org/10.1162/ling.2010.41.2.305

Lin, T-H. J. (2011). Finiteness of clauses and raising of arguments in Mandarin Chinese. *Syntax*, 14, 48–73. https://doi.org/10.1111/j.1467-9612.2010.00145.x

Lohmann, H., & Tomasello, M. (2003). The role of language in the development of false belief understanding: A training study. *Child Development*, 74, 1130–1144. https://doi.org/10.1111/1467-8624.00597

Low, J., & Watts, J. (2013). Attributing false beliefs about object identity reveals a signature blind spot in humans' efficient mind-reading system. *Psychological Science*, 24, 30–311. https://doi.org/10.1177/0956797612451469

Milligan, K., Astington, J., & Dack, L. (2007). Language and Theory of Mind: Meta-analysis of the relation between language ability and false-belief understanding. *Child Development*, 79, 622–646. https://doi.org/10.1111/j.1467-8624.2007.01018.x

Nelson, K. (2005). Language pathways into the community of minds. In J. Astington & J. Baird (Eds.), *Why language matters for Theory of Mind* (pp. 26–49). Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195159912.003.0002

O'Connor, A. Burgin, C., de Villiers, J. G., Speas, P., & Roeper, T. (2007). How children coordinate multiple deictic terms in sentences. Poster presented at the Society for Research in Child Development, Boston, April.

Onishi, K. H., & Baillargeon, R. (2005). Do 15-month-old infants understand false beliefs? *Science*, 308, 255–258. https://doi.org/10.1126/science.1107621

Oshima-Takane, Y. (1988). Children learn from speech not addressed to them: The case of personal pronouns. *Journal of Child language*, 15, 95–108. https://doi.org/10.1017/S0305000900012071

Oshima-Takane, Y. (1999). *The learning of first and second person pronouns in English*. Cambridge, MA: The MIT Press.

Oshima-Takane, Y., Takane, Y., & Shultz, T. R. (1999). The learning of first and second person pronouns in English: Network models and analysis. *Journal of Child Language*, 26, 545–575. https://doi.org/10.1017/S0305000999003906

Pearson, H. (2015). Attitude verbs (Unpublished manuscript). Queen Mary University of London. Available online at semanticsarchive.net

Perner, J., Zauner, P., & Sprung, M. (2005). What does 'that' have to do with point of view? The case of conflicting desires and 'want' in German. In J. W. Astington & J. Baird (Eds.), *Why language matters for Theory of Mind* (pp. 220–244). Oxford: Oxford University Press.

Potts, C. (2007). The expressive dimension. *Theoretical Linguistics*, 33, 165–197. https://doi.org/10.1515/TL.2007.011

Pyers, J. E., & Senghas, A. (2009). Language promotes false-belief understanding evidence from learners of a new sign language. *Psychological Science*, 20, 805–812. https://doi.org/10.1111/j.1467-9280.2009.02377.x

Quine, W. (1956). Quantifiers and propositional attitudes. *Journal of Philosophy*, 53, 177–187. https://doi.org/10.2307/2022451

Richards, B. (1976). Adverbs: From a logical point of view. *Synthese*, 32, 329–372. https://doi.org/10.1007/BF00484782

Roeper, T. (2007). *The prism of grammar*. Cambridge, MA: The MIT Press.

Roeper, T., & de Villiers, J. G. (1991). Ordered decisions in the acquisition of wh-questions. In H. Goodluck, J. Weissenborn, & T. Roeper (Eds.), *Theoretical issues in language development* (pp. 191–236). Hillsdale, NJ: Lawrence Erlbaum Associates.

Roeper, T., & de Villiers, J. G. (2011). The acquisition path for wh-questions. In J. G. de Villiers & T. Roeper (Eds.), *Handbook of generative approaches to language acquisition* (pp. 186–246). Dordrecht: Springer. https://doi.org/10.1007/978-94-007-1688-9_6

Ruffman, T., Slade, L., Rowlandson, K., Rumsey, C., & Garnham, A. (2003). How language relates to belief, desire and emotion understanding. *Cognitive Development*, 18, 139–158. https://doi.org/10.1016/S0885-2014(03)00002-9

Russell, J. (1987). "Can we say…?" Children's understanding of intensionality. *Cognition*, 25, 289–308. https://doi.org/10.1016/S0010-0277(87)80007-0

Schick, B., de Villiers, P. A., de Villiers J. G., & Hoffmeister, R. (2007). Language and false belief reasoning; A study of deaf children. *Child Development*, 78, 376–398.

https://doi.org/10.1111/j.1467-8624.2007.01004.x

Scott, R. M., Richman, J. C., & Baillargeon, R. (2015). Infants understand deceptive intentions to implant false beliefs about identity: New evidence for early mentalistic reasoning. *Cognitive Psychology*, 82, 32–56. https://doi.org/10.1016/j.cogpsych.2015.08.003

Segal, G. (1998). Representing representations. In P. Carruthers & J. Boucher (Eds.), *Language and thought* (pp. 146–161). Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511597909.010

Southgate, V., Senju A., & Csibra, G. (2007). Action anticipation through attribution of false belief by 2-year-olds. *Psychological Science*, 18, 587–92. https://doi.org/10.1111/j.1467-9280.2007.01944.x

Southgate, V. (2013). Early manifestations of mind reading. In S. Baron-Cohen & H. Tager-Flusberg (Eds.), *Understanding other minds: Perspectives from developmental social neuroscience* (pp. 3–18). Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199692972.003.0001

Speas, M. (2004). Evidentiality, logophoricity and the syntactic representation of pragmatic features. *Lingua*, 114, 255–276. https://doi.org/10.1016/S0024-3841(03)00030-5

Stephenson, T. (2007). Towards a theory of subjective meaning (Unpublished doctoral dissertation). MIT.

Sybesma, R. (2007). Whether we Tense-agree overtly or not. *Linguistic Inquiry*, 38, 580–587. https://doi.org/10.1162/ling.2007.38.3.580

Tager-Flusberg, H., & Joseph, R. (2005). How language facilitates the acquisition of false belief in children with autism. In J. W. Astington & J. A. Baird (Eds.), *Why language matters for Theory of Mind* (pp. 298–318). Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195159912.003.0014

Tanz, C. (1983). *Studies in the acquisition of deictic terms*. Cambridge: Cambridge University Press.

Thomas, E. (2010). *Pronoun differentiation in young children: An eyetracker study* (BA honors thesis). Smith College.

Van Cleave, M., & Gauker, C. (2010). Linguistic practice and false belief tasks. *Mind & Language*, 25, 298–328. https://doi.org/10.1111/j.1468-0017.2010.01391.x

SECTION IV

Logical interpretations

CHAPTER 11

The meaning of question words in statements in child Mandarin

Stephen Crain and Peng Zhou
Macquarie University / Tsinghua University

This paper reports the findings of three experiments investigating children's emerging knowledge of the semantics of information-seeking questions and declarative statements in Mandarin Chinese. Mandarin is particularly revealing about the semantic relationship between questions and statements, because it is a *wh-in-situ* language where question words appear in the same surface syntactic position in both questions and statements. The non-interrogative meanings that Mandarin-speaking children assign to question words lend weight to a unified approach to the semantics of existential expressions, including Free Choice Expressions, Negative Polarity Items, and disjunction words. When Mandarin question words appear in statements, children interpret them to be the semantic equivalents of existential expressions. The findings, therefore, support the unified approach to the semantics of existential expressions.

Keywords: questions, statements, existential expressions, English, Mandarin Chinese, child language acquisition

1. Introduction

Across languages, there are three general patterns of information-seeking questions and declarative statements. In the pattern represented by English, information-seeking questions begin with a *wh*-word: *What did Ted order*? *Who ordered a Holden*? These *wh*-questions presuppose the truth of statements with the existential expressions *something* and *someone*: *Ted ordered something*. *Someone ordered a Holden*. The semantics of the existential expressions *something* and *someone* invokes the existential quantifier from classical logic, ∃. In addition to *someone* and *something*, English existential expressions include *anything* and *anyone*, as well as the disjunction word *or*. The expressions *anything* and *anyone* are licensed by a class of linguistic operators, designated by Op in the graphic representations in (1)–(3).

https://doi.org/10.1075/tilar.24.11ste

The class Op includes Downward Entailing operators. The negative quantificational expression *nobody* is an operator that licenses the existential expression *any*. This explains why *Nobody ordered any dessert* is acceptable, but *Everybody ordered any dessert* is anomalous. The pattern of questions and statements for this class of languages is graphically depicted in (1).

(1) Questions: *wh*-words ______________________________
Statements: __________Op_________ ∃-expressions___

Another pattern characterises languages like Mandarin Chinese. In these languages, question words are positioned *in-situ* both in information-seeking questions and in declarative statements. In statements, question words are assigned meanings corresponding to English *anything* and *anyone*. In this class of languages the presence or absence of a linguistic operator determines the speech act being performed. In the absence of a licensor, the utterance is an information-seeking question; with a licensing operator, the utterance is a statement. Mandarin question words and the existential expression *renhe* 'any' are also licensed in declarative statements with the quantificational adverb *dou* 'all,' just as long as the question word precedes *dou*. If a question word follows *dou*, the utterance is an information-seeking question. The pattern for languages like Mandarin is depicted in (2).

(2) Questions: ______________________ *wh*-words ____________________
Statements1: __________Op _______+*wh* ∃-expressions _____________
Statements2: ________ + *wh* ∃-expressions _____________*dou* _________

A third pattern is characteristic of another class of languages, including several Australian languages (e.g., Diyari, Martuthunia, Panyjima). In these languages, question words appear both in information-seeking questions and in declarative statements, as in Mandarin. However, question words are 'fronted' in information-seeking questions, as in English. Like Mandarin, question words appear *in-situ* in statements, with meanings corresponding to the English expressions *anything* and *anyone*. These expressions are licensed by the same kinds of operators that license English *anything* and *anyone*. So, question words can be represented as +*wh* ∃-expressions. This third pattern of questions and statements is depicted in (3).

(3) Questions: *wh*-words __________________________________
Statements: __________Op _______ + *wh* ∃-expressions ___

It has recently been proposed that question words, existential expressions, and disjunction words form a natural class (Chierchia, 2013, 2017). On this unified account, the basic meaning of the class of expressions is an existential interpretation, so Chierchia refers to the class as ∃-items. It remains to explain why question words, existential expressions, and disjunction words license conjunctive/universal inferences when they appear with certain operators. According to the unified account, the operators that license conjunctive/universal inferences are the same across languages. For example, these inferences are licensed when ∃-expressions appear in sentences with modal verbs (e.g., English *can*, Mandarin *keyi* 'is allowed to'). In Mandarin, conjunctive/universal inferences for question words also require the quantificational adverb *dou* 'all.'

The present paper argues that the unified approach proposed by Chierchia (2013, 2017) explains the non-interrogative interpretations that children assign to question words in Mandarin. The findings of three experiments support this conclusion. Before we discuss Mandarin, however, it will be useful to review the basic syntactic and semantic properties of English question words, and how English question words are syntactically and semantically related to the existential expressions *any* and *some*, and to the disjunction word *or*.

2. Non-interrogative uses of question words in English

In English, some interrogative expressions (e.g., *what, whatever*) are assigned non-interrogative meanings. For example, the *wh*-word *what* appears in a free relative in Example (4).

(4) Grover likes what's on the plate.

Sentence (4) conveys the meaning that Grover likes everything on the plate. Findings from the research literature on child language invite the conclusion that English-speaking children younger than 6–7 years fail to assign a universal interpretation to sentences like (4). In a situation in which Grover has informed child participants of his dislike for onions, 4- and 5-year-old children were found to accept sentence (4) as a description of a plate that contained both onions and doughnuts (Caponigro, Pearl, Brooks, & Barner, 2012; for related findings, see Munn, Miller, & Schmitt, 2006, and Modyanova & Wexler, 2008). The same child participants assigned adult-like universal interpretations in responding to sentences with plural definite descriptions such as (5). Perhaps not incidentally, free relatives are far less frequent than plural definite descriptions in the adult input to English-speaking children.

(5) Grover likes the things on the plate.

The English *wh*-words *whoever* and *whatever* are also acceptable in declarative sentences, but these expressions are even less frequent than free relatives are in the adult input. A search of the transcripts of spontaneous production data for Adam and Sarah (Brown corpus, CHILDES; Brown, 1973) uncovered only 4 utterances with *whatever* and only 2 with *whoever* in the adult input to Adam between the ages of 2;3–5;2 (55 files), and there were only 2 examples with *whatever* and only 1 example with *whoever* in the adult input to Sarah between the ages of 2;3–5;1 (139 files) (MacWhinney, 2000). In child English, then, *wh*-words like *what* and *whatever* play a limited role in declarative statements, as attested by their scarcity in the adult input to English-speaking children. Question words are primarily used in information-seeking questions.

3. *Wh*-questions are derived from existential structures

The observation that *wh*-questions and existential expressions are syntactically linked was a basic tenet of transformational grammar. For example, Chomsky (1964) introduced a transformational rule that derived information-seeking *wh*-questions from structures underlying declarative sentences with indefinite nonspecific NPs. According to this analysis, the indefinite noun phrase *some* (*one*, *thing*) was the source for *wh*-questions with *who* and *what*. For example, the structure associated with the declarative sentence in (6) was the source of the *wh*-question in (7).

(6) He has something of yours.

(7) What does he have of yours?

The information-seeking question (7) and the declarative sentence (6) were related by a transformational rule. This rule explained the extensive symmetries between declaratives and *wh*-questions, as well as the pair-wise anomalies. Here are three examples.

First, both of the possessive phrases *someone's book* and *whose book* are acceptable in English, but the possessive noun phrases *something's cover* or *what's cover* are both deviant. Nevertheless, the semantically related possessive noun phrase *its cover* is acceptable. These observations underscore the conclusion that *wh*-questions and declarative statements are syntactically related. A second phenomenon that illustrates this relationship is witnessed in Examples (8)–(10). These examples exhibit the opposite pattern than Examples (6) and (7). Both the declarative sentence (8) and the *wh*-question (9) are unacceptable, but the semantically related sentence in (10) is acceptable.

(8) ?He found someone of yours.

(9) ?Who did he find of yours?

(10) He found a friend of yours.

A third linguistic phenomenon was explained by the transformational approach. This is a 'semantic gap' that pertains to both indefinite NPs and *wh*-phrases. The gap is created because the expressions *something* and *what* are referentially restricted to inanimate objects, whereas the expressions *someone* and *who* are referentially restricted to humans. This leaves "no natural way to refer to an unspecified animal" (Chomsky, 1964, p. 40). The semantic gap is illustrated by the unnaturalness of the statement (11), which carries over to the *wh*-question in (12).

(11) I watched something eating its dinner. (referring to a cat)

(12) What did you watch eating its dinner? (the cat or the dog?)

4. *Wh*-questions are semantically equivalent to existential statements

The semantic correspondence between information-seeking questions and existential statements has long been a bedrock of formal semantics. According to Hamblin (1973, p. 48) "questions set up a choice-situation between a set of propositions, namely those propositions that count as answers to it." In another seminal paper, Karttunen (1977) states that "…for semantic reasons, we make *wh*-phrases equivalent to existentially quantified noun phrases. For example, *who* and *what* … will have the same translation as *someone* and *something*." More formally, the meaning of a question is the (exhaustive) set of its answers (or true answers).

In these deliberations, we will appeal to several logical equivalences. One equivalence is between an information-seeking question and a disjunctive statement. Consider the English *wh*-question (13), and the answer in (14).

(13) Question: What kind of car did Ted buy? Option set = {VW, MG, BMW}

(14) Answer: Ted bought a VW, an MG, or a BMW.

In answering the question in (13), the interlocutor selects from a set {x: Ted bought a car of kind x}. When the set of options is explicit, the answer to a *wh*-question can be exhaustively, though not informatively, answered using a disjunctive statement. The disjunctive statement answer names each member of the option set as one of the disjuncts. Suppose Ted changed his mind, and decided against buying a car. In that case, the question can be answered using a negative statement with the polarity sensitive expression *any*, as in (15). This answer, in turn, is logically equivalent to a disjunctive statement that names the individuals in the option set, as in (16). Finally, the disjunctive statement (16) is logically equivalent to a conjunction of negative statements, one for each disjunct, as in (17).

(15) Ted didn't buy any kind of car.

(16) Ted didn't buy a VW, an MG, or a BMW.

(17) Ted didn't buy a VW, Ted didn't buy an MG, and Ted didn't buy a BMW.

5. Two analyses of polarity sensitive expressions

There have been two main analyses of polarity sensitive expressions such as English *any*. On one, *any* is analysed as an existential quantifier. In negative sentences, *any* is interpreted in the scope of negation. On the other analysis, *any* is a universal quantifier taking scope over negation. An advocate of the universal quantifier analysis of *any* was Quine (1960). According to Quine, the meaning of sentence (18) can be paraphrased as (19) (cf. Dayal, 2004; Shimoyama, 2011).

(18) I didn't see any birds.

(19) Every bird is such that I didn't see it.

Although the universal quantifier account of *any* has considerable empirical coverage, several counter-examples cast doubt on its validity. First, the universal quantifier analysis of *any* results in the incorrect interpretation of sentences like (20) (adapted from Carlson, 1980).

(20) The king decided not to meet a man who has any weapons.

(21) For every weapon, the king decided not to meet a man who has that weapon.

If *any* is analysed as a universal quantifier with scope over negation, then (20) receives the interpretation in (21). This is not the interpretation that English speakers assign to (20). Rather, it means that the king decided against meeting a man who is armed with any weapon. This interpretation suggests that negation takes scope over *any*, so *any* is better analysed as an existential quantifier (cf. Kadmon & Landman, 1993; Chierchia 2013).

A second counter-example to the universal quantifier approach was offered by Huang and Crain (2014). To see the force of the counter-example, consider sentence (22). Notice that sentence (22) is amenable to the universal quantifier account, so (22) can be paraphrased as (23), with the universal quantifier, *every*, taking scope over negation. Alternatively, sentence (22) can be analysed as negation taking scope over an existential quantifier, as in (24).

(22) Jack is not sure about anything that Max said.

(23) Everything that Max said is such that John is not sure about it. [$\forall > \neg$]

(24) There isn't a single thing that Max said that John is sure about. [$\neg > \exists$]

The conditional statement in (25) contains the declarative sentence (22) in the antecedent clause. The antecedent clause of (25) has two interpretations. One interpretation is similar to the interpretation assigned to (22), as indicated in (26).

(25) If Jack is not sure about anything Max said, he should say so.

(26) If everything Max said is such that Jack is not sure about it, he should say so.

However, the antecedent clause of (25) has another interpretation. Two variations of this interpretation are offered in example sentences (27) and (28).

(27) If something / anything Max said is such that Jack is not sure about, he should say so.

(28) If Jack is not sure about something / anything Max said, he should say so.

As sentences (27) and (28) illustrate, the alternative interpretation of (25) can be rendered either by using the polarity sensitive expression *anything*, or by using an existential quantifier *something*. Regardless of their position with respect to negation, *anything* and *something* are interchangeable in the antecedent of a conditional. The fact that the polarity sensitive expression *anything* exerts the same quantificational force as the existential expression *something* argues against the universal quantifier analysis of *anything*.

Despite the fact that *something* and *anything* are in complementary distribution in ordinary affirmative sentences, these expressions are interchangeable in several linguistic environments. This suggests that *something* and *anything* could be allomorphs. The proposal that these expressions are allomorphs was made explicitly in Klima (1964) in response to a puzzle involving sentences with verb phrase ellipsis. The puzzle is illustrated in (29).

(29) Donald will never say anything true, but Jack will.
 a. Donald will never say anything true, but Jack will < say *anything true >
 b. Donald will never say anything true, but Jack will < say something true >

Sentence (29) appears to violate a putative constraint on verb phrase ellipsis, called the parallelism constraint (e.g., Chomsky & Lasnik, 1993; Chomsky, 1995; Fiengo & May, 1994; Lasnik, 1972). The parallelism constraint dictates that an elided verb phrase in the second clause of a co-ordinate structure must be structurally parallel and must have the same referent as the verb phrase in the preceding clause (Fox, 1998; cf. Thornton & Wexler, 1999). The parallelism constraint appears to be violated in (29) because the second conjunct lacks a Downward Entailing licensing expression for the polarity sensitive expression *anything*. Lacking a licensing operator, *anything* is not tolerated in the second (elided) conjunct, as shown in (29a). The existential indefinite *something* is required instead, as shown in (29b). However, this is an apparent violation of the parallelism constraint.

Another apparent violation of the parallelism constraint in sentences with verb phrase ellipsis is illustrated in (30). In this case, the wrong interpretation is generated if the existential indefinite *something* is copied from the first conjunct into the second conjunct, as part of the elided verb phrase (see 30a). Because the second conjunct contains negation (… *won't*), the polarity sensitive expression *something* would be forced to take scope over negation at the level of semantic interpretation, resulting in an unintended meaning. The intended interpretation is indicated in (30b), which includes *anything* in the second conjunct instead of *something*.

(30) Donald will sometimes say something true, but Jack won't.
  a. Donald will sometimes say something true, but Jack won't < say something true >
  b. Donald will sometimes say something true, but Jack won't < say anything true >

The parallelism constraint is rescued, according to Klima (1964), if the abstract logical operator, ∃, is the underlying form for both the existential expressions *something* and *anything*. On this analysis, *something* and *anything* are allomorphs of ∃.

The parallelism constraint is salvaged by introducing a three-step process for verb phrase ellipsis. At the first step, the operator ∃ is introduced into both conjuncts of the coordinate structure. At the second step, ∃ is elided from the second conjunct. Therefore, the parallelism constraint is satisfied at the second step in the process. The third step takes place after the verb phrase in the second conjunct has been elided. At the third step, an allomorph of the operator ∃ (either *something* or *anything*) is phonologically realized in the first conjunct. The expression *anything* is inserted if ∃ is in the scope of a Downward Entailing expression, as in (29). Otherwise, *something* is inserted, as in (30).

The analysis of *something* and *anything* as allomorphs is independently motivated (Crain 2012). *Something* and *anything* are interchangeable, with altering truth conditions, in linguistic environments in which the polarity sensitivity of *something* is neutralized. For example, the polarity sensitivity of *something* is neutralized in the predicate phrase of sentences with the focus adverb ONLY. Therefore, sentences (31) and (32) are judged to be true in similar circumstances.

(31) Only Jack understood anything Donald said.

(32) Only Jack understood something Donald said.

The focus adverb ONLY has two meaning components, a presupposition and an assertion (Horn 1969). The presupposition pertains to the element in focus, Jack. The presupposition cannot be expressed using *anything*, as indicated in (33a). The second meaning component is the assertion. The assertion pertains to individuals who are being contrasted with the element in focus. The assertion entails that everyone in the contrast set lacks the property being attributed to the element in focus. As indicated in (33b), the assertion cannot be expressed using the polarity sensitive expression *something*. Because *something* takes scopes over the negation, (33b) generates the wrong interpretation for (31) and (32); *anything* yields the correct interpretation of the assertion. Both of the sentences (31) and (32) assert that every individual being contrasted with Jack did not understand anything that Max said.

(33) a. Presupposition: Jack understood something / *anything Donald said.
b. Assertion: Everyone else (besides Jack) didn't understand *something / anything Donald said.

Sentences (31) and (32) are a further illustration of the fact that *something* and *anything* are allomorphs. Moreover, this example further undermines the universal quantifier account of *anything*, because *anything* is semantically interchangeable with *something*, not *everything*.

6. The conjunctive/universal interpretation of ∃-expressions

There is one more topic before we turn to an analysis of question words in Mandarin Chinese. This is the conjunctive/universal interpretation of existential expressions. On the unified account proposed by Chierchia (2013, 2017), the conjunctive/universal interpretation of existential expressions is activated by a licensing operator. One licensing operator is the deontic modal verb, *is allowed to*. Consider sentences (34) and (35). Both of these sentences yield a conjunctive/universal interpretation. For example, the disjunctive statement (34) licenses the conjunctive inference that Kung Fu Panda is allowed to push the green car, and Kung Fu Panda is allowed to push the orange car; Kung Fu Panda is free to choose which car to push.

(34) Kung Fu Panda is allowed to push the green car <u>or</u> the orange car.

(35) Kung Fu Panda is allowed to push <u>any</u> of the cars.

Conjunctive/universal inferences are generated by an algorithm that makes repeated calls on an exhaustivity function, which we will designate ONLY (cf. Chierchia, 2013; Fox, 2007). We will outline how the algorithm applies to sentence (34). To get the derivation started, we will render the deontic modal verb phrase *is allowed to* using the possibility operator, ◊. This enables us to symbolically represent the example sentence (34) as ◊[G or R], where 'G' stands for *Kung Fu Panda pushes the* **G**reen *car* and 'R' stands for *Kung Fu Panda pushes the* **R**ed *car*.

At the first stage in the algorithm, the exhaustivity function ONLY applies to the subdomain alternatives to the original assertion ◊[G or R]. The subdomain alternatives are statements that the speaker might have made, but did not make. The relevant subdomain alternatives are ◊G, and ◊R. The recursive exhaustivity algorithm generates the inferences that follow from each of these subdomain alternatives. For alternative ◊G, the exhaustivity function ONLY(◊G) licenses the inference [◊G and ¬◊R]. For alternative ◊R, the exhaustivity function ONLY(◊R) licenses the inference [◊R and ¬◊G].

Three propositions are now under consideration, the original assertion ◊[G or R] and the two 'enhanced' alternatives: [◊G and ¬◊R] and [◊R and ¬◊G]. At this point, the exhaustivity function ONLY applies a second time. The second application of ONLY disposes of alternatives that are informationally stronger than the original assertion. Logically, both of the enhanced alternatives, [$\Diamond$G and ¬$\Diamond$R] and [$\Diamond$R and ¬$\Diamond$G], are stronger than the original assertion $\Diamond$[G or R]. Consequently, the negations of these enhanced alternatives are entered into the derivation: ¬[$\Diamond$G and ¬$\Diamond$R] and ¬[$\Diamond$R and ¬$\Diamond$G]. Each of these are logically equivalent to a conditional. The negated alternative ¬[$\Diamond$G and ¬$\Diamond$R] is logically equivalent to [$\Diamond G \rightarrow \Diamond R$] and the negated alternative ¬[$\Diamond$R and ¬$\Diamond$G] is logically equivalent to [$\Diamond R \rightarrow \Diamond G$]. Taken together, the two conditionals yield a bi-conditional [$\Diamond G \leftrightarrow \Diamond R$]. Two propositions remain: the original assertion $\Diamond$[G or R] and the bi-conditional [$\Diamond G \leftrightarrow \Diamond R$]. The original assertion $\Diamond$[G or R] entails that at least one of $\Diamond$G, $\Diamond$R is true, and the bi-conditional [$\Diamond G \leftrightarrow \Diamond R$] entails that the truth of either $\Diamond$G or $\Diamond$R guarantees the truth of the other. That is, at least one of $\Diamond$G, $\Diamond$R is true, and if one is true, then so is the other. It follows that $\Diamond$G and $\Diamond$R are both true.

7. Question words in Mandarin

This brings us to Mandarin. Mandarin question words such as *shei* 'who' and *shenme* 'what' commonly appear in non-interrogative sentences (e.g., Cheng, 1994, 1997; Li, 1992; Lin, 1998; Ladusaw, 1980). Mandarin-speaking children begin producing question words in declarative statements at 2 years of age (e.g., Huang & Crain, 2014; Li & Tang, 1991). Based on the theory proposed by Chierchia (2013, 2017), we have conducted several experimental studies that assess children's knowledge of the linguistic environments that license either the interrogative or non-interrogative meanings of question words. We have compared those linguistic environments with ones that license the Mandarin polarity sensitive expression *renhe* 'any' and we have compared the linguistic environments that license question words with those that permit the universal quantifier *mei* 'all.' We have also investigated the meaning that children and adults assign to sentences with the Mandarin disjunction word *huozhe* 'or.' The findings demonstrate that young Mandarin-speaking children understand both the interrogative and non-interrogative meanings of question words (see Crain, 2012).

For adult speakers of Mandarin, there is considerable overlap in the linguistic environments that license the non-interrogative use of Mandarin question words and those that license the polarity sensitive expression *renhe* 'any.' For example, both the question word *shenme* 'what' and the polarity sensitive expression *renhe* 'any' are licensed by the Downward Entailing negative quantifier *meiyouren* 'nobody.' This is illustrated in Examples (36) and (37). By contrast, the Mandarin universal quantifier *meigeren* 'everybody' is not Downward Entailing (on its external argument, the predicate phrase). If the universal quantificational expression *meigeren* 'everybody' replaces *meiyouren* 'nobody,' the sentence with *renhe* 'any' is ungrammatical, and the sentence with *shenme* 'what' becomes an information-seeking question.

(36) Meiyouren chi <u>shenme</u> shuiguo.
nobody eat what fruit
'Nobody ate any fruit.'
cf. Meigeren chi shenme shuiguo.
everybody eat what fruit
'What fruit did everybody eat?'

(37) Meiyouren chi <u>renhe</u> shuiguo.
nobody eat any fruit
'Nobody ate any fruit.'
cf. *Meigeren chi renhe shuiguo
everybody eat any fruit
'Everybody ate any fruit.'

8. The basic meaning of mandarin question words

Both distributional and interpretive facts from several linguistic structures indicate that the Mandarin question word *shenme* 'what,' the polarity sensitive expression *renhe* 'any,' and the disjunction word *huozhe* 'or' all generate the same interpretation. Consider a conversational context with three individuals: Aijun, Liao and Haihua. In this context, the three *ruguo*-conditionals in (38), (39) and (40) have the same interpretation. More specifically, if these conditional statements are produced by a puppet, Kermit the Frog, then both Mandarin-speaking adults and 4- to 5-year-old Mandarin-speaking children consistently reward Kermit with a coin in a situation in which only one person has eaten an apple (cf. Gualmini, Crain, & Meroni, 2000, who investigated the circumstances in which English-speaking children accept conditional statements).

(38) Ruguo <u>shenme</u> ren chi-le pinguo, ni jiu gei wo
if what person eat-ASP apple, you then give me
yi-ge yingbi.
one-CL coin
'If anyone ate an apple, then give me a coin.'

(39) Ruguo <u>renhe</u> ren chi-le pinguo, ni jiu gei wo
if any person eat-ASP apple, you then give me
yi-ge yingbi.
one-CL coin
'If anyone ate an apple, then give me a coin.'

(40) Ruguo Aijun, Liao huozhe Haihua chi-le pinguo, ni jiu gei
if Aijun Liao or Haihua eat-ASP apple, you then give
wo yi-ge yingbi.
me one-CL coin
'If Aijun, Liao or Haihua ate an apple, then give me a coin.'

Consider next the *ruguo*-conditionals in (41) and (42). In (41), the Mandarin universal quantifier *mei-ge-ren* 'every-CL person' has replaced either *shenme ren* 'what person' or *renhe ren* 'any person.' In (42), the Mandarin conjunction word *he* 'and' has replaced the disjunction word *huozhe* 'or.' Neither adults nor children give Kermit a coin in response to these *ruguo*-conditionals in experimental contexts in which only one person has eaten an apple. Therefore, these *ruguo*-conditionals are not the semantic equivalents of ones with question words.

(41) Ruguo mei-ge ren chi-le pinguo, ni jiu gei wo
if every-CL person eat-ASP apple, you then give me
yi-ge yingbi.
one-CL coin
'If everyone ate an apple, then give me a coin.'

(42) Ruguo Aijun, Liao he Haihua chi-le pinguo, ni jiu gei wo
if Aijun Liao and Haihua eat-ASP apple, you then give me
yi-ge yingbi.
one-CL coin
'If Aijun, Liao and Haihua ate an apple, then give me a coin.'

The substitutability of *shenme, renhe* and *huozhe*, *salva veritate*, can be explained in the framework proposed by Chierchia (2013, 2017). According to this framework, Mandarin *wh*-words and the polarity sensitive expression *renhe* 'any' are existential quantifiers which, in a finite domain, are logically equivalent to disjunctive statements.

Other predictions of the unified account are also borne out. The Mandarin polarity sensitive expression *renhe* 'any' and the disjunction word *huozhe* 'or' both license conjunctive/universal inferences in sentences with the deontic modal verb *keyi* 'is allowed to' and the quantificational adverb *dou* 'all.' This is illustrated in example sentences (43) and (44). Sentence (43) generates the conjunctive/universal inference that Kung Fu Panda is free to choose which car to push. Sentence (43) is logically equivalent to the disjunctive statement (44) in a conversational context with just a green car and an orange car.

(43) Renhe xiaoche Gongfu xiongmao **dou** **keyi** tui.
any car Kung Fu Panda all may push
'Kung Fu Panda may push any of the cars.'

(44)	Lüse	xiaoche	huozhe	juse	xiaoche	Gongfu	xiongmao	**dou**
	green	car	or	orange	car	Kung Fu	Panda	all
	keyi	tui.						
	may	push						

'Kung Fu Panda may push the green car or the orange car.'

(45)	Shenme	xiaoche	Gongfu	xiongmao	**dou**	**keyi**	tui.
	what	car	Kung Fu	Panda	all	may	push

'Kung Fu Panda may push any of the cars.'

In Mandarin, moreover, the question word *shenme* 'what' can replace the polarity sensitive expression *renhe* 'any,' without changing the truth conditions of the sentence. This is illustrated in (45). It should be noted, though, that the deontic modal verb *keyi* 'is allowed to' does not suffice to license the conjunctive/universal interpretation of these expressions. The quantificational adverb *dou* 'all' is obligatory in these sentences.

The derivations of the conjunctive/universal interpretations of sentences with the question word *shenme* 'what,' the polarity sensitive expression *renhe* 'any' and the disjunction word *huozhe* 'or' can be generated using the recursive exhaustification algorithm that we outlined earlier. Mandarin existential expressions are first converted into disjunctive statements. Then the exhaustivity operator ONLY is engaged twice. ONLY first generates enhanced subdomain alternatives to the original disjunctive statement. Then ONLY eliminates any of the enhanced alternatives that are informationally stronger than the original disjunctive statement. The conversion from question words or polarity sensitive expressions to disjunction is justified by the logical equivalence of existential expressions and disjunctive statements in a finite domain.

Previous studies have documented young children's ability to generate conjunctive/universal inferences in sentences with disjunction and existential expressions (e.g., English *or*, Mandarin *renhe* 'any'). The first study that we know of was by Gualmini, Meroni, & Crain (2000), who found that 4- to 5-year-old English-speaking children license a 'distributive' interpretation to a disjunction phrase in the subject NP of sentences with the epistemic modal verb *can*. For example, children interpreted the sentence *Tigger or Eeyore can fix the tire* to assert that Tigger can fix the tire and that Eeyore can fix the tire. The Gualmini et al. (2000) study also documented children's knowledge that the corresponding sentences with conjunction, i.e., *Tigger and Eeyore can fix the tire*, generated an additional 'collective' interpretation. According to the collective interpretation, sentences with *and* in the subject NP are true if both characters together are able to fix the tire, even if neither of them can fix the tire alone. The study demonstrated that young English-speaking children know that the collective interpretation is accessible for sentences with conjunction, but not for sentences with disjunction. This is evidence that children know that disjunction and conjunction make sentences true in different circumstances. Two previous studies, by Huang and Crain (2014) and by Tieu, Romoli, Zhou, and Crain (2016), found that 4- to 5-year-old Mandarin-speaking children license a conjunctive/universal inference in sentences with disjunction and a modal verb. The study by Tieu, Romoli, Zhou, and Crain (2016) also found that 4- to 5-year-old English-speaking children licensed a conjunctive/universal inference in response to sentences with a modal verb and the polarity sensitive expression *any*.

9. Experiment 1: Question words as existential expressions

This experiment was originally reported in Su, Zhou, and Crain (2012). It investigated Mandarin-speaking children's interpretation of the *wh*-word *shenme* 'what' in the predicate phrases of *meiyouren* 'nobody' versus *meigeren* 'everybody'. We tested twenty-eight monolingual Mandarin-speaking children ranging in age from 3;5 to 4;9 with a mean age of 4;2. The child participants attended the kindergarten at the Beijing Language and Culture University. In addition, twenty Mandarin-speaking adults served as controls. The adults were postgraduate students at the Beijing Language and Culture University.

Experiment 1 used the *Question/Statement task* (Zhou & Crain, 2011). This task involves two experimenters. One experimenter acts out stories in front of the child participant using toy characters and props. The second experimenter manipulates a puppet that watches the stories alongside the child. After each story, the puppet being manipulated by the second experimenter explains to the child participant what it thinks happened in the story. Sometimes the puppet produces a statement, which can either be right or wrong. Alternatively, when the puppet is unsure about the outcome of the story, the puppet produces an information-seeking question. Whenever, in the child's judgment, the puppet makes a statement about the outcome of the story, the child is instructed to indicate whether the puppet's statement is right ('Yes') or wrong ('No'). Children's Yes and No responses are taken as evidence that, in the child's mind, the puppet made a statement. If, in the child's judgment, the puppet asks an information-seeking question, the child is instructed to provide the requested information to the puppet, i.e. to answer the question. Children's responses that contain information are taken as evidence that, in the child's mind, the puppet asked a question rather than made a statement.

The experimental materials consisted of six test stories, which were presented to each child individually. Three of the test sentences that followed these stories contained the Mandarin question word *shenme* 'what,' which was positioned in the predicate phrase of the negative quantificational expression *meiyouren* 'nobody.' The other three test sentences positioned *shenme* 'what' in the predicate phrase of the positive quantificational expression *meigeren* 'everybody.' For adults, the test sentences evoke a statement when *shenme* 'what' appears in the predicate phrase of *meiyouren* 'nobody,' *but shenme* 'what' evokes a question reading when it appears in the predicate phrase of *meigeren* 'everybody.' The research question was whether child speakers of Mandarin know that the question word *shenme* 'what' changes the illocutionary force of utterances, depending on the licensing quantificational expression.

On a typical trial, the story was about three pandas that went to have breakfast. There were four choices: three types of fruit (lemons, pears and strawberries) and eggs. None of the pandas picked lemons, because they were too sour. Each Panda then took an egg. Two of the Pandas didn't like pears, but the third panda chose one big pear. Finally, they came to the strawberries. They all took a strawberry. Figure 1 illustrates the scene at the end of the story. After the story concluded, the puppet either produced test sentence (46) or test sentence (47). Participants were divided into two groups for the presentation of the test trials. The first group (fourteen children and ten adults) was tested with sentence (46), and the second group (fourteen children and ten adults) was tested with sentence (47). These two types of test sentences were counterbalanced for each group.

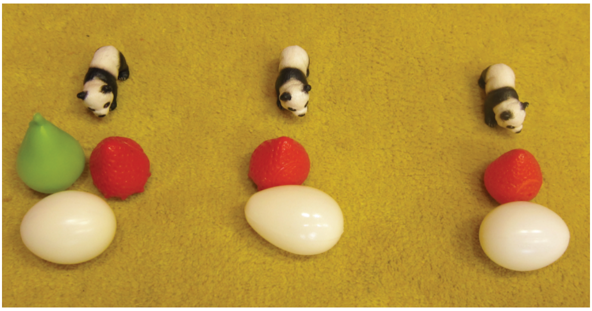


Figure 1. The end scene of the example story, Experiment 1

(46) Meiyou xiongmao chi shenme shuiguo.
not-have panda eat what fruit
'No panda ate any fruit.'

(47) Mei-zhi xiongmao dou chi-le shenme shuiguo?
every-CL panda all eat-ASP what fruit
'What kind of fruit did every panda eat?'

Two noteworthy experimental manipulations were incorporated into the experimental design to ensure that the participants judged the test sentences based on the linguistic structures, rather than other linguistic properties specific to Mandarin.

First, no sentence-final particles were used. Mandarin sentence-final particles include the modal particles *a* and *o*, and the question particles *ne* and *ma*. The modal particle *a* indicates affirmation, approval or consent, and the question particle *ma* converts declarative sentences into interrogative sentences. These particles are optional, so their absence in the test sentences ensured that participants relied on the sentence structure to determine whether they were questions or statements. Second, the experiment controlled for the effects of prosody. The experimenter who played the role of the puppet was instructed to utter the test sentences with level intonation, without assigning stress to the *wh*-word, for both types of test sentences. This ensured that the child participants were not able to use prosodic cues to determine which speech act the puppet was performing. To avoid any inadvertent bias towards either interpretive pattern, the experimenter who manipulated the puppet was not informed about the purpose of the experiment.

The main experimental finding was that children and adults analysed *shenme* as a negative polarity item in the predicate phrase of *meiyouren* 'nobody' 100% of the time, so they interpreted (46) as a statement, and rejected it on the grounds that every panda ate some fruit. By contrast, the same children, as well as adults, interpreted *shenme* in the predicate phrase of *meigeren* 'everybody' (e.g. (47)) as a question-marker 100% of the time, so they provided answers to this type of test sentence. In the following section, we turn to another linguistic structure in which question words are interchangeable with the existential expression *renhe* 'any.'

10. Experiment 2: Question words in the antecedent of conditionals

The second experiment assessed Mandarin-speaking children's interpretation of *wh*-words in the antecedent of conditionals. The experiment was originally reported in Zhou and Crain (2011). Examples (48) and (49) reveal that Mandarin question words and *renhe* 'any' are interchangeable when these expressions are positioned in the antecedent clause of a *ruguo*-conditional (English 'if…, then'). Both Examples (48) and (49) are statements about the circumstances in which the village head, Mr. Owl, will seek Batman's help. However, when question words are positioned in the consequent clause, as in (50), they revert to their semantic function as a question marker. So, (50) asks whose assistance the village head will seek in certain circumstances.

(48) Ruguo <u>renhe</u> ren qifu xiaotuzi, Maotouying cunzhang jiu
if any person bully rabbit owl village-head then
hui zhao Bianfuxia bangmang.
will find Batman help
'If anybody bullies Mr. Rabbit, Mr. Owl will ask Batman for help.'

(49) Ruguo shenme ren qifu xiaotuzi, Maotouying cunzhang jiu
if what person bully rabbit owl village-head then
hui zhao Bianfuxia bangmang.
will find Batman help
'If anybody bullies Mr. Rabbit, Mr. Owl will ask Batman for help.'

(50) Ruguo eyu qifu xiaotuzi, Maotouying cunzhang hui zhao
if crocodile bully rabbit owl village-head will find
shenme ren bangmang?
what person help
'If Mr. Crocodile bullies Mr. Rabbit, who will Mr. Owl ask for help?'

Mandarin has more than one kind of conditional structure. In addition to *ruguo*-conditionals, Mandarin has WH-conditionals, and *dou*-conditionals. It turns out that question words like *shenme* 'what' and *shei* 'who' are licensed in the antecedent of all three kinds of conditionals. Experiment 2 investigated Mandarin-speaking children's understanding of the meaning of question words in *dou*-conditionals. There were two test sentences. Both versions were presented following the same story, but to different participants. The difference is in the placement of the question words in the test sentences. In (51) the question word *shei* 'who' appears in the antecedent of a *dou*-conditional, whereas in (52) the same question word *shei* 'who' appears in the consequent clause. As the English glosses indicate, question words have existential import when they appear in the antecedent clause of *dou*-conditionals, but they function as question markers in the consequent clause.

(51) Eyu qu yao shei, maotouying cunzhang dou zhao-le
crocodile go bite who owl village-head all find-ASP
zhizhuxia bangmang.
Spiderman help
'If Mr. Crocodile went to bite anybody, Mr. Owl asked Spiderman for help.'

(52) Eyu qu yao xiaozhu, maotouying cunzhang dou zhao-le
crocodile go bite pig owl village-head all find-ASP
shei bangmang?
who help
'If Mr. Crocodile went to bite Mr. Pig, who did Mr. Owl ask for help?'

Again, this experiment used the Question/Statement task. We interviewed 30 Mandarin-speaking children ranging in age from 3;5 to 5;0, with a mean age of 4;3. The child participants were recruited from the kindergarten at the Beijing Language and Culture University. In addition, 20 Mandarin-speaking adults were tested as controls, all postgraduate students at the Beijing Language and Culture University. For each story, two types of test sentences were created. Altogether there were six test stories. We will illustrate Experiment 2 using the story that was associated with the test sentences (51) and (52).

The story was about a small village. There are four guys living in this village. One is the village head, Mr. Owl, and there are 3 villagers: Mr. Rabbit, Mr. Pig and Mr. Cat. The villagers lived a happy life until last week, when Mr. Crocodile started coming to the village. Mr. Crocodile always came at night to try to bite the villagers. When the village head, Mr. Owl, found out about this, he asked two of his superhero friends, Mr. Spiderman and Mr. Batman, to protect the villagers. The next night Mr. Crocodile went to Mr. Pig's house. Mr. Owl called Mr. Spiderman. Just when Mr. Crocodile was about to bite Mr. Pig, Mr. Spiderman came down from the roof and ran Mr. Crocodile out of the house. As a reward, Mr. Owl gave Mr. Spiderman a purple shell. The next night Mr. Crocodile wanted revenge, and went to Mr. Pig's house again. Mr. Owl called Mr. Spiderman again. When Mr. Crocodile jumped towards Mr. Pig, Mr. Spiderman came out from behind the door and ran Mr. Crocodile out of the house. Mr. Owl rewarded Mr. Spiderman with another purple shell. The following night, Mr. Crocodile decided not to try to bite Mr. Pig, so he went to Mr. Rabbit's house. This time, Mr. Owl called Mr. Batman. When Mr. Crocodile broke into the house, Mr. Batman ran Mr. Crocodile out of the house. Mr. Owl rewarded Mr. Batman with a white shell. After that, Mr. Crocodile left the village and never came back. Figure 2 illustrates the last scene of the story.

Figure 2. The end scene of the example story, Experiment 2

For the presentation of the two different sentence types, the child and adult participants were divided into two groups, with 15 children and 10 adults in each group. When the story concluded, the puppet produced the test sentence (51) to one group of children, and (52) to the second group. There were six test stories altogether. Each group was presented with only one type of test sentence for each story, and the two types of test sentences were counterbalanced. Again, the two research strategies described in Experiment 1 were adopted in Experiment 2, to ensure that the participants made their judgements solely based on the properties of the linguistic structures, as well as to control for potential prosodic effects on their judgements.

Here are the main findings. In response to sentences like (51), both children and adults consistently interpreted them as statements, i.e., children correctly rejected the sentences 88% of the time and adults did so 90% of the time. A Mann-Whitney Test showed that there was no significant difference between children and adults in their rejection rates ($Z = .25$, $p = 1$). On the example trial, both children and adults rejected sentence (51) by explicitly referring to the fact that Mr. Batman also came to help. In response to sentences like (52), both children and adults interpreted the sentences as questions, and provided answers 100% of the time. On the example trial, all the children and adults answered "Mr. Spiderman" in response to sentence (52).

In the introduction to Experiment 2, we observed that the Mandarin existential expression *renhe* 'any' and the Mandarin question marker *shenme* 'what' are both licensed, and have the same meaning, in the antecedent clause of a *ruguo*-conditional. We noted, however, that the Mandarin existential expression *renhe* 'any' is not tolerated in the consequent clause of a *ruguo*-conditional, and the Mandarin question word *shenme* 'what' functions as a question marker in the consequent clause. Experiment 2 investigated children's knowledge of how children interpreted the question word *shei* 'who' in the antecedent versus the consequent clause of a *dou*-conditional. The experimental findings indicated that young Mandarin-speaking children know that the question word *shei* 'who' has a non-interrogative meaning in the antecedent of a *dou*-conditional, but it can only function as a question marker in the consequent clause.

11. Experiment 3: Question words license conjunctive/universal inferences

The third experiment has not been reported previously. This experiment investigated another overlap in meaning between Mandarin question words and *renhe* 'any.' In the lead-up to Experiment 1, we discussed one meaning of *renhe* 'any', which is licensed by Downward Entailing operators. The Mandarin expression *renhe* 'any' can also be licensed in 'positive' linguistic environment and, when it is, it assumes a different semantic guise. For example, *renhe* 'any' is accepted in sentences that contain the deontic modal *keyi* 'may' and it is accepted in sentences with the passive structure *bei yunxu* 'to be allowed to.' In Mandarin *renhe* 'any' appears in combination with these linguistic expressions, it licenses a conjunctive/universal inference. To illustrate, it is perfectly acceptable to use Mandarin *renhe* in (53), where it is licensed by the Mandarin deontic modal verb *keyi* 'is allowed to.' The English existential expression *any* exhibits the same linguistic behaviour, as illustrated in Example (54). Both (53) and (54) assert that Kung Fu Panda is free to choose which car to push. By contrast, if the deontic modal is replaced by the perfective aspectual marker *-le*, the sentence is not acceptable, as illustrated by (55). When the English modal *may* is omitted, as in (56), a similar anomaly results.

(53) Gongfu xiongmao **keyi** tui renhe xiaoche.
Kung Fu panda may push any car
'Kung Fu Panda may push any of the cars.'

(54) Kung Fu Panda **may** push any of the cars.

(55) *Gongfu xiongmao tui-**le** renhe xiaoche.
Kung Fu panda push-ASP any car
'Kung Fu Panda pushed any of the cars.'

(56) *Kung Fu Panda pushed any of the cars.

A similar pattern of linguistic behaviour is manifested by disjunction words in Mandarin and in English. In Example (57) the Mandarin disjunction word *huozhe* 'or' appears with the deontic modal *keyi* 'is allowed to.' Although Example (57) contains the Mandarin disjunction word *huozhe* 'or,' its meaning can be paraphrased as *Kung Fu Panda is allowed to push the green car and Kung Fu Panda is allowed to push the orange car*. If the Mandarin deontic modal *keyi* is replaced by the perfective aspectual marker *–le*, the meaning reverts to a statement with disjunctive truth conditions. Example (58) is true, for instance, if Kung Fu Panda pushed just one of the cars.

(57) Gongfu xiongmao **keyi** tui lüse xiaoche huozhe juse xiaoche.
Kung Fu Panda may push green car or orange car
'Kung Fu Panda may push the green car or the orange car'

(58) Gongfu xiongmao tui-**le** lüse xiaoche huozhe juse xiaoche.
Kung Fu Panda push-ASP green car or orange car
'Kung Fu Panda pushed the green car or the orange car'

Mandarin question words *shei* 'who' and *shenme* 'what' also generate conjunctive/universal inferences in sentences with a deontic modal verb. There is a catch, however. Mandarin question words generate such inferences just in case they are followed by the quantificational adverb *dou*. This is illustrated in Examples (59) and (60), where both *renhe* and *shenme* are glossed as English *any*. Example (61) shows that the deontic modal *keyi* 'is allowed to' alone does not suffice to license a conjunctive/universal inference in sentences with question words. Lacking a proper licensor, the quantificational adverb *dou*, the question word *shenme* 'what' in (61) yields an information-seeking question. Question words also function as question markers when they follow the quantificational adverb *dou*, rather than precede it. This 'leftness' constraint is illustrated in (62), which can only be used to seek information.

(59)	Renhe	che	Gongfu	xiongmao	**dou**	**keyi**	tui.
	any	car	Kung Fu	Panda	all	may	push

'Kung Fu Panda may push any of the cars.'

(60)	Shenme	che	Gongfu	xiongmao	**dou**	**keyi**	tui.
	what	car	Kung Fu	Panda	all	may	push

'Kung Fu Panda may push any of the cars.'

(61)	Shenme	che	Gongfu	xiongmao	**keyi**	tui?
	what	car	Kung Fu	Panda	may	push

'What car may Kung Fu Panda push?'

(62)	Xiongmaomen	**dou**	**keyi**	tui	shenme	che?
	pandas	all	may	push	what	car

'What car may every panda push?'

Experiment 3 investigated children's knowledge of conjunctive/universal inferences generated by question words. The participants included 28 children, ranging in age from 4;1 to 4;9, with a mean age of 4;3. The child participants were recruited from the kindergarten at the Beijing Language and Culture University. In addition, 22 adults served controls and were recruited from the same university.

Experiment 3 presented children with sentences that contain question words in Topic position, both with and without the quantificational adverb *dou*. As in the first two experiments, a Question/Statement task was used to ascertain whether or not children analyse sentences that combine *dou* and the deontic modal *keyi* 'may' as statements that license Free Choice inferences, as in (60), but analyse sentences without *dou* as information-seeking questions, as in (61). Experiment 3 consisted of six test stories. For each story, two types of test sentences were constructed. One contained the *wh*-word *shenme* 'what,' the deontic modal *keyi* 'may' and the quantificational adverb *dou*, as in (60), and one contained the *wh*-word *shenme* 'what' and the deontic modal *keyi* 'may' but with the quantificational adverb *dou* removed, as in (61).

On a typical trial, Kung Fu Panda and Batman were engaged in a racing competition. Before the competition started, Mr. Owl (the judge) explained the rules. Mr. Owl told Kung Fu Panda that he was allowed to push the orange car and the green car, but not the purple car. Mr. Owl told Batman that he was only allowed to push the purple car, because he was not strong enough to push two cars. But Kung Fu Panda and Batman were very forgetful, and forgot the rules when the game was about to start. So they asked the puppet to restate the rules. Figure 3 illustrates the scene.

The puppet then either produced the test sentence in (60) or the one in (61). Each participant was instructed to respond to the puppet's utterance based on his/her understanding of whether the utterance was a statement or a question. For the presentation of the test sentences, the participants of the experimental and control groups were further divided into two subgroups, with 14 children and 11 adults in each subgroup. One subgroup heard test sentence (60) and the other heard test sentence (61). There were six test stories altogether. Each subgroup was presented with only one type of test sentences for each story, and the two types of test sentences were counterbalanced, so that each subgroup was presented with three each, of the two sentence types. Again, the same two research strategies used in Experiments 1 & 2 were adopted in Experiment 3, to ensure that the participants made their judgements solely based on the properties of the linguistic structures, and did not use prosodic information.

Figure 3. Example scene, Experiment 3

Let us review the experimental hypotheses. If children compute a conjunctive/universal inference in sentences in which the quantificational adverb *dou* is associated with a *wh*-word, as in (60), then they should judge (60) to be a false description of the story, because the sentence would be interpreted to mean that Kung Fu Panda had permission to push the green car, the orange car, and the purple car. However, in the story Kung Fu Panda was only given permission to push the green car and the orange car, but not the purple car. Children were therefore expected to reject (60) by pointing out that Kung Fu Panda was not allowed to push the purple car. By contrast, children were expected to judge the *wh*-word in (61) as a question marker, because conjunctive/universal inferences of *wh*-words are not licensed in the absence of the quantificational adverb *dou*. On the example trial, children were expected to respond to (61) with the answer "the green car and the orange car."

Here are the main findings. Children rejected sentences like (60) 96% of the time. Adults did so 100% of the time. A Mann-Whitney Test showed no significant difference in these rejection rates between groups ($Z = 1.91$, $p = .1$). When asked to justify their rejections, both children and adults pointed out that the character under consideration was only allowed to perform the action on two of the objects, but not on the third object. On the example trial, children and adults rejected the puppet's statement in (60) by making reference to the fact that Kung Fu panda was allowed to push two of the three cars, but not the third one. In response to sentences like (61), both children and adults interpreted the sentences as questions, and provided answers 100% of the time. On the example trial, all the children and adults answered "the orange car and the green car" in response to sentence (61).

12. Conclusion

To conclude, when the Mandarin question words *shenme* 'what' and *shei* 'who' are used in statements, their meanings directly correspond to the Mandarin existential expression *renhe* 'any'. The correspondence between Mandarin question words and existential expressions extends to both Downward Entailing linguistic environments, and to 'positive' linguistic environments with operators (such as modal verbs) that license conjunctive/universal inferences. A similar pattern is observed with the Mandarin word for disjunction.

Although *wh*-words are only rarely used as existential expressions in English, there are extensive parallels across languages in the interpretation of existential expressions. As we have discovered, Mandarin *renhe* 'any' and English *any* exhibit remarkably similar distributional properties, as do Mandarin *huozhe* 'or' and English *or*. Most importantly, when Mandarin question words like *shenme* 'what' and *shei* 'who' are used in declarative sentences, their semantic contribution to sentence meaning is remarkably similar to that of Mandarin *renhe* 'any' and English *any*. Like these existential expressions, Mandarin question words are polarity sensitive expressions in Downward Entailing linguistic environments, and they license conjunctive/universal inferences in sentences with a modal verb, for example.

The present study reported the findings of three experiments. These experiments revealed that young Mandarin-speaking children have impressive knowledge of the syntactic and semantic properties of questions words, existential expressions, and disjunction. The findings lend credence to an early proposal by Chomsky (1964) which derived information-seeking questions from their declarative counterparts by a transformational rule. The findings also lend credence to early theoretical proposals that analyzed information-seeking *wh*-questions and their corresponding answers as semantically equivalent (Hamblin, 1973; Karttunen, 1977). The findings of these experimental investigations also support the unified approach to existential expressions recently proposed by Chierchia (2013, 2017). Most broadly, the findings of these experiments add empirical weight to the biolinguistic approach to language. According to the biolinguistic approach, the goal of linguistic theory is the unification or amalgamation of disparate-looking phenomena, within and across languages. As with other scientific theories, linguistic theory seeks to demonstrate that phenomena that look different on the surface are just different combinations of the same basic building blocks. The findings from the present study make it clear that child language is fertile ground for pursuing the predictions of the biolinguistic approach to language.

References

Brown, R. (1973). *A first language: The early stages*. Cambridge, MA: Harvard University Press. https://doi.org/10.4159/harvard.9780674732469

Caponigro, I., Pearl, L., Brooks, N., & Barner, D. (2012). Acquiring the meaning of free relative clauses and plural definite descriptions. *Journal of Semantics*, 29, 261–293. https://doi.org/10.1093/jos/ffr014

Carlson, G. N. (1980). Polarity any is existential. *Linguistic Inquiry*, 11, 799–804.

Cheng, L. L-S. (1997). *On the typology of wh-questions*. New York, NY: Garland.

Cheng, L. L-S. (1994). Wh-words as polarity items. *Chinese Languages and Linguistics*, 2, 615–640.

Chierchia, G. (2013). *Logic in grammar*. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199697977.001.0001

Chierchia, G. (2017). Scalar implicature and their interface with grammar. *The Annual Review of Linguistics*, 3, 245–264.

Chomsky, N. (1964). *Current issues in linguistic theory*. The Hague: Mouton.

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

Chomsky, N., & Lasnik, H. (1993) The theory of principles and parameters. In J. Jacobs, A. von Stechow, W. Sternefeld, & T. Vennemann (Eds.), *Syntax: An international handbook of contemporary research*. Berlin: Walter de Gruyter.

Crain, S. (2012). *The emergence of meaning*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511842863

Dayal, V. (2004). Number marking and (in) definiteness in kind terms. *Linguistics and Philosophy*, 27, 393–450. https://doi.org/10.1023/B:LING.0000024420.80324.67

Fiengo, R., & May, R. (1994). *Indices and identity*. Cambridge, MA: The MIT Press.

Fox, D. (1998). Locality in variable binding. In P. Barbosa, D. Fox, P. Hagstrom, M. McGinnis, & D. Pesetsky (Eds.), *Is the best good enough? Optimality and Competition in Syntax*. Cambridge, MA: MIT Press/MITWPL.

Fox, D. (2007). Free choice disjunction and the theory of scalar implicatures. In U. Sauerland & P. Stateva (Eds.), *Presupposition and implicature in compositional semantics* (pp. 71–120). Houndmills: Palgrave Macmillan.

Gualmini, A., Meroni, L. & Crain, S. (2000). The inclusion of disjunction in child language: Evidence from modal verbs. In M. Hirotani, A. Coetzee, N. Hall, & J-Y. Kim (Eds.), *Proceedings of the North East Linguistic Society 30*. Amherst, MA: GLSA.

Gualmini, A., Crain, S., & Meroni, L. (2000). Acquisition of disjunction in conditional sentences. In S. C. Howell, S. Fish, & T. Keith-Lucas (Eds.), *Proceedings of the 24th annual Boston University Conference on Language Development* (pp. 367–378). Somerville, MA: Cascadilla Press.

Hamblin, C. L. (1973). Questions in Montague English. *Foundations of Language*, 10, 41–53.

Horn, L. R. (1969). A presuppositional approach to only and even. *Proceedings of the Chicago Linguistic Society*, 5, 98–107.

Huang, A., & Crain, S. (2014). Acquisition of the polarity sensitive item renhe 'any' in Mandarin Chinese. *Journal of Child Language*, 41, 861–889. https://doi.org/10.1017/S0305000913000275

Kadmon, N., & Landman, F. (1993). Any. *Linguistics and Philosophy*, 16, 353–422. https://doi.org/10.1007/BF00985272

Karttunen, L. (1977). Syntax and semantics of questions. *Linguistics and Philosophy*, 1, 1–44. https://doi.org/10.1007/BF00351935

Klima, E. S. (1964). Negation in English. In J. A. Fodor & J. J. Katz (Eds.), *The structure of language: Readings in the philosophy of language*. Englewood Cliffs, NJ: Prentice-Hall.

Ladusaw, W. (1980). *Polarity sensitivity as inherent scope relations*. New York, NY: Garland.

Lasnik, H. (1972) *Analysis of negation in English* (Unpublished doctoral dissertation). MIT.

Li, Y-H. A. (1992). Indefinite wh in Mandarin Chinese. *Journal of East Asian Linguistics*, 1, 125–155. https://doi.org/10.1007/BF00130234

Li, Y., & Tang, Z. (1991). *Shenme as an NPI from 2-years-old*. Wuhan: Middle China Normal University.

Lin, J-W. (1998). On existential polarity wh-phrases in Chinese. *Journal of East Asian Linguistics*, 7, 219–255. https://doi.org/10.1023/A:1008284513325

MacWhinney, B. (2000). *The CHILDES project: Tools for analysing talk* (3rd ed.), Mahwah, NJ: Lawrence Erlbaum Associates.

Modyanova, N., & Wexler, K. (2008). Maximal trouble in free relatives. In H. Chan, H. Jacob, & E. Kapia (Eds.), *BUCLD 32: Proceedings of the 32nd annual Boston University Conference on Child Language Development* (pp. 287–298). Somerville, MA: Cascadilla Press.

Munn, A., Miller, K., & Schmitt, C. (2006). Maximality and plurality in children's interpretation of definites. In D. Bamman, T. Magnitskaia, & C. Zaller (Eds.), *Proceedings of the 30th annual Boston University Conference on Child Language Development* (pp. 377–387). Somerville, MA: Cascadilla Press.

Quine, W. V. O. (1960). *Word and object*. Cambridge, MA: The MIT Press.

Shimoyama, J. (2011). Japanese indeterminate negative polarity items and their scope. *Journal of Semantics*, 28, 413–450. https://doi.org/10.1093/jos/ffr004

Su, Y., Zhou, P., & Crain, S. (2012). Downward entailment in child Mandarin. *Journal of Child Language*, 39, 957–990. https://doi.org/10.1017/S0305000911000389

Tieu, L., Romoli, J., Zhou, P., & Crain, S. (2016). Children's knowledge of free choice inferences and scalar implicatures. *Journal of Semantics*, 33, 269–298.

Thornton, R., & Wexler, K. (1999). *Principle B, VP ellipsis, and interpretation in child grammar*. Cambridge, MA: The MIT Press.

Zhou, P., & Crain, S. (2011). Children's knowledge of the quantifier *dou* in Mandarin Chinese. *Journal of Psycholinguistic Research*, 40, 155–176. https://doi.org/10.1007/s10936-010-9161-z

CHAPTER 12

Overt, covert, and clandestine operations

Ambiguity and ellipsis in acquisition

Kristen Syrett
Rutgers University

One of the major challenges on the path to becoming an adult speaker arises from ambiguous sentences – sentences that are in principle compatible with multiple interpretations. In this chapter, I review experimental evidence from a series of studies run with children age four to six years, focusing on three cases of sentential ambiguity. The first is the case of ambiguity of question-answer relations arising from the interaction of *wh*-phrases and universal quantifiers. The second is ambiguity resulting from covert movement of a quantificational phrase yielding different interpretations of verb phrase ellipsis. The third is the interpretation of comparative constructions involving pronominal reference. In each instance, a successful interpretation depends on one or more successful abstract syntactic-semantic *operations*, for which the child must deploy special forces. But as history tells us with such operations outside of the grammar, not every operation is a success, and failures (in this case, retrieval of a non-target interpretation, or the generation of interpretations beyond the target one), can either indicate the child's developing linguistic capacity or the extent to which the range of possibilities presented by the adult grammar may have been underestimated.

Keywords: movement, quantifiers, quantifier raising, verb phrase ellipsis, *wh*-questions, pronouns, comparative constructions

1. Introduction

My goal in this paper is to draw from three separate lines of language acquisition research to present evidence for the presence of abstract syntactic-semantic operations in child language. These lines of research collectively demonstrate how children arrive at interpretations of (uncontroversially or potentially) ambiguous sentences in which some sort of movement has taken place – whether signaled on the surface or not. They therefore provide a window into the structure of the child

https://doi.org/10.1075/tilar.24.12syr

grammar. But where the range of interpretations that children are able to access does not align with that of adults, we are also presented with evidence bearing on the composition of the adult grammar.

I have chosen to adopt terminology from special forces operations to refer to these three lines of evidence (*overt*, *covert*, and *clandestine* operations) – not because I want to draw special attention to the U.S. Department of Defense or think this organization should be connected in any way to language acquisition research – but because each line of research illustrates at least one of these types of operations, metaphorically speaking, and this terminology becomes rather useful when categorizing and referring to each of these lines of research. I will begin the paper with an introduction to the terminology, and how it applies to certain linguistic phenomena that will form the core of our three target case studies. I will then identify and review each of these case studies in turn, and then close by connecting them together and summarizing what they reveal collectively about the child and adult grammars.

2. Terminological background and selected case studies

Let me first define the terms *overt*, *covert*, and *clandestine* operations, and then describe how this terminology might apply to particular cases in language, which will feed into our case studies in language acquisition.

2.1 Overt operation

An ***overt*** operation is an operation conducted openly and without concealment; there's no doubt about what happened, who did it, or why. You know the shape of the operation and the forces involved, even if you don't know all the details, and even if the sponsor or operator takes great measures to maintain security and secrecy prior to and during the operation.

This is the *modus operandi* of *wh*-movement. You know it happened and you know who (or what) was responsible, because the *wh*-word was displaced, as illustrated in the questions in (1). In each, the *wh*-word is underlined, and its base position is signaled with a trace (*t*).

(1) a. <u>Where</u> did the 2008 Summer Olympics take place *t*?
b. <u>What</u> was the nickname of the 2012 USA female Olympics gymnastics team *t*?

Now, perhaps we do not know all of the details of what went on behind the scenes, because there is no trace of that on the surface (at least in English). But it is clear that *wh*-movement has taken place, because (a) the *wh*- word is no longer in its base position, and (b) the illocutionary force of the sentence is a question, not an assertion as signaled by a declarative statement.

2.2 Covert operation

A ***covert*** operation is one that is planned and executed in a way as to conceal the identity of, or permit plausible denial by, the sponsor. In a military covert operation, emphasis is placed on the concealment of the identity of the sponsor, rather than concealment of the operation itself. Thus, you have evidence that the operation took place, but cannot say who was responsible.

There are two kinds of covert operations represented in the research presented here. The first is Quantifier Raising (QR) of a quantificational phrase (QNP), as shown in (2).

(2) Aly's gold-winning floor routine impressed <u>everyone in the audience</u>.

In a Generative framework, a QNP must raise covertly from its base position for reasons of type mismatch (i.e., because it is the wrong semantic type to compose with the transitive verb taking it as an argument) (Heim & Kratzer, 1998). This covert movement shares many of the same features with *wh*-movement in terms of the path of movement through the structure, but the QNP is not noticeably displaced from its origin. Thus, a listener cannot tell that the QNP has been displaced in the representation; evidence that the operation has taken place is simply that the sentence is interpretable.

A QNP may also move covertly to take 'scope' at a different level of the representation (Fiengo & May, 1994; May, 1977, 1985, 1988), motivated by the possibility of generating a new interpretation (Fox 1995a, 2000), as in (3). Here, the QNP *every gymnast* in object position interacts scopally with the indefinite *a reporter* in subject position to generate two different interpretations, based on whether it takes 'narrow' or 'wide' scope with respect to the indefinite.

(3) <u>A reporter</u> interviewed <u>every female gymnast that competed in the 2012 Olympics</u>.
  a. indefinite > QNP: one reporter (the same one) interviewed every gymnast
  b. QNP > indefinite: every gymnast was interviewed by a (possibly different) reporter

The fact that there is a QNP in a sentence should be a relatively reliable indicator that a covert movement operation has taken place, since QR is obligatory to resolve type mismatch. However, one might ask who is harboring this displaced quantificational fugitive? That is, where did the QNP land when it raised? To know the answer to this question is to know something about the scope of the QNP and the possible interpretations of the sentence in which it occurs.

QR is also implicated in many sentences that do not have an overt quantifier such as *every*. For example, in comparative constructions such as the one in (4), what is being compared is the degree to which each gymnast's score was high, and each gymnast's score can be represented as a set of degrees *d* such that it is d-high. The degree phrase headed by the –*er* morpheme is taken to be a degree quantifier, which is claimed to obligatorily raise via QR to make the sentence interpretable.

(4) The score Aly earned on the floor routine was higher than Jordyn's score.

Thus, the set of degrees representing Jordyn's score was exceeded by (or was a subset of) the set of degrees representing Aly's score, as represented in (5) (Hackl, 2000; Heim, 1985; a.o.).

(5) [-er than Jordyn's score was d-high]$_1$ [the score Aly earned…was d$_1$-high]

The second type of covert operation is Verb Phrase Ellipsis (VPE). This is situation where an entire VP has been elided. You therefore know that the operation has taken place, because the sentence lacks an overt VP, but also because the absence of this VP on the surface may be signaled by a word such as *did* on the surface, as in (6), which should be interpreted in the same way as its antecedent, the underlined VP preceding it (Hankamer & Sag, 1976) (or at least highly similar way, allowing for possible mismatches in active/passive voice).

(6) Shawn <u>won four medals</u> and Dominique *did*, too.

As with QR, there is, however, an open question of what the scope of this operation was. If there are multiple possible antecedent VPs embedded in each other, as in (7), then there is an open question about which VP serves as the antecedent. As a result, the sentence is ambiguous.

(7) Shawn [<u>wanted to [win four medals]]</u>, just like Dominique *did*.
did = [wanted to win four medals] or [won four medals]?

QR and VPE interact in a construction known as Antecedent-Contained Deletion (ACD), illustrated in (8a). In this construction, the site of VPE (again, signaled by *did*) is contained within its antecedent VP (*competed*…).

(8) a. Aly [competed in [every event that Gabby *did*]].
b. Aly [every event that Gabby *did*] [competed in *t*].
c. Aly [every event that Gabby *competed in*] [competed in *t*].

The way to resolve this situation is for the QNP in which the VPE appears to raise out of this antecedent via QR, so that the VPE can be resolved (Fiengo & May, 1994; Fox, 2000; Kennedy, 1997; Merchant, 2000). There are independent reasons to think that the QNP lands in a position below the subject, as shown in (8b) (Fox 1995b; Merchant, 2000). Once the QNP has raised out of the antecedent, it can copy in the information from the antecedent to be interpreted similarly, as in (8c).

However, there may still be an open question as to the landing site of the QNP, especially when the VPE is embedded in multiple antecedents. The QNP could raise out of just the lower VP or the higher VP as well. As a result, a sentence such as (9) is ambiguous, as indicated by the two interpretations below it.

(9) Shawn [wanted to [win [every event that Nastia *did*]]].
a. *Embedded*: For every event that Nastia *won*, Shawn also wanted to win that event.
b. *Matrix*: For every event that Nastia *wanted to win*, Shawn also wanted to win that event.

2.3 Clandestine operation

A ***clandestine*** operation differs from a covert operation in that it is done in complete secrecy. It is planned and executed in such as way as to conceal the *operation itself*, rather than just the identity of the *sponsor*. Clandestine missions leave no trace that the operation ever happened. If and when you somehow realize that they have happened, it's too late.

To my knowledge, this terminology is not actively used in linguistics, as the other two terms are (for movement operations). However, there is a situation where it seems to me that this term could apply: the case of a pronoun in a site of (VP) ellipsis. Because the pronoun is contained in elided material, it is not pronounced. Unlike the entire VP, whose presence may be signaled by a word such as *did* or *does*, the pronoun – because it is contained in the VP – is not signaled by any specific lexical item. Nevertheless, it is interpreted along with the rest of the elided material, based on a relation with the overt material. However, the pronoun could stand in different identity relations with the overt material, as illustrated in Examples (10) and (11).

(10) Kyla$_i$ waved to her$_i$ mother, and Jordyn$_j$ did ⟨wave to her$_{\{i/j\}}$ mother⟩, too.

(11) The person who gave Gabby$_i$ a medal congratulated her$_i$, and the one who gave McKayla$_j$ a medal did ⟨congratulated her$_{\{i/j\}}$⟩, too.

In both of these examples, the pronoun can be interpreted in one of two ways. Under the first interpretation, the pronoun can stand in a 'strict' identity relation with the overt pronoun (the anaphor) in the VP antecedent, and be interpreted precisely as the overt pronoun is, referring back to the previous name. Both pronouns then would share the same index. In (10), both pronouns would make reference to Kyla, and in (11), both would make reference to Gabby.

Under the second interpretation, the elided pronoun is not interpreted in a way that is identical to the pronoun in the VP antecedent, but is similar in some features (e.g., gender) and is coindexed with another entity. In this case, the identity relation is termed 'sloppy' (Fiengo & May, 1994; Reinhart, 1983; Ross, 1967; Sag, 1976). Each pronoun, then, would bear it own index, and would be linked to a different name. In (10), the first pronoun would be linked to Kyla and the second to Jordyn, and in (11), the first would be linked to Gabby and the second to McKayla.

One way to account for this difference is to posit that when the VP is copied into the site of ellipsis, either the original pronoun is copied in (resulting in strict identity), or else a variable is copied in (allowing for sloppy identity to arise). However, there is no way of knowing what operation took place in the ellipsis site, since the material is elided, and there is no cue on the surface about which interpretation is favored. The listener would have to reconstruct the most plausible interpretation based on the discourse context – and even then may not get it right.

2.4 Three case studies

The three lines of research I will present in this paper collectively cover these three types of operations (overt, covert, and clandestine), and implicate the specific linguistic phenomena described above: *wh*-movement, QR, VPE, comparatives, and pronominal reference in ellipsis. In each instance, a successful interpretation depends on a successful operation, for which the child must deploy so-called special forces. Thus, children's successful interpretations of sentences implicating these operations reveal that children must be appealing to abstract syntactic-semantic operations in their grammar in order to arrive at a particular meaning.

But as history tells us outside of language, not every special forces operation is a success, and in this domain, not every interpretation is a success. When the child's interpretation diverges from the adult interpretation, we are left with the question of why this is the case. A failure to retrieve the target interpretation, or the ability to generate those beyond the intended one, can either reflect the child's developing grammar or processing capability, or the possibility that the range of possibilities presented by the adult grammar may have been underestimated. Thus, to the extent that children's interpretations either converge with or diverge from those of adults, we can better understand the path of language development and the nature of the adult grammar. Combined, these three case studies shed light on both child and adult language, and illustrate the value of language acquisition research for further understanding the grammar.

The first case study, illustrating both overt and covert operations, involves the possible answers available in response to questions involving the interaction of a *wh*-phrase and a universal quantifier, as shown in (12a–b).

(12) a. Which musical piece did [every/each] gymnast select?
  b. Which gymnast performed [every/each] routine?

(12a) is said to allow for both a single answer (the one musical piece all of the gymnasts selected) and a 'pair-list' answer (a list of musical piece-gymnast pairings), while (12b) is said to allow for only a single answer (the one gymnast) (Agüero-Bautista, 2001; Chierchia, 1993; May, 1985, 1988), unless the quantifier is *each* (Beghelli, 1997; Beghelli & Stowell, 1997; Szabolcsi, 1997). The range of children's answers to such questions – which is a superset of those adults provide in one instance and a proper subset in another – highlights the abstract mechanisms at play in the child grammar and lexical entries for universal quantifiers that are refined over the course of language development.

The second case study, implicating the covert operations of QR and VPE, is the case of embedded ACD, as shown in (13).

(13) a. Shawn [wanted to [win [<u>every event that Nastia *did*</u>]]].
  Embedded reading: *did* = win…
  Matrix reading: *did* = wanted to win…
 b. Shawn [said that she [won [<u>every event that Nastia *did*</u>]]].
  Embedded reading: *did* = won…
  Matrix reading: *did* = said that she won…

In (13a), ACD is embedded in a non-finite clause; in (13b), ACD is embedded in a finite, or tensed, clause. While the sentence in (13a) is uncontroversially ambiguous, allowing for both the embedded and matrix readings, the availability of the matrix reading for (13b) has, until recently, been seen as ruled out, or at the very least deemed highly questionable, because of a boundary introduced by the tensed clause (Farkas 1981; Hornstein, 1994; Larson & May 1990; May 1985). However, the contribution of data presenting children's interpretations of these sentence types, which are subsequently observed in adults as well, has been key to revisiting the influence (or lack thereof) of the tensed clause boundary, the range of possible readings generated by the grammar, and the status of the matrix reading for (13b).

The third case study is comparative constructions involving pronominal reference, as shown in (14a–b).

(14) a. More players encouraged him$_{i/*j}$ to talk to Alicia than to Brady$_j$'s teammate.
b. He$_{*i/j}$ gave more attention to the coach than to Brady$_j$'s teammate.

In both (14a–b), there is elided material between the head of the standard phrase signaled by *than* and *to* (roughly, in (14a): [d-players encouraged him to talk], and in (14b): [he gave d-much attention]). And in both instances, there is a pronoun in the site of the ellipsis. This pronoun – even while absent on the surface – stands in a particular syntactic (c-command) relation to the R-expression at the end of the sentence, leading it to be interpreted as not standing in a coreference relation with that individual (Bhatt & Takahashi, 2011; Lechner, 2004).[1] What's more, the elided pronoun should be interpreted in the same way as the overt pronoun preceding it. Interestingly, children appear to be unhindered by these constraints. Thus, there appears to be something happening behind the scenes with these elided pronouns, leading us to question children's representations of these comparative constructions and speculate about how this unpronounced pronoun is being interpreted.

In the sections that follow, I will present these three case studies in turn. For each, I will summarize the theoretical background in slightly more detail than above and present relevant evidence from experimental investigations in child language that speaks to children's interpretation of the target sentences and the nature of the child and adult grammar.

3. Case 1: Questions involving *wh*-phrases and universal quantifiers

3.1 Theoretical background

Our first case of special forces operations in acquisition comes to us from questions in which a *wh*-word interacts with universal quantifier, as shown in (12), repeated here. In (12a), the quantifier is in subject position, while in (12b), it is in object position. This difference in syntactic position is critical for the availability of certain answer types.

1. Follow-up psycholinguistic investigations with adults by Gor and Syrett in which prosody has been manipulated reveals that a significant percentage of participants (and I) find (a) acceptable under a co-construal reading, given focus on *Alicia* and *teammate*, thereby giving rise to a so-called 'grammatical illusion' that flies in the face of what structural relations would predict is possible.

(12) a. Which musical piece did [every/each] gymnast select?
b. Which gymnast performed [every/each] routine?

As stated earlier, (12a) allows for both a 'single answer' and a 'pair-list answer', as shown in (15).

(15) Question: Which song did [every/each] gymnast select?
Single Ans.: song X
Pair-list Ans.: gymnast A selected song X, B selected Y, and C selected Z.

(12b), however, only allows for the pair-list answer when the quantifier is *each*, although some speakers might consider this judgment marginal.

(16) Question: Which gymnast performed [every/each] routine?
Single Ans.: gymnast X
Pair-list Ans.: gymnast A performed routine X, B performed Y, and C performed Z.

This difference between the subject and object quantifier questions arises, because in order to generate the pair-list reading, the universal quantifier must take scope over the *wh-* phrase. This can happen if the phrase undergoes covert movement via QR to take scope over the *wh-* phrase, as discussed earlier. Thus, children's ability to provide or accept a pair-list answer is evidence of the presence of QR in their grammar. However, as we have also discussed, the *wh-* phrase also moves – overtly and prior to the covert movement – to the specifier of the Complementizer Phrase.

Under one account (May 1985, 1988), the problem is that in the subject quantifier questions, the path of movement of the quantified phrase is nested within the movement of the *wh-* phrase in object position to spec, CP, as shown in (17). However, with the object quantifier questions, the paths cross, as shown in (18).

(17) Which song did [every/each] gymnast select?

$[_{CP}\text{ Which song}_i\ [_{IP}\text{ every gymnast}_j\ [_{TP}\ t_j\ [\text{select}\ [_{DP}\ t_i]]]]]$

(18) Which gymnast performed every routine?

$[_{CP}\text{ Which gymnast}_j\ [_{IP}\text{ every routine}_i\ [_{TP}\ t_j\ [\text{performed}\ [_{QNP}\ t_i\ ]]]]]$

Under another account (Chierchia, 1993), the problem is not that the paths cross, but rather that the quantified object phrase must cross over the pronominal trace left by the *wh-* phrase, which gives rise to a Weak Crossover Violation (Postal, 1971; Wasow, 1972).

However, this cannot be the whole story, since object quantifier questions do seem to allow for pair-list answers when the quantifier is *each*. Agüero-Bautista (2001) accounts for this state of affairs by arguing that the QNP headed by *each* QRs to a higher position than one headed by *every* (Beghelli, 1997; Beghelli & Stowell, 1997; Szabolcsi, 1997), a reflection of its strong distributivity (Tunstall, 1998; Vendler, 1962). As a result, the *wh*- phrase in subject position can take scope under the quantifier, even after reconstructing to its base position inside the VP.

So, the state of affairs is this: subject quantifier questions allow both single answers and pair-list answers for both *every* and *each* QNPs, while object quantifier questions allow single answers for both *every* and *each* QNPs but only pair-list answers with *each* QNPs. If children access pair-list answers, they must be accessing a representation in which the quantifier outscopes the *wh*- phrase. Furthermore, if they allow for pair-list answers in certain instances (with subject quantifier questions and with *each* object quantifier questions), but not in other (with *every* object-quantifier questions), then this would demonstrate a knowledge of the structural and lexical factors that influence the availability of certain question-answer pairings.

3.2 Evidence from child language acquisition

Previous research by Roeper and de Villiers (1993) showed that children overproduce pair-list answers in response to both subject and object quantifier questions. However, there is a catch: Roeper and de Villiers had included *who* as the *wh*-phrase in their linguistic stimuli, and since *who* is not marked for number, it could have induced pair-list answers by leaving open the possibility that it referred to a plurality. Moreover, the scenarios employed by Roeper and de Villiers seemed to have made the single answer difficult to access – an observation followed up on in experimental work by Yamakoshi (2002). In turn, Yamakoshi's research highlighted the fact that responses provided by children should be analyzed carefully, as they may be constrained by production and/or subject to a conversational maxim to be concise. That is, responses that are not complete pair-list answers as in (16) may as yet be (attempts at) pair-list answers and not single answers. Thus, the previous literature had left open the question of whether or not children do in fact overproduce pair-list answers, and whether or not they display a subject-object asymmetry, but had not addressed the question of whether the type of universal quantifier interacting with the *wh*-phrase made a difference on the rate of pair-list answers.

Achimova et al. (2013) and Achimova et al. (2017) were interested in investigating precisely these questions. In our experimental paradigm, participants (adults and children age four to five) viewed a series of animated stories, each of which was followed by a question about the events that had transpired, posed by a puppet. In each story, there were three characters and three objects. Each of the three characters interacted with one of the objects, and then two moved on separately to interact with one of the two other objects (e.g., friends playing board games, or eating different flavors of ice cream). Target questions were subject and object quantifier questions of the form in (12), presented within a session, and participants were randomly assigned to a condition in which they heard questions with either *every* or *each*. Children's response to the puppet's question was transcribed and analyzed for its status as a single or pair-list answer. Given previous findings that children often do not produce complete pair-list answers, non-single answers were further subcategorized based on whether they were (non-) exhaustive, involved pairings or lists of single entities, or listed the correct or incorrect entities.

Adults consistently produced pair-list answers to subject and object quantifier questions with *each*, but single answers to subject and object quantifier questions with *every* (even though pair-list answers were licensed when the quantifier appeared in subject position). Thus, adults did not exhibit a subject-object asymmetry. However, since they had the choice to respond with either a single or a pair-list answer in some circumstances, a choice to respond with a single answer with *every* does not mean that a pair-list answer is barred. What is interesting is that the pair-list answer was produced for object-quantifier questions with *each*. The pattern displayed by children was strikingly different: children allowed pair-list answers for <u>both</u> *every* and *each* in subject position, and preferred single answers for <u>all</u> object quantifier questions. Thus, relative to adults, they *over*produced pair-list answers in response to *every* subject quantifier questions and *under*produced pair-list answers in response to *each* object quantifier questions.

The emerging picture from Achimova et al. (2013) and Achimova et al. (2017) is that children are sensitive to the structural position of the quantifier and exhibit a subject-object asymmetry, but do not appear to be sensitive to the type of quantifier – and specifically, its potential requirement of strong distributivity. Thus, they allow the quantifier and *wh*-phrase to interact in the abstract representation in such a way as to license a pair-list answer, interacting with the structural constraints, but lack the lexical knowledge that adults have about differences among universal quantifiers. The path of language development must then involve children enriching the lexical entries for *each* in such a way that not only distinguishes it from *every*, but that means that its strongly distributive nature allows it to give rise to pair-list answers in both subject and object position. The next case study complements this one in illustrating the similarity between children and adults in their syntax-semantics representations.

4. Case 2: Embedded antecedent-contained deletion (ACD)

4.1 Theoretical background

The second case of special forces operations comes from research on embedded ACD as illustrated in (13), repeated here.

(13) a. Shawn [wanted to [win [every event that Nastia *did*]]].
Embedded reading: *did* = win…
Matrix reading: *did* = wanted to win…
b. Shawn [said that she [won [every event that Nastia *did*]]].
Embedded reading: *did* = won…
Matrix reading: *did* = said that she won…

As described above, the QNP *every event that Nastia did* must undergo QR for reasons of type mismatch, but in order for the instance of VPE to be interpretable, the QNP hosting the ellipsis site must raise out of its antecedent in order to be parallel in form to that antecedent and copy in the information provided by the antecedent for interpretation.

In both (13a) and (13b), there are two VPs (one embedded in the other), and therefore (in principle) two possible antecedents. In order to target the matrix VP as the antecedent, the QNP would have to move out of both the embedded and the matrix antecedent so that it is no longer contained in its antecedent. Now, in (13a), ACD is embedded in a non-finite clause, while in (13b), ACD is embedded in a finite, or tensed, clause. The problem that seems to be presented by the latter situation is that the tensed clause has been claimed to present a boundary to covert movement (Farkas, 1981; Hornstein, 1994; Larson & May, 1990; May, 1985). Thus, while in both (13a) and (13b) the embedded antecedent can be targeted, it has been thought that the matrix antecedent can only be targeted in (13a), since in (13b), the QNP cannot escape the matrix VP and therefore cannot target it as an antecedent.

4.2 Evidence from child language acquisition

Syrett and Lidz (2011) were interested in investigating whether children are able to access both the embedded and matrix readings of sentences such as these, with the goal of determining whether children's QR operation was constrained or not. The prediction was that if children can only target the shortest, or nearest, landing site, then they should only be able to access the embedded reading of such sentences. Under the theoretical assumption that the matrix reading is available for when ACD is embedded in a non-finite clause and not available when it is embedded in a finite clause, the prediction would be that the matrix reading would be available in the former case, but not the latter.

Syrett and Lidz administered a truth value judgment task (Crain & Thornton, 1998) to both adults and children. In each trial, an experimenter acted out a scenario with toys, at the end of which a puppet (who had watched the story unfold alongside the child) delivered a target statement, which the children were asked to evaluate. If the puppet's statement was correct, he was rewarded with his preferred sweet; if it was incorrect, he was given something less desirable. There were two experiments involving embedded ACD. In Experiment 1, participants heard sentences with the structure in (13a); in Experiment 2, they heard sentences similar to (13b). In each, participants were randomly assigned to a condition in which either the embedded or the matrix reading of the target sentence was favored in the story.

In an example story in Experiment 1, Kermit the Frog has two set of cars: an old set he has driving many times and is tired of driving, and a newer set he has just received for his birthday and is not yet allowed to drive. Miss Piggy is interested in his cars, and asks him to drive them. Kermit complains, because in order to oblige, he is stuck driving the old cars, which he does not want to drive. He drives each of these grudgingly. Thus, participants see that there are cars that Kermit drives, but does not want to drive (the old ones), and cars that he wants to drive, but does not drive (the new ones). After Kermit drives the cars, he gives Miss Piggy the choices of which cars she would like to drive. In the condition favoring the embedded VP reading, she expresses an interest in the old cars, and in the condition favoring the matrix VP reading, she expresses an interest in the new cars. (19) illustrates this difference.

(19) Miss Piggy wanted to drive every car that Kermit did.
  a. Embedded VP: ⟨*drive*⟩
  b. Matrix VP: ⟨*want to drive*⟩

In a minimally different experiment, after Kermit drives the cars, their friend Fozzie Bear expresses an interest in driving a set of cars. They decide to play a game in which Miss Piggy and Kermit will leave for a few minutes, and in their absence Fozzie will drive a set of cars, Miss Piggy and Kermit will return, and they will then guess which cars Fozzie drove. Participants never actually see Fozzie drive any cars. When Miss Piggy and Kermit return, Kermit guesses first, and says that Fozzie drove the cars that Kermit did not drive (the new ones). The two conditions then diverge based on what Miss Piggy guesses. In the condition favoring the embedded VP reading, Miss Piggy guesses that Fozzie drove the cars Kermit drove (the old ones), while in the condition favoring the matrix VP reading, she guesses that he drove the cars that Kermit also guessed that he drove (the new ones). The target sentence and the two possible interpretations are illustrated in (20).

(20) Miss Piggy said that Fozzie drove every car that Kermit did.
a. Embedded VP: ⟨*drive*⟩
b. Matrix VP: ⟨*say that Fozzie drove*⟩

Remarkably, Syrett and Lidz found that not only were children able to access both the embedded and matrix reading for (19), but they were also able to do so for (20). Adults also accessed both the embedded and matrix reading for (19), and largely resisted the matrix reading in favor of the embedded reading in (20), but a small cohort of adults did access this reading. Here lies the conundrum: if children are overly permissive in their QR and allow it to move beyond a tensed clause boundary to access a matrix reading that is barred by the grammar, they are over-generating interpretations, and producing a superset of those produced by adults. This poses a challenge from the perspective of learnability (Berwick, 1985; Gold, 1967; Manzini & Wexler, 1987; Pinker, 1989). At some point, children would have to learn that this reading is ungrammatical, and prune it from their grammar to be more conservative, but what kind of evidence could lead them to filter out a matrix reading of ACD sentences? How often do they encounter such sentences, and what would be the universally available evidence for this shift? This pattern of results lead Syrett and Lidz to entertain the possibility that there is not, in fact, an arbitrary clause-boundedness constraint on QR.

This possibility was pursued in subsequent work exclusively targeting adult participants, in order to probe the availability of the matrix reading of sentences in which ACD was embedded in a finite clause, which should impose a barrier to movement, if previous theoretical claims are correct. Already, however, there was a hint that this might not be the case, for reasons reviewed in Syrett (2015a, b). Syrett (2015b) administered an experiment that was similar in design to those with a finite clause presented in Syrett and Lidz (2011), with a few key changes in design. First, since the experiment was geared towards adults, participants were presented with animations on slides, rather than toys (still accompanied by narration). Second, the sentences were modified in order to facilitate the processing of them: the complementizer preceding the embedded clause was removed, and a pronoun appeared as the embedded clause subject instead of a proper name, as shown in (21).

(21) Woody said $\boxed{\text{ø he}}$ jumped over every frog that Jessie did

Finally, and crucially, elements of the discourse signaled the focus on the reporting (i.e., what the characters guessed), rather than the actual actions (e.g., the jumping events).

This time, and in two separate experiments, adults robustly accessed the matrix reading of the target sentences. While the rate of accessing this reading was as low as 37% in some trials, it was high as 95% in others. Indeed, across the majority of cases, adult participants accessed the matrix reading, and provided unambiguous justifications for why they did so, in the context of the story. These findings were elaborated upon as Syrett (2015a) pushed the boundaries even further, investigating whether adults could not only allow the QNP to raise out of the matrix VP, but travel even further to scope over an indefinite in subject position, as in (22) – a possibility spelled out in theoretical work by Cecchetto (2004).

(22) Someone said $\boxed{\text{ø he}}$ could jump over every frog that Jessie did

Here, too, adults demonstrated a robust ability to access the reading that would require the QNP to have raised out of the matrix VP to take scope not only over the matrix verb, but also the indefinite subject. It appears that the manipulations of the linguistic stimuli (i.e., the lack of complementizer and type of embedded subject), and the structured support for the matrix and extra wide scope readings in their respective discourse contexts influenced the accessibility of these readings for adults, who might otherwise be constrained by reasons of economy to raise the QNP only as far as it would need to go to generate a licensed interpretation (i.e., perform the 'shortest move') (Fox, 1995a, 2000).

Thus, what started as a query into the nature of QR in the child grammar transformed into a series of experiments in which robust empirical evidence indicated that we should abandon the assumption that there is an arbitrary boundary imposed by tense preventing QR out of a finite clause. Thus, it appears that the covert operations launched in the child grammar are highly similar in nature to those in the adult grammar, and there is no learnability challenge to be addressed. Here, the judgments of four-year-olds prompted a series of experiments resulting not only in a simplified picture of language acquisition, but also a better understanding of the adult grammar.

The third case study, however, does just the opposite. It highlights a situation in which children unexpectedly diverge from adults in the interpretations they access, and demonstrate that there may be a clandestine operation available in their grammar, which they perform, which adults are not performing in the same sentences. But of course, since this is a clandestine operation, the evidence for it is not immediately apparent, because it is very cleverly orchestrated behind the scenes.

5. Case 3: Pronominal reference in comparatives

5.1 Theoretical background

The third case study thus brings us to evidence for a clandestine operation: one where you have no idea that it is being carried out, or by whom. Such a case comes to us from children's interpretations of comparative constructions involving pronominal reference, as in (23). In each of these examples, there is a pronoun in the matrix clause. In (a) and (b), it appears in the object position, and in (c), it appears as the subject.

(23) a. More blocks connected **him**$_{i*/j}$ to Minnie than to Flynn$_i$'s horse.
b. Nemo delivered more presents from **him**$_{i*/j}$ to Flounder than to Eric$_i$'s dog.
c. **She**$_{*i/j}$ gave more cones to Winnie-the-Pooh than to Sleeping Beauty$_i$'s godmother.

As is common in comparatives, there is material elided in the standard clause, as is (albeit roughly) illustrated in (24). The content of this elided material is crucial for the interpretation of the comparative construction.

(24) a. More blocks connected **him**$_{*i/j}$ to Minnie than
[blocks connected **him** $_{=\text{ matrix him}}$] to Flynn$_i$'s horse
b. Nemo delivered more presents from **him**$_{i*/j}$ to Flounder than
[Nemo delivered presents from **him** $_{=\text{ matrix him}}$] to Eric$_i$'s dog.
c. **She**$_{*i/j}$ gave more cones to Winnie-the-Pooh than
[**she** $_{=\text{ matrix she}}$ gave cones] to Sleeping Beauty$_i$'s godmother.

In each of these instances, the position of the pronoun in the elided clause is such that it 'c-commands' the name (R-expression) in the standard clause. This relation prevents the pronoun and the R-expression from being coreferential, since the R-expression cannot be bound by a pronoun c-commanding it (Chomsky, 1981; Lasnik, 1976; Reinhart, 1976). That is not the case with the sentences in (25), since the elided pronoun does not c-command the R-expression.

(25) a. More lambs walked from Belle to him$_{i/j}$ than
[lambs walked] from Harris$_i$'s brother [to him $_{=\text{ matrix him}}$].
b. King Triton gave more lizards to her$_{i/j}$ than
Olivia$_i$'s mother [gave lizards to her $_{=\text{ matrix her}}$].

In addition, as indicated in each instance above, the pronoun in the elided standard material should be interpreted in exactly the same way as the pronoun in the matrix clause.

Now, this is not always the case with pronouns appearing in ellipsis. In the case of VPE in which a pronoun appears, there is the possibility of either strict or sloppy identity, as shown in (26).

(26) Olivia$_i$ gave a present to her$_i$ mother, and Minnie$_j$ did, too.
 a. strict identity: …and Minnie$_j$ gave a present to her$_i$ mother, too
 b. sloppy identity: …and Minnie$_j$ gave a present to her$_j$ mother, too

It is not entirely clear what disallows similar 'sloppy' reference in the case of the comparative constructions above. It is the case that the structures in (24) differ from those typically appealed to in order to illustrate the phenomenon of sloppy identity; in the latter, the R-expression binds the pronoun, as in (26). In addition, in (24) and (25), there is no 'functional' relationship, as there is in (26) for 'mother of x'. However, even when both do hold, there is still a strong preference at least that the pronouns be interpreted parallel, which seems to be stronger when the VPE is *not* overtly signaled in the sentence, as shown in (27).

(27) a. Olivia$_i$ gave more presents to her$_i$ mother than Minnie.
 b. Olivia$_i$ gave more presents to her$_i$ mother than Minnie did.

Regardless of whether or not the two pronouns *must* always be interpreted the same way in any or all of these cases, the examples in (23) and (25) make it clear that this phenomenon goes beyond classical instances of sloppy identity, and that the pronoun absolutely cannot be interpreted as freely as it might in other cases of elided content. Returning to the issue of learnability, it is also not clear what would indicate to children acquiring language that such 'sloppy' reference is allowed with VPE but something like that is not in these other comparatives. In fact, this is precisely the puzzle that arises when we consider the results of an experiment conducted by Gor & Syrett (2015) with children and adults.

5.2 Evidence from child language acquisition

Gor and Syrett (2015) were initially interested in determining whether children and adults performed as predicted by the theoretical literature and rejected coreference in sentences where the elided pronoun c-commands the R-expression in the standard. However, an unexpected pattern surfaced in the behavioral responses of the child participants in our experiments, which bears on how the elided material was interpreted. We will therefore focus less here on the results that bear on the question of c-command and coreference (although we will give a nod to them), and instead pay closer attention to the pattern of responses that bears on the interpretation of the elided pronoun and whether it would be coindexed with the overt pronoun.

To determine how children (age four and a half to six) and adults interpret the target sentences in (23) and (25) above, and whether structural constraints guided these interpretations, Gor and Syrett administered an experimental task incorporating an act-out procedure. Each test trial was set up as a truth value judgment task right up until the delivery of the target sentence. One experimenter set some props on the table in front of the participant, and told a brief story about them. A puppet (played by a second experimenter) watched the story alongside the child. Each set of props and each story involved four characters, two of whom were intended to be made salient as possible referents for the pronoun appearing in the target sentence. At the end of the story, the experimenter delivered the target sentence. Then, instead of having the participant assess the truth value of the proposition expressed by this utterance in the context, the puppet repeated the target sentence, and asked the child to configure the props so that they made the target sentence true. (Adults only interacted with an experimenter, not a puppet.)

For example, for sentence (23c) repeated here as (28), the story involved the friends Sleeping Beauty and Hello Kitty who had some toy cones, which they decided to distribute to Winnie-the-Pooh and Sleeping Beauty's godmother. There were therefore two salient candidate antecedents for *she* (Sleeping Beauty and Hello Kitty).

(28) She$_{*i/j}$ gave more cones to Winnie-the-Pooh than to Sleeping Beauty$_i$'s godmother.

Because of the structural c-command relation in the sentence, *she* and *Sleeping Beauty* cannot be coreferential, since the pronoun in subject position c-commands the R-expression in the standard, so *she* must be coreferential with Hello Kitty, and not with Sleeping Beauty. Since the elided pronoun must be interpreted in the same way as the overt pronoun, the elided *she* should be also interpreted as Hello Kitty. Thus, the scene acted out in the end should have it so that Hello Kitty gives more cones to Winnie-the-Pooh than she gives to Sleeping Beauty's godmother.

Unexpectedly, across trials and a significant percentage of the time, children – but not adults – acted out the final scene so that both Sleeping Beauty and Hello Kitty gave out their cones, and the number of cones that Hello Kitty gave to Winnie-the-Pooh was greater than the number that Sleeping Beauty gave to her godmother, or in a way that the number of cones that Sleeping Beauty gave to Winnie-the-Pooh was greater than the number that Hello Kitty gave to Sleeping Beauty's godmother (an configuration that was surprising, since it violates c-command constraints)). In some cases, when props were re-set, and children were asked to act out the sentence again, they provided clear justifications as they placed the props on the table in a particular configuration, indicating that the interpretation they were accessing was similar to the one captured in (29) or (30).

(29) She$_{\text{= Hello Kitty}}$ gave more cones to Winnie-the-Pooh than
[she$_{\text{= Sleeping Beauty}}$ gave cones] to Sleeping Beauty's godmother.

(30) She$_{\text{= Sleeping Beauty}}$ gave more cones to Winnie-the-Pooh than
[she$_{\text{= Hello Kitty}}$ gave cones] to Sleeping Beauty's godmother.

Indeed, as Gor and Syrett noted, this response pattern was not rare by any means: 19 of the 26 child participants acted out such a scene at least once out of the five test trials. The adults did not do so.

In a follow-up experiment with 18 new children, the experimenters modified the story slightly so that it was clear that only one of the two characters should be distributing the cones (or performing whatever action was indicated in the story). Still, approximately half of the children still provided what Gor and Syrett have termed the 'two-giver response'. That is, even though the story made clear that only *one* of the two girls should distribute the cones, children still made *both* characters do so, and compared the quantities across the two instances of giving by different givers to different recipients.

Had we not administered an act-out task, we never would have known that such an interpretation of these sentences existed. Linguistic theory would not have led us to predict this response pattern. Moreover, without a methodology such as this, many other frequently-used methodologies in child language research would have masked this response pattern. A forced choice task that targets c-command as a contributor to coreference constraints would have compared two scenes to determine the referent for the pronoun (one in which Sleeping Beauty gave more cones to Winnie-the-Pooh than to her Godmother and another in which Hello Kitty gave more cones to Winnie-the-Pooh than to Sleeping Beauty's Godmother), and never would have included the 'two-giver' scene as an option. Even if it had, if children had selected one of the other scenes instead of this one, that would not tell us that this interpretation was unavailable – just that it was dispreferred relative to the other scenes. A truth value judgment task would have favored the ungrammatical response (Sleeping Beauty giving more cones to Winnie-the-Pooh than to her Godmother) in order to elicit rejections from the children, and also never would have involved the two-giver scene as a comparison of quantities *across* rather than *within* subjects. Appealing to the terminology of this chapter, one could say that we would have had no idea that children are capable of conducting such a clandestine operation, had we not convinced them to leak their intelligence in an act-out task.

In this case study, we have thus observed children accessing an unexpected interpretation relative to the adult grammar, leading us to ponder exactly what the nature of their representation is, and how they arrive at the adult state. The question arising from the pattern exhibited by the children is what allows their grammar to generate such a response. Gor and Syrett entertained one possibility: that instead of a coindexed *pronoun* in the elided material, children's representation instead has a *function* in this slot, mapping individuals to individuals, each represented by a variable. Such a representation, modeled after cases of functional, or E-type, pronouns (Evans, 1977, 1980; Elbourne, 2001, 2005; Heim, 1990), would allow the interpretation of the recipient to co-vary with the giver. Thus, the situation is similar in interpretational output to sloppy identity, but does not arise for the same reasons. If indeed children are accessing such a representation, we must ask what prompts the developmental shift in the interpretation of pronouns, and what bars such functional pronouns and 'sloppy' reference in the adult representation.

6. Conclusion

I began this paper by appealing to terminology from U.S. special forces – overt, covert, and clandestine operations – in order to discuss a particular set of case studies from child language studies that employ one or more of these operations. My goal was to demonstrate that investigating how these operations are carried out in the child grammar is beneficial not only from the point of language acquisition, but also from the point of better understanding the adult grammar (the end state).

The first case study was an investigation into whether children produce single or pair-list answers in response to *wh*-questions involving a universal quantifier – sentences that involved both overt and covert operations. There we saw that children demonstrate an expected structural subject-object asymmetry in their response pattern and that adults allowed pair-list answers not only with subject quantifier questions, but also those with *each* as the object quantifier. Both findings are in line with the theoretical predictions. Where children diverged from adults, however, was in the treatment of *each*, where they undergenerated pair-list answers. Thus, we learned that children's lexical representations of this quantifier are not yet adult-like in this realm.

The second case study was an investigation of children's ability to access multiple interpretations of sentences with ACD embedded in either a non-finite or finite clause – sentences involving multiple covert operations. Children's ability to access seemingly ungrammatical interpretations posed a learnability problem, prompting further investigations into adult interpretations. Those subsequent experiments revealed that adults, too, access such interpretations, thus leading to revisiting the state of the relevant theory and the status of supposed constraints on movement. The combined pattern of responses across children and adults presents a more fine-grained picture of the adult grammar and a more parsimonious description of the path of language development.

The final case study was an investigation of children's interpretation of comparative constructions featuring an elided pronoun that stood in a particular relation to an R-expression in the standard clause – sentences that involved covert and clandestine operations. Here, children demonstrated an ability to access a rather unexpected interpretation, leading us to ask how they represent elided material, and how they could come to be adult-like. Unlike the last case study, these results *do* pose a learnability puzzle: how do children come to represent elided pronouns correctly, and impose limits on 'sloppy' reference or functional interpretation?

Combined, these studies illustrate two main take-away points. First, there are abstract grammatical mechanisms at work in children's grammar, which allow them to generate a host of interpretations. These mechanisms, which appear to be largely adult-like, interact with other lexical and structural material. Second, experiments in language acquisition have value beyond determining the nature of the child grammar and the path of language development. They can also lead to a better understanding of the adult grammar, and our linguistic competence.

References

Achimova, A., Syrett, K., Musolino, J., & Déprez, V. (2017). Children's developing knowledge of *wh-*/quantifier question-answer relations. *Language Learning and Development*, 13, 80–99.

Achimova, A., Crosby, C., Déprez, V., Syrett, K., & Musolino, J. (2013). Which account of *wh*-quantifier interaction should everyone adopt? A new take on a classic developmental puzzle. In S. Baiz, N. Goldman, & R. Hawkes (Eds.), *Proceedings of the 37th annual Boston University Conference on Language Development* (pp. 1–12). Somerville, MA: Cascadilla Press.

Agüero-Bautista, C. (2001). Cyclicity and the scope of wh-phrases (Unpublished doctoral dissertation). MIT.

Beghelli, F. (1997). The syntax of distributivity and pair-list readings. In A. Szabolcsi (Ed.) *Ways of scope taking* (pp. 349–408). Dordrecht: Kluwer. https://doi.org/10.1007/978-94-011-5814-5_10

Beghelli, F., & Stowell, T. (1997). Distributivity and negation. In A. Szabolcsi (Ed.) *Ways of scope taking* (pp. 71–107). Dordrecht: Kluwer. https://doi.org/10.1007/978-94-011-5814-5_3

Berwick, R. C. (1985). *The acquisition of syntactic knowledge*. Cambridge, MA: The MIT Press.

Bhatt, R., & Takahashi, S. (2011). Reduced and unreduced phrasal comparatives. *Natural Language and Linguistic Theory*, 29, 581–620. https://doi.org/10.1007/s11049-011-9137-1

Cecchetto, C. (2004). Explaining the locality conditions of QR: Consequences for the theory of phases. *Natural Language Semantics*, 12, 345–397. https://doi.org/10.1007/s11050-004-1189-x

Chierchia, G. (1993). Questions with quantifiers. *Natural Language Semantics*, 1, 181–234. https://doi.org/10.1007/BF00372562

Chomsky, N. (1981). *Lectures on government and binding*. Dordrecht: Foris.

Crain, S., & Thornton, R. (1998). *Investigations in Universal Grammar: A guide to research on the acquisition of syntax and semantics*. Cambridge, MA: The MIT Press.

Elbourne, P. D. (2001). E-Type anaphora as NP-deletion. *Natural Language Semantics*, 9, 241–288. https://doi.org/10.1023/A:1014290323028
Elbourne, P. D. (2005). *Situations and individuals*. Cambridge, MA: The MIT Press.
Evans, G. (1977). Pronouns, quantifiers, and relative clauses (I). *Canadian journal of Philosophy*, 7(3), 467–536.
Evans, G. (1980). Pronouns. *Linguistic Inquiry*, 11, 337–362.
Farkas, D. (1981). Quantifier scope and syntactic islands. In R. A. Hendrick, C. S. Masek, & M. F. Miller (Eds.), *Papers from the Seventeenth Regional Meeting of the Chicago Linguistic Society* (pp. 59–66). Chicago, IL: CLS.
Fiengo, R., & May, R. (1994). *Indices and identity*. Cambridge, MA: The MIT Press.
Fox, D. (1995a). Economy and scope. *Natural Language Semantics*, 3, 283–341. https://doi.org/10.1007/BF01248820
Fox, D. (1995b). Condition C effects in ACD. In R. Pensalfini & H. Ura (Eds.), *Papers on minimalist syntax* (pp. 105–120). Cambridge, MA: MITWPL.
Fox, D. (2000). *Economy and semantic interpretation*. Cambridge, MA: The MIT Press.
Gold, E. M. (1967). Language identification in the limit. *Information and Control*, 10, 447–474. https://doi.org/10.1016/S0019-9958(67)91165-5
Gor, V., & Syrett, K. (2015). Picking up after sloppy children: What pronouns reveal about children's analysis of English comparative constructions. In E. Grillo & K. Jepson (Eds.), *Proceedings of the 39th annual Boston University Conference on Language Development* (pp. 191–203). Somerville, MA: Cascadilla Press.
Hackl, M. (2000). Comparative quantifiers (Unpublished doctoral dissertation). MIT.
Hankamer, J., & Sag, I. 1976. Deep and surface anaphora. *Linguistic Inquiry*, 7, 391–428.
Heim, I. (1985). Notes on comparatives and related matters (Unpublished manuscript). University of Texas, Austin.
Heim, I. (1990). E-type pronouns and donkey anaphora. *Linguistics and Philosophy*, 13, 137–177. https://doi.org/10.1007/BF00630732
Heim, I., & Kratzer, A. (1998). *Semantics in generative grammar*. Oxford: Blackwell.
Hornstein, N. (1994). An argument for minimalism: The case of antecedent-contained deletion. *Linguistic Inquiry*, 25, 455–480.
Kennedy, C. (1997). Antecedent-contained deletion and the syntax of quantification. *Linguistic Inquiry*, 28, 662–688.
Larson, R., & May, R. (1990). Antecedent containment or vacuous movement: Reply to Baltin. *Linguistic Inquiry*, 21, 103–122.
Lasnik, H. (1976). Remarks on coreference. *Linguistic Analysis*, 2, 1–21.
Lechner, W. (2004). *Ellipsis in comparatives*. Berlin: Mouton de Gruyter. https://doi.org/10.1515/9783110197402
Manzini, M. R., & Wexler, K. (1987). Parameters, binding theory, and learnability. *Linguistic Inquiry*, 18, 413–444.
May, R. (1977). The grammar of quantification (Unpublished doctoral dissertation). MIT.
May, R. (1985). *Logical form: Its structure and derivation*. Cambridge, MA: The MIT Press.
May, R. (1988). Ambiguities of quantification and *wh*: A reply to Williams. *Linguistic Inquiry*, 19, 118–135.
Merchant, J. (2000). Antecedent-contained deletion in negative polarity items. *Syntax*, 3, 144–150. https://doi.org/10.1111/1467-9612.00029
Pinker, S. (1989). *Learnability and cognition: The acquisition of argument structure*. Cambridge, MA: The MIT Press.

Postal, P. M. (1971). Remarks on weak crossover effects. *Linguistic Inquiry*, 24, 539–556.
Reinhart, T. (1976). The syntactic domain of anaphora (Unpublished doctoral dissertation). MIT.
Reinhart, T. (1983). Coreference and bound anaphora: A restatement of the anaphora questions. *Linguistics and Philosophy*, 6, 47–88. https://doi.org/10.1007/BF00868090
Roeper, T., & de Villiers, J. G. (1993). The emergence of bound variable structures. In E. Reuland & W. Abraham (Eds.), *Knowledge and language, Vol. 1: From Orwell's Problem to Plato's Problem* (pp. 105–139). Dordrecht: Kluwer. https://doi.org/10.1007/978-94-011-1840-8_6
Ross, J. R. (1967). Constraints on variables in syntax (Unpublished doctoral dissertation). MIT.
Sag, I. 1976. Deletion and Logical Form (Unpublished doctoral dissertation). MIT.
Szabolcsi, A. (1997). Quantifiers in pair-list readings. In A. Szabolcsi (Ed.), *Ways of scope taking* (pp. 349–408). Dordrecht: Kluwer. https://doi.org/10.1007/978-94-011-5814-5_9
Syrett, K. (2015a). Experimental support for inverse scope readings of finite-clause embedded Antecedent-Contained Deletion sentences. *Linguistic Inquiry*, 46, 579–592. https://doi.org/10.1162/LING_a_00194
Syrett, K. (2015b). QR out of a tensed clause: Evidence from Antecedent-Contained Deletion. In N. Nansen & E. Borg (Eds.), *Investigating Meaning*. Special issue of Ratio, 28, 395–421.
Syrett, K., & Lidz, J. (2011). Competence, performance and the locality of Quantifier Raising: Evidence from 4-year-old children. *Linguistic Inquiry*, 42, 305–337. https://doi.org/10.1162/LING_a_00043
Tunstall, S. L. (1998). The Interpretation of quantifiers: Semantics and processing (Unpublished doctoral dissertation). University of Massachusetts, Amherst.
Vendler, Z. (1962). Each and every, any and all. *Mind: A Quarterly Review of Psychology and Philosophy*, LXXI, 145–160. https://doi.org/10.1093/mind/LXXI.282.145
Wasow, T. (1972). *Anaphoric relations in English* (Unpublished doctoral dissertation) MIT. Revised version published in 1979 as Anaphora in Generative Grammar. Ghent: Story-Scientia.
Yamakoshi, K. (2002). The acquisition of wh/every interaction in English. In B. Skarabela, S. Fish & A. H-J. Do (Eds.), *BUCLD 26: Proceedings of the 26th annual Boston University Conference on Language Development* (pp. 769–780). Somerville, MA. Cascadilla Press.

SECTION V

The relation between semantics and pragmatics

CHAPTER 13

Developmental insights into gappy phenomena

Comparing presupposition, implicature, homogeneity, and vagueness

Lyn Tieu[i,ii], Cory Bill[ii], Jérémy Zehr[iii], Jacopo Romoli[ii,iv] and Florian Schwarz[iii]

[i]Western Sydney University / [ii]Centre of Excellence in Cognition and its Disorders, Macquarie University / [iii]University of Pennsylvania / [iv]Ulster University

In natural language, we encounter various sentence types that, under certain circumstances, are evaluated as neither true nor false. For instance, it is intuitively difficult to assess the *truth value* of a sentence whose *presupposition* is not satisfied in the context. A common theoretical approach is to characterize the status of such sentences with a third value of one kind or another. In this chapter, we consider children's acquisition of four linguistic phenomena that can give rise to 'gappy' judgments that correspond neither to True nor False: scalar implicature, presupposition, homogeneity, and vagueness. We discuss how young children's interpretations of such sentences can provide insight into how these phenomena should be treated within semantic theories.

Keywords: scalar implicature, presupposition, homogeneity, vagueness, truth value gaps

1. Introduction

In the formal study of meaning, the notion of truth conditions – the conditions under which a sentence is true – plays a crucial role: to know the meaning of a sentence like (1) is to know under what conditions (1) would be true. A speaker of English can be expected, for instance, to recognize that (1) is true in a context in which it is in fact raining and false if it is not. Such a speaker can provide a *truth value judgment* for (1) on the basis of its *truth conditions* and knowledge about the situation in which (1) is evaluated.

(1) It is raining.

https://doi.org/10.1075/tilar.24.13tie

In order to understand whether children understand sentences in an adult-like way, one common methodology involves targeting children's knowledge of the conditions that must hold in order for the sentence to be true. For instance, the Truth Value Judgment Task involves presenting young children with short stories, after which they must judge whether a sentence is true or not given the events that unfolded in the story (Crain & Thornton 1998, 2000). This task has been used successfully with children as young as three years of age, to test their knowledge of a range of syntactic and semantic phenomena.

But now consider the sentence in (2), which is superficially not a huge leap from (1).

(2) Jack knows it is raining.

Whether or not it is actually raining not only has a bearing on whether the sentence is true or false, but also on whether it can be uttered felicitously. The standard idea is that the sentence *presupposes* that it is raining, and *asserts* that Jack has knowledge of this state of affairs. In a context in which it is raining, we can evaluate whether (2) is true or false, depending on Jack's knowledge state. However, in a context in which it is not raining, the presupposition of the sentence is not satisfied, and it is intuitively more difficult to assess what the truth value of the sentence should be. A common theoretical approach to dealing with such cases of *presupposition failure* is to characterize the status of the sentence with a third value of one kind or another.

In addition to *presupposition failure*, there are various other phenomena in natural language that cause sentences to be neither clearly true nor clearly false in a given state of affairs. In this chapter, we will consider four such cases, and argue that young children's interpretations of sentences in such situations can provide insight into how these phenomena should be treated in a semantic theory of the adult grammar. We will begin by introducing the four phenomena, as well as an adult psycholinguistic study that will serve as a starting point for our discussion of the acquisition studies.

1.1 The phenomena

The phenomena we will address are presupposition, scalar implicature, homogeneity, and vagueness. Let us begin by considering examples of each that will be pertinent to our discussion of the acquisition studies. First, imagine a scenario in which Jack was a spectator at a race but never actually ran in the race. In such a scenario, the sentence in (3) is neither clearly true nor clearly false.

(3) *Context: Jack did not run in the race.*
Sentence: Jack stopped running.

Such cases are standardly treated as examples of *presupposition failure*. Assuming the verb *stop* in (3) triggers the presupposition that *Jack was running* previously, the sentence is neither clearly true nor false when the presupposition is not satisfied.

Another example of lack of clear Truth or Falsity, much more widely studied in the developmental literature, involves the use of scalar terms. While the literal meaning of the sentence in (4) is in principle compatible with a situation in which four out of four apples are red, adult speakers nevertheless generally find the sentence to be a dissatisfying description of the situation, which often leads to the rejection of the sentence.

(4) *Context: Four of four apples are red.*
Sentence: Some of the apples are red.

This is usually taken to be the case because the sentence in (4) triggers the *scalar implicature* that *Not all of the apples are red* (Grice 1975). When the literal meaning of the sentence is true but the implicature is false, an oddness arises (for discussion, see Magri 2009, 2014).

A third case involving lack of clear Truth or Falsity involves so-called *homogeneity effects* that arise from the use of plural definite descriptions. In a scenario where two out of four apples are red, neither the positive (5a) nor the negative (5b) appear to be clearly true, or clearly false.

(5) *Context: Two of the apples are red and two of the apples are green.*
a. *Sentence*: The apples are red.
b. *Sentence*: The apples aren't red.

Finally, certain instances of *vague* predicates also give rise to a similar effect. Vague predicates are typically described as having fuzzy boundaries and, as a consequence, it is possible to observe borderline cases within these fuzzy regions. As illustrated in (6), "big" appears to be a vague predicate (it is not clear what precise size is required to qualify as big), and a bear that is slightly above average-sized can constitute a borderline case for "big". In turn, the sentences in (6) are neither clearly true nor clearly false.

(6) *Context: The bear is slightly above average-sized.*
a. *Sentence*: The bear is big.
b. *Sentence*: The bear is not big.

The examples above provide us with four kinds of sentences that, in certain situations, do not correspond clearly to either of the two truth values True and False. Following terminology in Cremers, Križ, & Chemla (2015), we will refer to the contexts described above as 'gappy' contexts, and the sentences that they render neither true nor false, as 'gappy' sentences.

1.2 The starting point

Each of the phenomena outlined in Section 1.1 has been studied in great detail in previous theoretical research. While we describe them in a uniform way with respect to the 'gappiness' that they can give rise to, much of the existing theoretical research has investigated the phenomena independently of each other. One main reason to consider children's development of these phenomena together, however, stems from some recent theoretical attempts to unify subsets of them; for instance, Chemla (2009) and Romoli (2014) attempt to unify certain cases of presupposition and scalar implicature, Magri (2014) attempts to unify scalar implicatures with homogeneity, and Zehr (2014) explores potential unifications of presupposition and vagueness. These are areas where careful, theoretically informed empirical research can be highly informative, and indeed recent experimental research has offered new ways of empirically characterizing the relationships among these phenomena. Such empirical methods and data are useful for assessing whether some or all of the phenomena above should receive a unified treatment within linguistic theories.

One example of such empirical research is a study reported in Cremers, Križ, & Chemla (2015). These authors collected probability judgments from adult native speakers, using the treatment of gaps as a diagnostic for differentiating gappy phenomena (the paper, dataset, and R analysis script for this study are available online at: http://semanticsarchive.net/Archive/DZjNWY0N/Cremers-Kriz-Chemla-ProbasAndGaps.html). For example, participants would see contexts like (7a), represented in the form of three cards: one card would contain a yellow square, another would contain a green square, and the third would contain an orange circle. Participants were told that one of these three cards would be selected at random. The question was what the probability would be that the randomly selected card would correspond to the one described by the test sentence, e.g., (7b). In our example in (7), the presuppositional sentence in (7b) would be clearly *true* of the card with the yellow square, and it would clearly be *false* if describing the card with the green square. Importantly, the sentence would be *gappy* if it were meant to describe the card with the non-square shape: the orange circle.

(7) a. *Context: The card will either contain a yellow square, a green square, or an orange circle.*
b. *Sentence*: The square is yellow.

Participants were given a multiple choice task in which they had to select the probability of the sentence being true for the randomly selected card. The participants' choice of probability would crucially reveal how the participant treated the gap case, i.e. the orange circle in (7). For instance, if the participant decided the probability of selecting a yellow square was 1/3, then the gap case of the orange circle counted as a failure to make (7b) true; that is, the sentence was considered *false* of the orange circle, just as it was false of the green square. If the participant said the probability was 1/2, then one could infer that for that participant, the gap case (the orange circle) was ignored for the purposes of calculating the probability; only the two squares were under consideration. But if the participant said the probability of selecting a yellow square was 2/3, then the gap case presumably counted as a success; that is, the sentence would be considered *true* of both the yellow square and of the orange circle.

Analogously with the above example, Cremers et al. (2015) investigated how participants treated potential truth value gaps associated with *presupposition, implicature, homogeneity*, and *vagueness*. They found that implicature and presupposition patterned differently from each other and from vagueness and homogeneity, whereas vagueness and homogeneity patterned together. That is, participants treated 'gappy' instances of vagueness and homogeneity in the same way, whereas they treated gappy instances of implicature and presupposition distinctly from each other, and distinctly from the gappy instances of vagueness/homogeneity.

Such results are *prima facie* at odds with accounts that attempt to unify scalar implicature and presupposition, such as Chemla (2009) and Romoli (2014). On the other hand, the results also suggest a parallel between vagueness and homogeneity, which is unexpected on both presuppositional (Gajewski 2005) and scalar implicature (Magri 2014) accounts of homogeneity.

Psycholinguistic work with adults is one way to get at potential parallelisms and differences among the various gappy phenomena. With successful experimental designs and methods, researchers can draw out existing differences among the phenomena, as reflected in different behavioral and/or processing measures from adults. Yet another rich source of information that can allow us to get at the same questions is child language data. Child language provides a useful tool for investigating the nature of these semantic phenomena, and also has the potential to adjudicate between competing analyses. The time course of acquisition, for example, can provide hints to common underlying interpretive mechanisms across phenomena. In the remainder of this chapter, we will discuss recent studies that have turned to acquisition to shed light on the potential connections between subsets of the gappy phenomena described in Section 1.1. In particular, two recent studies have respectively compared scalar implicature with presupposition, and homogeneity with scalar implicature. We also describe how developmental data may shed light on the potential connection between vagueness and presupposition.

For each study that we describe in the subsequent sections, we will begin by introducing the two phenomena being compared, present a brief theoretical background, describe the relevant acquisition experiments, and end with the implications of the study for the relevant theories. As we move through these studies, we will touch upon questions (8a) and (8b), and in passing, (8c).

(8) a. How are similarities and differences among gappy phenomena reflected in child language?
b. Are young children sensitive to truth value gaps, or do they display strictly bivalent truth values?
c. What methods allow us to tap into children's sensitivity to truth value gaps?

2. Presupposition and implicature

Let us start with an experimental comparison of presupposition and implicature. Bill, Romoli, Schwarz, & Crain (2016) set out to compare scalar implicatures such as the one in (9) and presuppositional sentences such as the one in (10). The goal of obtaining such comparative data was to test theories such as those put forth in Chemla (2009) and Romoli (2014), both of which attempt to provide a unified explanation for the two phenomena.

(9) a. Not all of the giraffes have scarves.
b. ⇝ *At least some of the giraffes have scarves*

(10) a. The bear didn't win the race.
b. ⇝ *The bear participated in the race*

2.1 Theoretical background

The traditional perspective on scalar implicatures and presuppositions treats them as very different from each other: scalar implicatures are traditionally considered to arise from reasoning about the speaker's intentions (see Grice 1975 and much subsequent work), while presuppositions are typically analyzed as appropriateness conditions to be satisfied in the conversational context (see Stalnaker 1974; Karttunen 1974; Heim 1982, among others).

For presuppositions, the idea is that a sentence like (10a) is only felicitous in a context in which the presupposition in (10b) is already assumed to be in the common ground (Stalnaker 1974; Karttunen 1974; Heim 1982, 1983; Beaver & Geurts To appear). According to this perspective, presuppositions are always present in sentences where their triggers (e.g., "win") are used.

In some cases it appears possible to suspend the presupposition, as in (11a), where the continuation directly contradicts the presupposition that the bear participated in the race. This suspension of the presupposition gives rise to a meaning that can be paraphrased as in (11b).

(11) a. The bear didn't win the race… he didn't even participate!
b. It's not true that the bear both participated in and won the race.

To account for this possibility of suspension, the approaches mentioned above assume an extra mechanism, through which the presupposition is 'locally accommodated' in the scope of negation (Heim 1983; see also von Fintel 2008). The application of this mechanism gives rise to the meaning paraphrased in (11b), which is compatible with the continuation in (11a).

Turning to implicature, traditional approaches treat scalar implicatures as an independent phenomenon, following works like Grice (1975) and Horn (1972). On the traditional approach, the source of scalar implicatures involves general principles that are invoked when we interact with each other in conversation, in the following manner. Consider the implicature in (12b), which arises from the use of "some" in (12a). The hearer will assume that the speaker is being as informative as she can be. Given this, the fact that the speaker uttered (12a) rather than the more informative utterance in (13) leads the hearer to infer that the speaker's reason for not uttering the stronger alternative containing "all" is that the speaker believes this stronger alternative to be false. A further step of strengthening leads to the conclusion that (13) must be false, hence the inference in (12b).

(12) a. Some of the giraffes have scarves.
b. $\rightsquigarrow$ *Not all of the giraffes have scarves*

(13) All of the giraffes have scarves.

Implicatures also arise when strong scalar terms like "all" are embedded under negation, as illustrated in (14). (14b) is referred to as an *indirect scalar implicature*, and its derivation can be explained analogously with that of the direct scalar implicature in (12): replacing the negated universal "not all" with a negated existential "not some" effectively yields the "none" alternative in (15). This alternative is stronger than (14a), and when negated (*It's not the case that none of the giraffes have scarves)*, yields (14b).

(14) a. Not all of the giraffes have scarves.
b. $\rightsquigarrow$ *At least some of the giraffes have scarves*

(15) None of the giraffes have scarves.

In contrast to this traditional approach to scalar implicatures and presuppositions, recent accounts of these inferences have attempted to bring them closer together. In particular, some accounts treat certain presuppositions, such as the presupposition associated with the verb “win” (10b), as a scalar implicature of some kind (Simons 2001; Abusch 2002, 2010; Chemla 2009; Romoli 2012, 2014). The main argument for this analysis comes from differences that have been observed between the presupposition of “win” and those of other presupposition triggers, related to the ease with which the different presuppositions can be suspended, and to their behavior in quantificational sentences (see Abusch 2010 and Romoli 2014 for discussion).

The basic idea is that the inference (10b) is derived from (10a) as a scalar implicature, following the same line of reasoning as above. On this approach, a stronger alternative to (10a) is (16). Given that the speaker chose to utter the weaker (10a) rather than the more informative (16), the hearer infers that the latter must be false, deriving the inference in (10b).

(16) The bear didn’t participate in the race.

This approach unifies scalar implicatures like (12b) and (14b) with presuppositions like (10b), deriving the two kinds of inferences using the same line of reasoning. It therefore predicts that, everything else being equal, the two should display similar developmental trajectories in young children. Of course, the assumption that *everything else is equal* is crucial in formulating such *uniformity predictions*, and indeed the assumption may turn out to be untenable in certain cases. Such cases can nevertheless be informative for understanding the respective developmental trajectories of two phenomena (see Tieu et al. 2016 for an example of how divergent performance on standard implicatures vs. *free choice inferences* provides information about the role of *lexical alternatives* in children’s development of scalar inferences). We turn next to a study that tested one such uniformity prediction for scalar implicatures and presuppositions.

2.2 Experiment: Bill, Romoli, Schwarz, & Crain (2016)

Bill, Romoli, Schwarz, & Crain (2016) tested 20 monolingual English-speaking adults and 30 monolingual English-speaking children on the interpretation of sentences like those in (10), (12), and (14). The children were split into two age groups, consisting of sixteen 4- to 5-year-olds and fourteen 7-year-olds. The experiment used a *covered picture task* (Huang et al. 2013), with participants being shown a series of scenes involving cartoon animals participating in races.

Each trial consisted of three pictures: a first picture that set the scene and made the subsequent use of negation felicitous, and then two test pictures side by side. On the left was a visible picture, and on the right was a covered picture (an image that was hidden by a black box). The participant was presented with a short description of the context picture, followed by a test sentence that participants were told described only one of the two test pictures (either the visible one or the covered one). The participant's task was to decide which of these two test pictures the test sentence was describing, and then to provide a short justification for their decision. In the covered picture task, the participant can only see one of two possible situations depicted; the participant cannot see what is hidden by the black box. The rationale is that the participant will consider whether the visible picture is an adequate match for the target sentence. If they can imagine a scenario that is a 'better' match for the sentence – whether in terms of truth or in terms of felicity/appropriateness – they should choose the covered picture.

Examples of Bill et al.'s presupposition and direct scalar implicature targets are provided in (17) and (18), respectively.

(17) *Context (Visible test picture)*: The bear is at home and did not participate in the race.
Sentence: The bear didn't win the race.

(18) *Context (Visible test picture)*: All of the elephants are holding balloons.
Sentence: Some of the elephants have balloons.

Crucially, the visible pictures in the test trials, while consistent with the literal meaning of the test sentences, were incompatible with the relevant inference. For example, the visible picture paired with (18) depicted all of the elephants having balloons, and so was not consistent with the scalar implicature *Not all of the elephants have balloons*. Selection of the covered picture on such trials was thus interpreted as evidence for generation of the associated inference. The authors also included control trials to make sure that participants were capable both of selecting the covered picture and of selecting the visible picture, when these were consistent with the relevant inferences.

Bill et al. (2016) reported that adults selected the visible picture in the presupposition condition more so than in the scalar implicature condition; children, in contrast, were more likely to select the covered picture in the presupposition condition, compared to the scalar implicature condition. These results indicate that neither group treated presupposition and scalar implicature alike.

2.3 Implications

The data reported in Bill et al. (2016) show that children and adults do not treat presupposition and implicature alike. While children treat the two phenomena differently from the way that adults do, however, the two groups nevertheless differentiate between the two phenomena in their respective behavioral patterns. Children's selections of the covered pictures indicate that they generated the presupposition at much higher rates than the scalar implicature interpretation. On the other hand, adults appeared to generate scalar inferences much more often, while in the presupposition condition they responded as though the presupposition were not present. Bill et al.'s explanation of the presupposition results is that adults, but not children, were able to locally accommodate the presupposition under negation, leading to the interpretation in (19b).

(19) a. The bear didn't win the race.
b. It's not true that the bear both participated in and won the race.

As mentioned earlier, a unified approach to these phenomena would seem to predict that, all else being equal, participants might have computed the inferences at similar rates across the two conditions. As we have seen, however, this prediction was not borne out by the results. On the other hand, the present findings are more in line with the traditional perspective, which treats presupposition and implicature as distinct phenomena, derived through different mechanisms. This approach is compatible with an asymmetry in participants' behavioral responses to the two kinds of inferences. Children access the basic meanings of the relevant sentences: sentences containing scalar terms are interpreted literally, on the weak meaning of the scalar expression, and sentences containing presupposition triggers are interpreted presuppositionally. Adults, on the other hand, can access derived meanings, computing scalar implicatures from the scalar expressions, and accommodating presuppositions locally under negation. As things stand, unified approaches cannot capture this discrepancy between the two groups.

In sum, the developmental data reported in Bill et al. (2016) do not provide support for analyses that unify the derivation of scalar implicatures and presuppositions; rather, they appear to favor treating the two distinctly. Minimally, unified theories would have to be supplemented with additional assumptions.

3. Homogeneity and implicature

The next developmental comparison we turn to involves homogeneity and scalar implicature. Sentences containing plural definite descriptions give rise to so-called *homogeneity* effects (see, among others, Löbner 1987; Schwarzschild 1994; Breheny 2005; Gajewski 2005; Büring & Križ 2013; Spector 2013; Magri 2014). Imagine some scenarios involving four colored toy trucks, as described in (20).

(20) Critical contexts
a. *4 of 4 trucks are blue*
b. *0 of 4 trucks are blue*
c. *2 of 4 trucks are blue*

The positive (21a) is clearly true in a context like (20a), while the negative (21b) is clearly true in a context like (20b).

(21) a. The trucks are blue.
b. The trucks aren't blue.

But there is a gap between these two possible situations, namely the case in (20c): imagine that two of the trucks are blue and two are yellow. In such a context, the positive (21a) and negative (21b) are considered to be neither true nor false, corresponding either to a third truth value or to none at all. Križ & Chemla (2015) provide experimental evidence for such a *truth value gap*. Their experiment, conducted with adult English speakers, reveals that adults perceive sentences like (21a) and (21b) as neither completely true nor completely false descriptions of contexts that violate homogeneity, e.g., (20c).

3.1 Theoretical background

The earliest proposals regarding homogeneity treat it as a presupposition (Schwarzschild 1994; Löbner 2000; Gajewski 2005). On such accounts, sentences like (21a) and (21b) carry a presupposition that either all of the trucks are blue or none of the trucks are blue. In a gap context, this presupposition is not satisfied, and therefore the sentences are associated with a truth value gap.

An alternative approach is to say that the definite description itself is either existential or universal, but crucially its interpretation involves a kind of indeterminacy or vagueness. On such approaches, a sentence only has a definite truth value if it has that same truth value no matter how this indeterminacy is resolved (Spector 2013; Križ & Spector 2017). For example, assume "the trucks" in (21) has the two possible interpretations in (22), an existential one and a universal one.

(22) a. Some of the trucks are blue.
b. All of the trucks are blue.

The sentence in (21a) would then be true if both (22a) and (22b) are true, i.e. if all of the trucks are blue, and false if both (22a) and (22b) are false, i.e. if none of the trucks are blue. In a gap scenario, neither condition is satisfied, and so (21a) can be neither true nor false. Likewise, (21b) can be neither true nor false, since the negations of (22a) and (22b) would be neither both true nor both false.

Yet another approach treats homogeneity as a kind of scalar implicature. According to Magri (2014), plural definites have a literal existential meaning that can be strengthened to the universal meaning through an implicature. As we have seen, scalar implicatures arise through the comparison of assertions with alternatives that could have been uttered but were not. One way to formally capture this process is to invoke the application of a covert, grammaticalized exhaustification operator EXH (Fox 2007; Chierchia, Fox, & Spector 2011). Consider its application in (23), using our scalar implicature example from (12):

(23) EXH(Some of the giraffes have scarves)
= Some of the giraffes have scarves and NOT(all of the giraffes have scarves)

In (23), EXH takes the proposition containing "some" and affirms this proposition while negating the stronger alternative containing "all" (for further discussion, see Groenendijk et al. 1984; van Rooij & Schulz 2004; Spector 2007; Fox 2007; Chierchia et al. 2011). In the case of plural definite descriptions, Magri assumes that "some" is an alternative to the definite (just as "all" is an alternative to "some" in (23)). By applying the exhaustification process recursively, he derives what is effectively a universal meaning for the plural definite description:

(24) EXH(EXH(The trucks are blue))
= EXH(The trucks are blue) and NOT(EXH(some of the trucks are blue))
= Some of the trucks are blue and NOT(some but not all of the trucks are blue)
= All of the trucks are blue

Of the three existing accounts of homogeneity, the scalar implicature account makes a very specific and testable prediction with respect to the timecourse of acquisition: given that the "some-but-not-all" implicature is a *subcomputation* of the implicature required for homogeneity, this implicature should emerge in development *at least as early* as homogeneity. That is, we might expect to observe the concurrent emergence of the "some-but-not-all" implicature and homogeneity, or we might observe the "some-but-not-all" implicature emerge prior to homogeneity, but crucially we should not observe homogeneity emerging prior to the implicature.

3.2 Experiment: Tieu, Križ & Chemla (2015)

Tieu, Križ, & Chemla (2015, 2017) report two experiments conducted in French, one using a standard Truth Value Judgment Task (TVJT), and one using a ternary judgment task (Katsos & Bishop 2011) (the paper, dataset, and R analysis script for this study are available online at: http://semanticsarchive.net/Archive/DM5YjA1M/Tieu-Kriz-Chemla-AcqHomogeneity.html). We will describe the materials here in English. Tieu et al. presented 4- and 5-year-old children with pictures of simple objects of different colors. On critical homogeneity target trials, children saw pictures depicting gap contexts like the one in (20c), and were asked to judge sentences containing plural definite descriptions like those in (21).

If children do not initially treat plural definite descriptions as imposing homogeneity, one might expect them instead to interpret the definite descriptions as existential or universal, and to interpret negative sentences containing the definite description in a negation-preserving manner. For instance, in a gap context, children might interpret the pair in (21) along the lines of (25) or (26).

(25) a. Some of the trucks are blue.
b. None of the trucks are blue.

(26) a. All of the trucks are blue.
b. Not all of the trucks are blue.

This means that we could expect three possible outcomes for children's interpretation of the plural definite description in gap contexts, as indicated in Table 1 (Križ & Chemla 2015 report that adults judge homogeneity-violating sentences as *non-true*; in a binary *yes/no* judgment task, this typically manifests itself as *rejections* of the relevant sentences, hence the Reject-Reject pattern for the homogeneous interpretation).

Table 1. Expected responses to positive and negative gap sentences, according to the interpretation of the plural definite description

Interpretation	Positive gap sentence	Negative gap sentence
Homogeneous	Reject	Reject
Existential	Accept	Reject
Universal	Reject	Accept

Collecting children's *pairs of responses* to positive and negative gap sentences therefore allows us to determine what interpretation(s) children assign to the plural definite description (see Karmiloff-Smith 1979; Munn et al. 2006 and Caponigro et al. 2012 for relevant data pertaining to children's interpretation of plural definite descriptions; crucially, these previous studies did not include plural definite descriptions under negation, and therefore do not allow us to determine whether children start out with homogeneous, existential, or universal interpretations of plural definite descriptions). To assess Magri's (2014) scalar implicature theory of homogeneity, Tieu et al. (2015, 2017) included a comparison with scalar implicature targets. On scalar implicature target trials, children would see pictures of four blue trucks, for example, and be asked to judge existentially quantified sentences such as "Some trucks are blue."

Recall that the implicature theory of homogeneity derives homogeneous interpretations by strengthening a literal existential meaning of the plural definite description to a universal one through an implicature, which includes as a sub-computation the *not all* implicature of "some". Given this, we should not expect to see homogeneous readings of plural definite descriptions occurring less often than the *not all* scalar implicature (especially given the experimental materials for the two conditions were made as visually comparable as possible). Yet Tieu et al. observed that children and adults alike rejected the positive plural definite descriptions in gap contexts significantly more often than they rejected the underinformative scalar implicature targets, suggesting a greater presence of homogeneity effects than of the scalar implicature. This would appear to run counter to the predictions of the implicature theory.

The implicature theory of homogeneity also makes the further developmental prediction that homogeneity effects can only emerge as early as the *not all* scalar implicature; it should not arise prior to the implicature, since the *not all* implicature is a required sub-computation of the implicature that derives homogeneity effects. Yet Tieu et al. observed the presence of a group of children who displayed homogeneity effects while at the same time failing to compute the scalar implicature of "some". The presence of such a group of children would appear to run counter to the predictions of the implicature theory.

Some recent research has suggested that binary judgment tasks like the TVJT may not be sensitive enough to assess children's ability to compute scalar implicatures. Specifically, Katsos & Bishop (2011) argue that binary tasks cannot distinguish between a greater pragmatic tolerance for underinformative descriptions and a true inability to compute implicatures. They report that when 5-year-olds are presented with three response options (a small strawberry, a medium strawberry, and a big strawberry), they consistently choose the intermediate reward for literally true but underinformative scalar implicature targets. In order to get a more sensitive measure of children's knowledge of homogeneity and scalar implicatures, Tieu et al.'s second experiment made use of a ternary judgment task adapted from Katsos & Bishop (2011). Children were given the option to reward the puppet with one strawberry, two strawberries, or three strawberries. Adapting the expected binary responses in Table 1 to a ternary judgment task, participants were categorized as displaying a *homogeneous* response pattern if they gave minimal or intermediate rewards to positive and to negative homogeneity targets; they were characterized as giving *existential* responses if they gave maximal rewards to positive homogeneity targets and minimal rewards to negative homogeneity targets; and finally, they were characterized as giving the *universal* response pattern if they gave minimal rewards to positive homogeneity targets and maximal rewards to negative homogeneity targets. As in the first experiment, children and adults were also given a scalar implicature test, on which minimal or intermediate rewards were counted as evidence of implicatures.

The results of the ternary judgment task experiment replicated the essential findings of the binary truth value judgment task experiment: collapsing minimal and intermediate rewards in the ternary paradigm, the authors again observed a subgroup of children with homogeneous readings of plural definite descriptions and, nonetheless, no implicatures.

3.3 Implications

On Magri's implicature theory of homogeneity, the "some-but-not-all" implicature is a required subcomputation of the implicature that generates homogeneous readings of plural definite descriptions. This theory therefore makes the prediction that homogeneity should not emerge in acquisition earlier than the more basic "some-but-not-all" implicature. Tieu et al.'s experimental findings are inconsistent with this timecourse prediction, revealing that at least some children manage to display homogeneity effects even without the "some-but-not-all" scalar implicature. While such results do not tell us what *is* the adult means of generating homogeneity, they do allow us to assess the plausibility of one existing proposal, and thereby narrow down the theoretical possibilities in light of the data. In this case, the developmental data appear to speak against the implicature account as the means of deriving homogeneity.

4. Presupposition and vagueness

Let us now turn to a final comparison that has begun to receive attention in the adult psycholinguistics literature, but has been very little investigated in acquisition work. As we saw in the Introduction, borderline instances of vague predicates (27) and cases of presupposition failure (28) are two examples where adult native speakers are typically unwilling to qualify a sentence as clearly true or clearly false; instead, the four sentences provided in (27) and (28) are seen as *inappropriate* in some sense. As in the preceding cases we have discussed, these positive and negative sentences are *gappy* in the given contexts.

(27) *Context: The bear is slightly above average-sized.*
  a. The bear is big.
  b. The bear isn't big.

(28) *Context: The bear didn't participate in the race.*
  a. The bear won the race.
  b. The bear didn't win the race.

While vagueness and presupposition have traditionally been treated distinctly in the theoretical literature, some recent approaches attempt to unify the two phenomena, in light of certain observable commonalities.

4.1 Theoretical background

Both in cases of vagueness (Mehlberg 1958; Fine 1975; Tye 1994) and of presupposition (van Fraassen 1966; Fox 2012; George 2008), the lack of a clear truth value judgment has been analyzed with the help of trivalent logics. In these systems, sentences like those in (27) and (28) are modeled with propositions that receive a third, non-bivalent truth value in the described contexts. Despite appealing to the same tools to account for these two phenomena, however, vagueness and presupposition have traditionally been conceived of as clearly distinct phenomena (Fine 1975 analyzed vagueness with the supervaluationist system previously developed by van Fraassen 1966 for presuppositions, but did not suggest any connection between the two phenomena; Tye 1994 may have been the first to explicitly compare the two within such a perspective).

Recent approaches diverge from the traditional view, and more closely investigate the interactions between the two phenomena. Zehr (2014) models the two phenomena within a single truth-functional system containing five logical values: in addition to the traditional values *true* and *false*, three additional values correspond to the set of propositions that involve unsatisfied presuppositions, borderline cases of vagueness, or both. A principle then states that infelicity occurs whenever a proposition of any of these non-bivalent values is used.

Importantly, the five values proposed by Zehr (2014) are ordered and thus define one dimension of logical truth, and a set of semantic principles determines which of these values a given proposition receives, based on whether it is about borderline cases or unsatisfied presuppositions. An alternative proposal can be found in Spector (2015). Spector derives a 7-valued system meant to handle interactions between vagueness and presupposition. Spector provides a semantics for presuppositional expressions on the one hand, in terms of *true*, *false*, and a specific value representing presupposition failures; on the other hand, he models vague predicates as potentially generating ambiguity, resulting in four additional values representing possible combinations of the other three.

More generally, these two formalisms can be thought of as exemplifying two radically opposed positions. On Zehr's *monist* treatment, vagueness and presupposition are simultaneously processed by a single mechanism. On Spector's *dualist* treatment, vagueness and presupposition are processed by distinct mechanisms. A monist view along the lines of Zehr (2014) might lead us to expect that certain linguistic operations can affect that single mechanism, and thus be applicable to both vague and presuppositional expressions; for instance, some form of local accommodation could convert any non-bivalent proposition (resulting from vagueness or presupposition) to a bivalent one. Given that the same mechanism would deal with the infelicity of a borderline usage of a vague predicate and with the infelicity arising from a presupposition failure, one might expect that children's sensitivity to the 'gappiness' of the two phenomena might emerge concurrently.

In contrast, a dualist view along the lines of Spector (2015) might lead us to expect linguistic operations like local accommodation to be phenomenon-specific. Since vagueness and presupposition would be dealt with by distinct mechanisms, we might expect differences between the two phenomena to be reflected in language development, with no predicted relation between the acquisition of one and the acquisition of the other.

4.2 Experimental background

As we have just seen, unified approaches to vagueness and presupposition have only recently developed, and as a consequence, few experiments have compared the two phenomena. To investigate vagueness and presupposition, Zehr (2014) adapted Križ & Chemla's (2015) paradigm, originally conceived to identify truth value gaps associated with homogeneity. Zehr's aim was twofold: first to elicit truth value gaps arising from vagueness and presupposition, and second to test the prediction of the 5-valued system that presuppositions should yield different truth value judgments depending on the polarity of the sentence, whereas vagueness should be insensitive to negation.

Zehr (2014, 2015) presented adult participants with positive and negative vague and presuppositional sentences like (27) and (28), accompanied by pictures depicting borderline cases or cases of presupposition failure. As in Križ & Chemla (2015), participants were asked to assess the sentences as one of the following: *Completely false*, *Completely true*, or *Neither true nor false*. Zehr (2015) reports that participants made use of the *Neither true nor false* option in response to both presuppositional and vague sentences, regardless of polarity. In this experimental context then, speakers' behavior suggests that vagueness and presupposition share a certain 'gappiness' in truth value.

Although *Neither true nor false* responses were observed for both vagueness and presupposition, however, the two nevertheless differed with respect to the distribution of the selected response options. While vagueness predominantly triggered *Neither true nor false* judgments across both polarities, the judgments for positive presuppositional descriptions (cf. (28a)) were evenly distributed between

Completely false and *Neither true nor false*. This contrast experimentally supports a distinction between vague and presuppositional sentences, but doesn't by itself rule out the possibility that certain linguistic operations (like local accommodation) may target both vagueness and presupposition.

In fact, the following aspect of the results is in line with just this possibility: negation yielded an increase in *Completely true* answers *for both types of descriptions*. To understand how this suggests the existence of a general operation of local accommodation, recall that a positive vague or presuppositional proposition receives a non-binary value in the critical gappy contexts. The operation of local accommodation would turn the non-binary propositions into false propositions, and negation would in turn yield a true proposition. More descriptively, this result might reveal that negation can target 'borderline-ness' in the same way that it can target a presupposition. If negation can indeed target both types of content in the same way, this would provide further evidence that vagueness and presupposition can be given a similar representation, at least at the level at which a process like local accommodation operates.

These similarities in responses to presupposition and vagueness stand in contrast to the results of Cremers et al. (2015). As discussed in the Introduction, Cremers et al. (2015) compared vagueness and presupposition (in addition to homogeneity, implicatures, and conditionals) in a probability assignment task, and observed that participants displayed different response patterns for presupposition failures and borderline instances of vagueness, in particular treating presupposition failures as negative outcomes, while treating borderline instances of vagueness as ambiguous between positive and negative outcomes.

4.3 Potential insights from acquisition

The two alternative views of vagueness and presupposition discussed above make different predictions for child language development. If vagueness and presupposition involve a single mechanism operating on representations of the same type (as in Zehr's (2014) ordered 5-valued system), one might expect to see the effects of this mechanism on presupposition and vagueness emerging around the same time. On the other hand, if vagueness and presupposition are dealt with by distinct mechanisms (as in Spector (2015)'s unordered 7-valued system), one might instead expect no particular relation between children's treatment of vague sentences and of presuppositional sentences.

A potential future investigation involves adapting the design of Zehr (2014) (already adapted from Križ & Chemla (2015)) for use with children. Building on the data from adults, two developmental observations would be of particular interest: whether children give any non-binary responses at all, and whether negation increases the rate of *true* responses. A monist view along the lines of Zehr (2014) would predict a general correlation between vagueness and presupposition in both cases: if children make use of non-binary responses, they should do so for both vague and presuppositional descriptions, as these are assumed to be equally represented as non-bivalent; moreover, assuming the process behind adults' increase in *true* responses to negative descriptions applies equally to vagueness and presupposition, we may expect to see this response pattern emerging concurrently for vagueness and presupposition in children. In contrast, a dualist view along the lines of Spector (2015) predicts no such correlation: some children could give non-bivalent responses to one phenomenon while exhibiting an exclusively bivalent behavior for the other, and negation might yield an increase in *true* judgments for one but not the other.

Some defenders of the latter view could also anchor the distinction in a semantic vs. pragmatic opposition. One way to read Spector's (2015) system is to regard presupposition as being treated at the semantic level, and to regard vagueness as an ambiguity in truth value that must be pragmatically resolved. Future developmental work could investigate the implications of such a semantic/pragmatic divide for the relative timecourse of acquisition.

Finally, looking at what bivalent answers, if any, children give for each phenomenon will also be informative, in particular for derivational analyses. For instance, Abusch (2002) and Romoli (2014) propose mechanisms that derive presuppositions from propositions that are semantically *false* in case of presupposition failure. Sudo (2012) and Klinedinst (2010), on the other hand, suggest that some presuppositional sentences are semantically true in situations of presupposition failure. From a derivational perspective, looking at children's judgments may shed light on the semantic representations from which presuppositions are derived.

Further investigation into the development of vagueness also holds great potential. Positivist views according to which any entity is either in the positive or negative extension of a vague predicate (e.g., Williamson 1994) usually draw the borderline as touching upon both the negative and positive extensions. Understanding how children characterize borderline cases and where they themselves draw the line could help to shed light on the underlying representation of vagueness and how it is acquired by child learners.

5. General discussion

In this chapter, we have made the argument that child language data provide a very useful perspective with which to assess semantic theories about gappy phenomena. By comparing how children perform on the different gappy phenomena, we obtain not just a relative timeline of when different phenomena are acquired, but also insight into potential connections among the phenomena. Developmental studies can therefore shed light on how these phenomena should be treated within linguistic theories. We have seen that children differentiate presupposition from scalar implicature, providing support for those theories that posit different underlying mechanisms for the two phenomena. In another case study, we have observed developmental evidence against the view that homogeneity is derived via scalar implicature. Finally, we have suggested that further developmental study may shed light on the relationship between presupposition and vagueness.

The studies we have discussed also suggest that gappy phenomena are not all treated alike by the child learner. This is relevant for the question of how child learners initially navigate different gappy phenomena. The experimental studies we have discussed suggest that already for young children, presupposition and implicature are distinct, as are homogeneity and implicature. If indeed children already posit different analyses for two phenomena from as early as we can test them, an open question is what leads them to do so. Presumably children do not encounter (many) instances of *gappiness* in their input. They should only rarely encounter presupposition failures, borderline instances of vagueness, homogeneity violations, etc. How then do they become sensitive to gappiness, let alone distinguish among different gappy phenomena? What kinds of evidence would be relevant in helping the child to navigate gappy phenomena?

One intriguing direction would be to investigate the role of negation. Consider again Zehr's positive sentences involving borderline instances of vague predicates and presupposition failure. Zehr (2014) reports that these positive propositions receive a non-binary truth value in the critical gappy contexts. Yet negation does something interesting: it has a common effect on the two phenomena, increasing the proportion of *Completely true* judgments in adults. That is, negation actually converts the non-binary value to a clear *True*. If the set of true sentences that children hear in their input includes such negative presuppositional or vague sentences, such cases of *local accommodation* could form part of the dataset that help children to sort out the various gappy phenomena. Along these lines, comparing the potentially distinct patterns (under negation) for the different gappy phenomena may be quite instructive. For example, the child might observe that presuppositions and vagueness can be locally accommodated under negation, whereas implicatures commonly disappear under negation; meanwhile they should only rarely encounter homogeneity violations in the presence of negation. An alternative way to characterize the distinct patterns is to say that children might sometimes observe bivalent uses of vague predicates and presuppositional expressions in gappy contexts (specifically under negation), whereas they should not encounter bivalent instances of homogeneity under negation in gappy contexts. Such contrasts could turn out to be informative evidence for the learner, who is tasked with sorting out the various phenomena.

Returning to our semantic theories, the child data we have described provide but one piece of the puzzle. Such data can be considered hand in hand with experimental work that has been conducted with adult speakers. Both kinds of empirical work provide useful insights into how we should analyze various semantic phenomena. Our main premise is that fundamental similarities and differences will be reflected in the developmental trajectory of the respective phenomena, such that we can use child language as a means to better understand how these phenomena should be dealt with in our linguistic theories. Future work should continue to refine our understanding of how various semantic phenomena are alike and different. Additionally, as we have touched upon in the previous sections, future work should also be devoted to refining the experimental methods at our disposal, which allow us to tap into young children's intuitions about truth values and the relevant truth value gaps.

Before closing, we would note a further distinction that has been somewhat glossed over in the discussion. We have raised the question of when and how children might acquire *gappiness* (i.e. non-bivalent truth values), and gappy phenomena more generally. How do they come to realize that "win" triggers a presupposition, that "some" triggers an implicature, that plural definite descriptions impose homogeneity, and that "big" is a vague predicate? But this is a distinct question from asking what children actually do with these phenomena once they have discovered their gappiness (thanks to Alexandre Cremers for discussion of this point). For example, as Bill et al.'s (2016) study shows, 4-year-old children appear to be aware of the presuppositional status of "win", yet they do not locally accommodate the presupposition the way that adults do. As more gappy phenomena are systematically studied and compared in development, we will not only be able to address the question of when children's sensitivity to each gappy phenomenon emerges, we will also begin to uncover the development of the mechanisms that are involved in the interpretation and treatment of the different gappy phenomena.

Acknowledgements

For helpful discussion of the issues in this chapter, we would like to thank Emmanuel Chemla, Stephen Crain, Alexandre Cremers, and Manuel Križ.

References

Abusch, D. (2002). Lexical alternatives as a source of pragmatic presupposition. In B. Jackson (Ed.), *Semantics and Linguistic Theory (SALT) 12* (pp. 1–19).

Abusch, D. (2010). Presupposition triggering from alternatives. *Journal of Semantics*, 27(1), 1–44. https://doi.org/10.1093/jos/ffp009

Beaver, D., & Geurts, B. (to appear). Presuppositions. In C. Maienborn, K. von Heusinger, & P. Portner (Eds.), *Semantics: An international handbook of natural language meaning* (Vol. 3). Berlin: Mouton de Gruyter.

Bill, C., Romoli, J., Schwarz, F., & Crain, S. (2016). Scalar implicatures versus presuppositions: The view from acquisition. In *Presuppositions: Philosophy, linguistics, and psychology*. Special issue of Topoi, 35(1), 57–71.

Breheny, R. (2005). Exhaustivity, homogeneity, and definiteness. In P. Dekker & M. Franke (Eds.), *Proceedings of the Fifth Amsterdam Colloquium* (pp. 59–65).

Büring, D., & Križ, M. (2013). It's that and that's it! Exhaustivity and homogeneity presuppositions in clefts (and definites). *Semantics & Pragmatics*, 6(6), 1–29.

Caponigro, I., Pearl, L., Brooks, N., & Barner, D. (2012). Acquiring the meaning of free relative clauses and plural definite descriptions. *Journal of Semantics*, 29, 261–293. https://doi.org/10.1093/jos/ffr014

Chemla, E. (2009). Similarity: Towards a unified account of scalar implicatures, free choice permission, and presupposition projection (Unpublished manuscript). LSCP.

Chierchia, G., Fox, D., & Spector, B. (2011). The grammatical view of scalar implicatures and the relationship between semantics and pragmatics. In C. Maienborn, K. von Heusinger, & P. Portner (Eds.), *An international handbook of natural language meaning* (Vol. 3). Berlin: Mouton de Gruyter.

Crain, S., & Thornton, R. (1998). *Investigations in Universal Grammar: A guide to experiments on the acquisition of syntax and semantics*. Cambridge, MA: The MIT Press.

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

Cremers, A., Križ, M., & Chemla, E. (2015). Probability judgments of gappy sentences (Unpublished manuscript). Ecole Normale Supérieure.

Fine, K. (1975). Vagueness, truth and logic. *Synthese*, 30(3–4), 265–300. https://doi.org/10.1007/BF00485047

Fox, D. (2007). Free choice and the theory of scalar implicatures. In U. Sauerland & P. Stateva (Eds.), *Presupposition and implicature in compositional semantics* (pp. 71–120). Houndmills: Palgrave.

Fox, D. (2012). Presupposition projection from quantificational sentences: Trivalence, local accommodation, and presupposition strengthening (Unpublished manuscript). The Hebrew University of Jerusalem.

Gajewski, J. (2005). Neg-raising: Polarity and presupposition (Unpublished doctoral dissertation). MIT.

George, B. R. (2008). Presupposition repairs: A static, trivalent approach to predicting projection (Unpublished MA thesis). UCLA, Los Angeles, CA.

Grice, P. (1975). Logic and conversation. In D. Davidson & G. H. Harman (Eds.), *The logic of grammar* (pp. 64–75). Encino, CA: Dickenson.

Groenendijk, J. A. G., Janssen, T. M. V., & Stokhof, M. B. J. (1984). *Truth, interpretation and information: Selected papers from the third amsterdam colloquium*. Dordrecht: Foris. https://doi.org/10.1515/9783110867602

Heim, I. (1982). The semantics of definite and indefinite noun phrases (Unpublished doctoral dissertation). University of Massachusetts, Amherst.

Heim, I. (1983). On the projection problem for presuppositions. In D. P. Flickinger (Ed.), *Proceedings of WCCFL 2* (pp. 114–125). Stanford, CA: CSLI.

Horn, L. (1972). On the semantic properties of logical operators in English (Unpublished doctoral dissertation). UCLA, Los Angeles, CA.

Huang, Y. T., Spelke, E., & Snedeker, J. (2013). What exactly do numbers mean? *Language Learning and Development*, 9, 105–129. https://doi.org/10.1080/15475441.2012.658731

Karmiloff-Smith, A. (1979). *A functional approach to child language: A study of determiners and reference*. Cambridge: Cambridge University Press.

Karttunen, L. (1974). Presupposition and linguistic context. *Theoretical Linguistics*, 1, 181–194. https://doi.org/10.1515/thli.1974.1.1-3.181

Katsos, N., & Bishop, D. V. (2011). Pragmatic tolerance: Implications for the acquisition of informativeness and implicature. *Cognition*, 120, 67–81. https://doi.org/10.1016/j.cognition.2011.02.015

Klinedinst, N. (2010). Totally hardcore semantic presuppositions (Unpublished manuscript).

Križ, M., & Chemla, E. (2015). Two methods to find truth-value gaps and their application to the projection problem of homogeneity. *Natural Language Semantics* 23(3), 205–248. https://doi.org/10.1007/s11050-015-9114-z

Križ, M., & Spector, B. (2017). Interpreting plural predication: Homogeneity and non-maximality (Unpublished manuscript). Institut Jean-Nicod, CNRS, Paris.

Löbner, S. (1987). The conceptual nature of natural language quantification. In I. Rusza & A. Szabolcsi (Eds.), *Proceedings of the '87 Debrecen Symposium on Logic and Language*. Budapest: Akadémiai Kiadó.

Löbner, S. (2000). Polarity in natural language: Predication, quantification and negation in particular and characterizing sentences. *Linguistics and Philosophy*, 23, 213–308. https://doi.org/10.1023/A:1005571202592

Magri, G. (2009). A theory of individual-level predicates based on blind mandatory scalar implicatures (Unpublished doctoral dissertation). MIT.

Magri, G. (2014). An account for the homogeneity effects triggered by plural definites and conjunction based on double strengthening. In S. Pistoia Reda (Ed.), *Pragmatics, semantics and the case of scalar implicatures* (pp. 99–145). Houndmills: Palgrave Macmillan.

Mehlberg, H. (1958). *The reach of science*. Toronto: University of Toronto Press.

Munn, A., Miller, K., & Schmitt, C. (2006). Maximality and plurality in children's interpretation of definites. In D. Bamman, T. Magnitskaia, & C. Zaller (Eds.), *BUCLD 30: Proceedings of the 30th annual Boston University Conference on Child Language Development* (pp. 377–387). Somerville, MA: Cascadilla Press.

Romoli, J. (2012). Soft but strong: Neg-raising, soft triggers, and exhaustification (Unpublished doctoral dissertation). Harvard University.

Romoli, J. (2014). The presuppositions of soft triggers are obligatory scalar implicatures. *Journal of Semantics*, 32(2), 173–219. https://doi.org/10.1093/jos/fft017
Schwarzschild, R. (1994). Plurals, presuppositions and the sources of distributivity. *Natural Language Semantics*, 2(3), 201–248. https://doi.org/10.1007/BF01256743
Simons, M. (2001). On the conversational basis of some presuppositions. In R. Hastings, B. Jackson, & Z. Zvolenszky (Eds.), *Semantics and Linguistic Theory (SALT) 11* (pp. 431–448).
Spector, B. (2007). Aspects of the pragmatics of plural morphology: On higher-order implicatures. In U. Sauerland & P. Stateva (Eds.), *Presupposition and implicature in compositional semantics* (pp. 243–281). Houndmills: Palgrave.
Spector, B. (2013). Homogeneity and plurals: From the strongest meaning hypothesis to supervaluations. (Presented at Sinn und Bedeutung 18).
Spector, B. (2015). Multivalent semantics for vagueness and presupposition. *Topoi*, 1–11. https://doi.org/10.1007/s1124590149929291
Stalnaker, R. (1974). Pragmatic presuppositions. In M. Munitz & D. Unger (Eds.), *Semantics and philosophy* (pp. 197–213). New York, NY: New York University Press.
Sudo, Y. (2012). On the semantics of Phi features on pronouns (Unpublished doctoral dissertation). MIT.
Tieu, L., Križ, M., & Chemla, E. (2015). On the acquisition of homogeneity in plural definites. Poster presented at the 40th Boston University Conference on Language Development.
Tieu, L., Križ, M., & Chemla, E. (2017). Children's acquisition of homogeneity in plural definite descriptions. (Unpublished manuscript). Ecole Normale Supérieure, Institut Jean Nicod, Paris.
Tieu, L., Romoli, J., Zhou, P., & Crain, S. (2016). Children's knowledge of free choice inferences and scalar implicatures. *Journal of Semantics*, 33(2), 269–298.
Tye, M. (1994). Sorites paradoxes and the semantics of vagueness. *Philosophical Perspectives*, 8, 189–206. https://doi.org/10.2307/2214170
van Rooij, R., & Schulz, K. (2004). Exhaustive interpretation of complex sentences. *Journal of Logic, Language and Information*, 13, 491–519. https://doi.org/10.1007/s10849-004-2118-6
van Fraassen, B. C. (1966). Singular terms, truth-value gaps, and free logic. *Journal of Philosophy*, 63(17), 481–495. https://doi.org/10.2307/2024549
von Fintel, K. (2008). What is accommodation, again? *Philosophical Perspectives*, 22(1), 137–170. https://doi.org/10.1111/j.1520-8583.2008.00144.x
Williamson, T. (1994). *Vagueness*. London: Routledge.
Zehr, J. (2014). Vagueness, presupposition and truth value judgments (Unpublished doctoral dissertation). Ecole Normale Supérieure, Institut Jean Nicod, Paris.
Zehr, J. (2015, July). Vagueness, presupposition and truth-value gaps: An empirical investigation. Poster presented at Experimental Pragmatics 2015.

CHAPTER 14

Four-year-old children compute scalar implicatures in absence of epistemic reasoning

David Barner[i], Lara K. Hochstein[i], Miriam P. Rubenson[ii] and Alan Bale[iii]
[i]University of California, San Diego / [ii]University of Southern California / [iii]Concordia University

Children exhibit sophisticated mental state (epistemic) reasoning abilities from an early age, but it remains unclear what role these abilities play in the development of pragmatic inference in language acquisition. Here, we examined the role of epistemic reasoning in scalar implicature. Experiment 1 found that most 4-year-olds successfully computed a sub-type of scalar implicature (*ad hoc* implicatures), despite failing to compute so-called "ignorance implicatures", a type of epistemic inference that Gricean models of pragmatics deem necessary for scalar implicature. In Experiment 2, we tested 4- and 5-year-old children, and replicated the finding that children compute *ad hoc* implicatures more readily than ignorance implicatures, and that children's difficulties with ignorance implicature are unrelated to performance on Theory of Mind tasks.

Keywords: pragmatics, scalar implicature, ignorance implicature, language acquisition

1. Introduction

When children acquire language, they not only acquire the literal meanings of words but also the ability to compute the meanings of utterances in context – an ability that is generally thought to involve epistemic inference – i.e., reasoning about a speaker's mental states (e.g., Grice, 1975). Although there is evidence that children can reason about other people's minds from an early age (e.g., Baillargeon et al., 2010; Behne et al., 2005; Carpenter et al., 1998; O'Neill, 1996), it remains unclear exactly when children begin to apply this understanding to enrich the literal meanings of utterances pragmatically. In this paper, we explore children's ability to compute *ad hoc* scalar implicatures, pragmatic inferences which some claim are

https://doi.org/10.1075/tilar.24.14bar

acquired early in development, and we ask whether these implicatures plausibly involve epistemic reasoning.

Over the first two years of life, children develop a basic understanding of other people as intentional beings whose behaviors are mediated by beliefs and desires. For instance, by as early as 9 months of age, children differentiate between a person who deliberately refuses to cooperate and one who is simply unable to (Behne et al., 2005), and by 12 months of age, children point more to the location of an object when an adult searching for this object is ignorant of its location than when the adult knows exactly where the object is to be found (Liszkowski et al., 2008). By 14 to 18 months, children are much more likely to imitate intentional actions than unintentional ones (Carpenter et al., 1998). Recently, a number of researchers have even argued that children are able to impute false beliefs before the age of 2 (e.g., Onishi & Baillargeon, 2005; Surian et al., 2007; Southgate et al., 2007). For instance, Onishi and Baillargeon (2005) found that 15-month-olds are surprised when an agent searches for an object in the wrong place despite knowing its correct location, and also when an agent searches in the correct location despite falsely believing the object is elsewhere. These results suggest that even 15-month-olds expect agents to act in accordance with their beliefs – whether true or false – and that they are surprised when they do not.

While there is increasing evidence that children have relatively sophisticated epistemic reasoning abilities from an early age, it remains a matter of significant debate how such abilities are related to the ability to compute pragmatic inferences in language. By the age of 2, children can compute mutual exclusivity inferences, which by some accounts depend critically on epistemic reasoning – specifically, on the ability to reason about a speaker's intentions (Clark, 1990; Gathercole, 1989; Bloom, 2000; Diesendruck & Markson, 2001; Woodward & Markman, 1998; Wynn, 1992). According to such accounts, when shown two objects – one known and one novel – children assume that an unfamiliar word applies to the novel object based on the Gricean assumption that a speaker would have used a familiar word if he had wanted to reference a familiar item. For example, they might reason, "If the experimenter had wanted me to pick up the cup, she would have asked me to show her the cup. Since she asked me for the dax instead, she must want me to give her the other object" (e.g., Diesendruck & Markson, 2001). However, the role of epistemic reasoning in mutual exclusivity has been a matter of significant debate, with some researchers arguing that this inference is based on a simpler assumption that objects only have one label, thus precluding the need for reasoning about epistemic states (e.g., Markman & Wachtel, 1988; de Marchena et al., 2011; Preissler & Carey, 2005).

To further explore the role of epistemic reasoning in children's early inferences in language, researchers have investigated children's ability to compute scalar implicature, a form of inference that, by some accounts, is computationally similar to mutual exclusivity (e.g., Barner & Bachrach, 2010). Scalar implicatures arise when a speaker utters a statement as in (1), and listeners infer that the stronger alternative statement in (2) is false.

(1) Saya bought some of Dimitri's paintings.

(2) Saya bought all of Dimitri's paintings.

By neo-Gricean accounts of pragmatic inference (e.g., Horn, 1972; Gazdar, 1979; Geurts, 2009, 2010; Hirshberg, 1985), this inference is based, first, on the assumption that words like *some* and *all* form a substitution class (known as *Horn scales*; Horn, 1972) and thus that (2) is an alternative to (1); and second, on the assumption that the speaker is being cooperative and is providing an appropriate amount of information given what he knows to be true (Grice, 1975). Thus, if the speaker knew that Saya bought all of Dimitri's paintings, he would have simply said so. Because he chose not to utter this stronger, more informative alternative statement, listeners infer that he does not know or believe it to be true. As with mutual exclusivity, however, the role of epistemic reasoning in the computation of scalar implicature has come under debate, with some arguing that this inference, while generally conditioned on epistemic considerations (e.g., the competence assumption) is nevertheless not fundamentally an epistemic inference, and instead is computed using a grammatical operator that is independent of reasoning about beliefs and desires (e.g., Chierchia 2004, 2006; Chierchia, Fox, & Spector, 2012; Fox 2007; Fox & Hackl 2006; Landman 1998, among others).

Relevant to this debate, a number of studies have shown that, surprisingly, children up to 9 years of age fail to compute scalar implicatures (e.g., Noveck, 2001; Papafragou & Musolino, 2003; Paris, 1973; Braine & Rumain, 1983; Smith, 1980; inter alia). Recent work has explored whether these failures are related to pragmatic deficits in children (e.g., Katsos & Bishop, 2011), or are instead attributable to other factors like general processing limitations (e.g., Chierchia et al., 2001; Pouscoulous et al., 2007; Reinhart, 2004) or an inability to access the relevant scalar alternatives (e.g., Barner & Bachrach, 2010; Barner et al., 2011; Chierchia et al., 2001; Foppolo, Guasti & Chierchia 2012; Skordos & Papafragou, 2016; among others). In relation to this, Hochstein et al. (2016) argued that children's difficulties are not likely epistemic in nature. Their study, which we describe in detail below, showed that although 5-year-olds fail to compute scalar implicatures, they successfully compute so-called ignorance implicatures. This is important because ignorance implicatures involve all of the epistemic reasoning that Neo-Griceans argue is needed for scalar implicature (Sauerland, 2004).[1] Thus, evidence that children can compute ignorance implicatures would suggest that their problem is unlikely to be related to epistemic inference.

To understand this, consider the steps involved in ignorance implicature. An ignorance implicature arises when, for instance, a speaker utters a disjunctive statement like the one in (3), and the listener infers ignorance or uncertainty on the part of the speaker with respect to non-disjunctive statements as in (4) and (5).

(3) Bob went to Peru or Brazil for his vacation.

(4) Bob went to Peru for his vacation.

(5) Bob went to Brazil for his vacation.

The standard Gricean explanation of these inferences starts with the assumption that the speaker is being cooperative and giving the appropriate amount of information. If the speaker knew that Bob went to Peru, she should have simply uttered the stronger statement in (4). Further, if she knew that Bob didn't go to Peru but yet believes the original disjunctive statement in (3) to be true, then she should have uttered the stronger sentence in (5). Given that the speaker chose to utter neither (4) nor (5) but instead chose (3), it follows that she believes that (3) is true, but is simply ignorant about which of the two disjuncts is true.

By most Gricean accounts, scalar implicatures result when a hearer computes all of the steps involved in an ignorance implicature, but then strengthens this inference by taking what Sauerland (2004) calls the "epistemic step." To understand this, consider a situation in which a speaker utters a weaker statement *p* instead of a stronger alternative *q* (e.g., "Bob went to Peru or Brazil for his vacation" instead of "Bob went to Peru"). In this case, the hearer can immediately conclude that it is not the case that the speaker believes *q* to be true. This conclusion is consistent either with the speaker being ignorant about the truth or falsity of *q* or with the speaker knowing that *q* is false. If the hearer further assumes that the speaker is knowledgeable with respect to *q* – i.e., either the speaker believes *q* is true or believes *q* is false – then the hearer can conclude that the speaker believes that *q* is false. The strengthening of the expression from "it is not the case that the speaker believes *q*" to "the speaker believes *q* is false" is called the epistemic step. In the case of the alternatives in (4) and (5) above, the assumption that the speaker is knowledgeable is not possible. Were the hearer to assume that the speaker was knowledgeable, then pragmatic reasoning would lead to a contradiction – i.e., the speaker believing

1. A notable exception is Meyer (2013), who derives ignorance implicatures through grammatical operators.

that Bob didn't go to Brazil and that Bob didn't go to Peru, but yet stating that Bob went to either Peru or Brazil. Thus, the speaker does not assume that the speaker is knowledgeable with respect to *q*, and derives an ignorance implicature instead of a scalar implicature (see the discussion in Sauerland, 2004; Spector, 2003). However, in this same context, the hearer may conclude without risk of contradiction that the speaker is knowledgeable with respect to the stronger statement *p* & *q* (i.e., that Bob went to both places for his vacation), and take the epistemic step to conclude that Bob went to Peru or Brazil, but not to both.

Hochstein et al. (2016) showed that whereas 5-year-old children readily compute ignorance implicatures, they struggle to compute scalar implicatures, despite the fact that, on Gricean accounts, they differ only with respect to a single step. To show this, they presented children with two puppets. In one condition (which tested ignorance implicature), one of the puppets wore a blindfold and thus was ignorant about things going on in his surroundings, while the second puppet could see and was knowledgeable. In a second, scalar-implicature condition, which tested a different group of children, one puppet was described as smart, and the other as silly and prone to errors. In the ignorance implicature condition, children were presented, e.g., two objects, and a third puppet, a bear, approached and took one or both of the objects. Then, the child was told that one of the two puppets said, e.g., "The bear took an apple or a banana". They were then asked which of the two puppets – the seeing puppet or the blindfolded puppet – had said this sentence. Five-year-old children were near ceiling in identifying the blindfolded puppet as the speaker – since clearly only someone who was ignorant would use the disjunction when a stronger statement was possible. In the scalar implicature condition, children were presented, e.g., with two animals, each of whom had either one or two objects, and heard a sentence like "Each animal has an apple or a banana". When both animals had a flower and a sticker, children selected the smart puppet as the most likely speaker, and thus failed to compute the scalar implicature that disjunctive utterance implies that the animal does not have both items.

These findings suggest that, to the extent that these children struggled with scalar implicature, it was not likely due to difficulties with epistemic reasoning, since ignorance implicature requires all of the epistemic reasoning ability that is required by scalar implicature (and perhaps more, if scalar implicature is assumed to involve chiefly grammatical processes). Also, as Hochstein et al. argue, children's difficulty with scalar implicature cannot be attributed to a failure to take the "epistemic step." In their study, the knowledge states of speakers were provided contextually and did not need to be inferred; thus, no "step" was required by children. Furthermore, if children's difficulty with scalar implicature stemmed solely from a failure to take the epistemic step, we would still expect them to compute ignorance implicatures for underinformative statements. However, in Hochstein et al.'s study, children attributed underinformative statements involving scalar implicature to smart, knowledgeable speakers, suggesting that they did not compute either scalar or ignorance implicatures for these statements. On the basis of this, Hochstein et al. concluded that children failed to compute scalar implicatures because they failed to spontaneously access the relevant alternative statements and hence did not engage in pragmatic computations involving such statements. Whereas computing a scalar implicature for disjunctive statements required children to access stronger statements involving the conjunction – which were not contextually available – the alternatives required for computing ignorance implicatures were provided contextually, and were contained within the original disjunctive statement itself.

Critically, while this type of evidence speaks to why children might struggle with scalar implicature, it does not address the question of whether scalar implicature depends upon epistemic reasoning in development. If the ability to compute ignorance implicatures always precedes the ability to compute scalar implicatures – as predicted by Neo-Gricean models of pragmatic inference – then we might conclude that the former is required for the latter. Two facts are relevant to this. First, in their study of ignorance implicature, Hochstein et al. found that whereas 5-year-olds easily ascribed disjunctive statements to ignorant speakers, 4-year-olds were unable to do so, and performed at chance on both ignorance implicature and scalar implicature. Second, recent studies have argued that children as young as 3 or 4 can compute *ad hoc* scalar implicatures – i.e., inferences that are based on contextual scales (e.g., Papafragou & Tantalou, 2004; Stiller et al., 2011; see also Syrett & Arunachalam, 2016). By most accounts, *ad hoc* implicatures are formally identical to scalar implicatures except that the alternative utterances involved are derived from contextual scales (i.e., scales that arise from the context of a conversation) rather than from Horn scales (e.g., Hirschberg, 1985). For instance, consider the dialogue in (6):

(6) Speaker A: Are Monica, Jess, and Abby going to the party?
Speaker B: Monica is going to the party.

In (6), we infer that Jess and Abby are not going to the party. By standard Gricean accounts, this inference is based on the assumption that if Speaker A knew that Jess and Abby were going to the party, she would have simply said so; it therefore involves the same reasoning about speaker knowledge and intentionality as ignorance implicature. Given this, evidence that children can compute *ad hoc* implicatures before ignorance implicatures would be challenging to explain on Gricean accounts.

The evidence that children can compute *ad hoc* implicatures by 3 or 4 years of age remains inconclusive, however. For example, in one study sometimes taken to demonstrate an early ability to compute *ad hoc* implicatures, Barner et al. (2011) found that 4-year-old children can access alternatives based on *ad hoc* scales more readily than they can access members of Horn scales like *some* and *all*. In their study, children were shown a picture of a cat, dog, and cow sleeping, and asked either "Are only some of the animals sleeping?" or "Are only the cat and the dog sleeping?". Children said "yes" to the former question but "no" to the latter, whereas adults said "no" to both. Critically, though, no implicature computation was involved in this task, since the word "only" entails the falsity of stronger statements, as demonstrated by the fact that statements involving "only" are not defeasible (e.g., it is not possible to say, "Only some of the animals are sleeping; in fact, all of them are").[2] This study therefore provided evidence that children can access alternatives based on *ad hoc* scales before they can access Horn scale alternatives, but it did not address whether or not children could actually compute implicatures based on *ad hoc* scales.

Stiller et al. (2011) set out to test *ad hoc* implicature computation more directly using a forced choice paradigm. They presented 3- and 4-year old children with sets of three pictures that differed in terms of the number of features they exhibited. For instance, children were shown three smiley faces: one with no accessories, one with only glasses, and one with glasses and a top hat – thus forming the following contextual scale: >glasses, a hat, a hat and glasses>. The experimenter then told children, "My friend has glasses. Can you show me my friend?" The authors argued that, in this context, "My friend has glasses and a top hat" would have been a stronger and more informative description of the face with glasses and a top hat. Thus, the utterance "My friend has glasses" implies that the friend does not have a top hat. And, indeed, children as young as 3 years of age in this study reliably chose the face with only glasses in this context. However, utterances like, "My friend has glasses. Can you show me my friend?" presuppose that *only* one of the three possible referents is the correct one. Thus, children could have chosen the face with only glasses not because they ruled out the face with glasses and a top hat as a potential referent (via implicature), but because they had a slight preference for the face with glasses when forced to choose between the two, on the basis of informativeness alone (e.g., because it was a minimal match to the description they were given). Thus, although the forced choice task used by Stiller et al. demonstrates that children are sensitive in some way to utterance informativeness, it does not show whether children actually consider a statement like "My friend has glasses" to be an infelicitous description of the man with glasses and a hat (for a similar point, see Katsos & Bishop, 2011). This, coupled with the fact that forced-choice

2. Furthermore, questions are generally not believed to generate implicatures, a fact which was also borne out by Barner et al.'s (2011) study in contexts where the word "only" was omitted from questions. When asked, "Are some of the animals sleeping?" both children and adults said "yes" even if all of the animals were sleeping.

paradigms may make stronger alternative statement more available to children, make it difficult to compare these results to others in the literature, or to Hochstein et al.'s ignorance implicature data, which did not involve a forced choice between utterances or interpretations, but instead between likely speakers.

Finally, in a study by Papafragou and Tantalou (2004), 4- to 6-year-old children were presented with a scene in which a character was instructed to complete a task (e.g., to wrap a toy parrot and a doll). When the character was asked whether it had completed the task (e.g., "Did you wrap the gifts?"), it responded by mentioning only one of the required items (e.g., "I wrapped the parrot"). Children were then instructed to reward the character only if it had successfully completed the task. Papafragou and Tantalou found that their subjects refused to reward the character 90% of the time on these trials, which they took as evidence that 4- to 6-year-olds can successfully compute *ad hoc* implicatures (i.e., children interpreted "I wrapped the parrot" as implying the falsity of the stronger alternative utterance, "I wrapped the parrot and doll"). However there are several problems in interpreting this finding. First, relevant to the present study, the mean age of their subjects was 5;3, making it unclear whether 4-year-olds can generally compute these implicatures. Furthermore, in a subsequent study, Sullivan et al. (2011) provided evidence that children's behavior on Papafragou and Tantalou's (2004) task may not actually reflect scalar implicature at all. In their study, Sullivan et al. replicated Papafragou and Tantalou's methods and found that in addition to refusing rewards for characters who only mentioned a subset of the required items, children also refused rewards for characters who mentioned *more* items than were required. For instance, if a character who was instructed to wash a hat was asked, "Did you wash the hat?" answered, "I washed the hat and the shirt," children refused to reward the character, despite the fact that the character had clearly met the requirements of the task. Thus children not only rejected statements that implied the falsity of stronger alternative utterances, they also rejected statements that entailed the truth of weaker alternative utterances. These findings suggest that although children are certainly able to make pragmatic inferences, these inferences may not be full-fledged implicatures, and may instead be explained by simpler mechanisms like mutual exclusivity or the principle of contrast (for discussion, see also Bale & Barner, 2013).

It remains unclear, then, exactly when children can successfully compute *ad hoc* implicatures and how this ability relates to the ability to compute ignorance implicatures in a similar context. Although results from previous studies are consistent with the idea that *ad hoc* implicatures emerge before ignorance implicatures, no previous study has directly compared these two forms of inference using comparable methods. To address this, we conducted two experiments. In Experiment 1, we tested 4-year-old children's ability to compute both *ad hoc* scalar implicatures and ignorance implicatures in almost identical situations, and found that children's ability to compute *ad hoc* implicatures precedes their ability to make Gricean epistemic inferences. In Experiment 2, we replicated our results from Experiment 1 and also included a group of 5-year-olds in order to replicate Hochstein et al.'s finding that 5-year-olds compute ignorance implicatures. Also, we asked how tests of scalar implicature and ignorance implicature relate to general measures of Theory of Mind.

2. Experiment 1

2.1 Method

2.1.1 *Participants*

Based on Hochstein et al. (2016), we recruited 50 monolingual English-speaking 4-year-olds (25 females, *M* = 4;4, *range* = 4;0–4;11), who were recruited either by phone or through daycares in San Diego, CA and Comox, B.C. Two additional children were excluded; one due to experimenter error and the other due to failure to complete the task.

2.1.2 *Procedure & stimuli*

Subjects were seated at a small table next to the experimenter and were administered one of two tasks: an *Ad hoc Implicature* task or an *Ignorance Implicature* task, between subjects. In each task, subjects watched a scene on a computer, were presented with a sentence, and were asked to determine which of two characters most likely uttered the given statement.

2.1.2.1 *Ignorance task*

In this task, subjects were introduced to two plastic action figures, Farmer Brown and Captain Blue. The experimenter wrapped a blindfold around Captain Blue's eyes, and explained that, "Captain Blue has a blindfold on, so he can't see. He can still hear, but he can't see anything, so he might say things that are funny or not true." Each subject then received 4 warm-up trials followed by 10 total test trials.

On each of the four warm-up trials, there were three or four objects on the lower right-hand corner of the screen (e.g., 3 toy cars) and an animal (e.g., a cow) in the upper left-hand corner. The animal introduced itself, identified the items on the screen, announced its intention to take something, and then took all but one of the items (e.g., "It's me, cow! Look, 3 cars! Look what I'm taking!"). The experimenter then looked at Farmer Brown and Captain Blue and said, "Hey puppets, what did the [cow] take?" while leaning in to "listen" to the puppets whisper their responses. On two of the trials, subjects were then presented with a sentence that mentioned more items than were actually taken (e.g., "The cow took 3 cars" when in fact it took 2) and were asked, first, whether this was "the best thing to say" and, second, whether it was Captain Blue or Farmer Brown who uttered this sentence. On the other two trials, subjects were presented with a sentence that accurately described the number of items the animal had taken (e.g., "The cow took 2 cars") and were asked, first, whether this was "the best thing to say", and, second, whether it was Captain Blue or Farmer Brown who uttered it. These warm-up trials were designed to familiarize subjects with attributing sentences to one or the other character. Subjects were given feedback on their performance on these trials.

The 10 test trials were identical to the warm-up trials, except that there were two different objects on the lower right-hand corner of the screen (e.g., a banana and a toy car) and on each trial the animal either took one or both of the items. These test trials consisted of 5 different types, as shown in Table 3.1 below, and these trials were presented in one of two counter-balanced orders. There were four critical trials and two trials for each control trial-type (four total).

Table 3.1 Test trials in the *Ignorance Implicature* task

Condition	Choices	Animal takes	Someone says…	Correct response
True (Control)	Orange / Banana	Orange & Banana	The bear took an orange and a banana.	Seeing doll
False (Control)	Orange / Banana	Orange	The bear took an orange and a banana.	Blindfolded doll
Ignorance (Critical)	Orange / Banana	One or both items	The bear took an orange or a banana.	Blindfolded doll

The *True* trials were attributable to the seeing doll (Farmer Brown), as he was the only one who knew exactly what the animal took on each trial. The *False* trials were attributable to the blindfolded doll (Captain Blue), as he was the only one in a position to guess incorrectly. Although the *Ignorance* statements were literally true, we expected that subjects would attribute them to the blindfolded doll if they were able to compute ignorance implicatures.

2.1.2.2 Ad hoc implicature *task*

This task was identical to the *Ignorance Implicature* task, except that subjects attributed sentences to smart vs. silly speakers rather than to knowledgeable vs. ignorant ones, and the critical trials differed slightly (see Table 3.2 below). At the beginning of the task, subjects were introduced to two stuffed animals: "Smart Puppet", who "always says things that are just right", and "Silly Puppet", who "always says things that are a little weird or silly." Each subject then received four warm-up trials followed by eight total test trials.

Table 3.2 Trials in the *Ad hoc Implicature* task

Trial Type	Choices	Animal takes	Someone says…	Correct response
True (Control)	Orange / Banana	Orange & Banana	The bear took an orange and a banana.	Smart puppet
False (Control)	Orange / Banana	Orange	The bear took an orange and a banana.	Silly puppet
Ad Hoc (Critical)	Orange / Banana	Orange & Banana	The bear took an orange.	Silly puppet

The *True* trials were attributable to the Smart Puppet and the *False* trials to the *Silly Puppet*. Though literally true, we expected the *ad hoc* sentences to be attributed to the *Silly Puppet* if children were able to compute *ad hoc* implicatures.

2.2 Results

Mixed effect logistic regression analyses were conducted using the lme4 package (Bates et al., 2015b) in the R environment (R Core Team, 2014) and p values for regression coefficients were obtained using the RePsychLing package (Version 0.0.4). We follow procedures proposed in Bates et al. (2015a), first fitting a maximal model, and then simplifying the random effects structure to avoid over-parameterization. All models reported below contain the maximal random effect structure supported by the data.

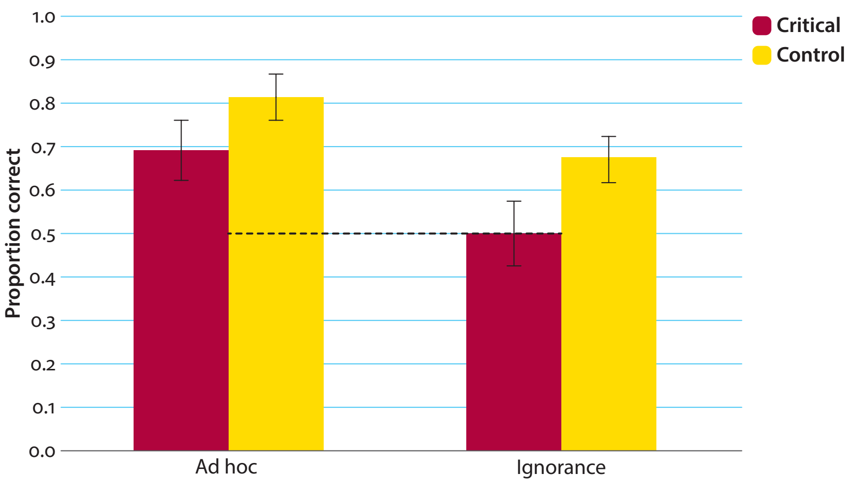


Figure 3.1 Proportion of correct responses on critical vs. control trials in the *Ad hoc* and Ignorance Implicature tasks in Experiment 1. Error bars represent standard error of the mean

Figure 3.1 displays performance on the critical and control trials in both tasks. The main question we sought to answer was whether 4-year-olds exhibit evidence of computing *ad hoc* implicatures before *ignorance* implicatures, or vice versa. To this end, we fit a model containing fixed effects of Task (*Ignorance Implicature* vs. *Ad hoc Implicature*), Trial Type (*Critical* vs. *Control*) and the interaction of Task by Trial Type. We also fit a random intercept for subject and a random Trial Type by-subject slope. This model found a significant main effect of Task ($\beta = -.95$, $SE = 0.45$, $p = .034$, $OR = .39$, 95% $CI = [.16, 93]$), but no significant main effect of Trial Type ($p = .79$) or interaction ($p = .48$). Thus, subjects performed better overall on the *Ad hoc Implicature* task than on the *Ignorance Implicature* task, but similarly on the control and critical trials within each task. Performance was significantly above chance on both trial types for the *Ad Hoc* Implicature task (*p's* < .0001) but only on the control trials for the *Ignorance* Implicature task ($p = .0018$).

Taken together, these results suggest that the 4-year-olds in our study performed better overall on critical and control trials for the *Ad hoc Implicature* task relative to the *Ignorance Implicature* task, and that children differed from chance only on the *Ad hoc* task. Very generally, children appeared to struggle with the *Ignorance* task, whether this involved critical or control trials, perhaps because in both cases mental state ascription – i.e., epistemic reasoning – was involved. Overall, these findings suggest that children exhibit an ability to compute one form of scalar implicature despite struggling at the same age to reason about the epistemic states, which, on neo-Gricean models, are necessary for scalar implicature.

3. Experiment 2

Although Experiment 1 demonstrated a difference between two types of inference, it did so using a between-subjects design, raising the possibility that the result was due to sampling error. To address this, and to replicate the finding in Experiment 1, we conducted a second experiment using the identical methods, but now in a within-subjects design.

Experiment 2 also explored how implicature computation is related to more general tests of Theory of Mind, in order to explore the question of why ignorance implicature is harder for children. To explore this, we included a group of 5-year-olds, since past studies have found that children this age successfully compute ignorance implicatures. In this way we were able to ask how children's emerging ability to compute ignorance implicatures is related to more general Theory of Mind abilities.

3.1 Method

3.1.1 *Participants*

Based on Hochstein et al. (2016) and Experiment 1 of this paper, we sought to test 50 4- and 5-year-olds in a within-subjects design. We successfully tested 25 monolingual English-speaking 4-year-olds (14 females, *M* = 4;6, *range* = 4;0–4;11) and 24 English-speaking 5-year-olds (16 females, *M* = 5;5, *range* = 5;0–5;11) recruited either by phone or through daycares in San Diego, CA. Five additional children were excluded; four due to experimenter error and one due to failure to complete the task.

3.1.2 *Procedure & stimuli*

Children were tested on both the *Ignorance* Task and *Ad hoc* Task described in Experiment 1, within subjects. Testing procedures and stimuli in Experiment 2 were identical to those in Experiment 1, with the exception that all children received both tasks in counterbalanced order. In addition, all children in Experiment 2 participated in a Theory of Mind battery assessment adapted from Wellman and Liu (2004) after completing the *Ignorance* and *Ad hoc* tasks. This allowed us to test whether children's failures with ignorance implicature were specific to the problem of computing linguistic implicatures, or instead were attributable to more global problems with epistemic reasoning. The battery included 7 scenarios from Wellman and Liu (2004), described and acted out using dolls, pictures, and toys. Children were asked 2-alternative forced-choice questions about the mental states of the characters in each scenario. The scenarios were presented in random order, with the experimenter describing each as below and then asking one or two questions about each. Children were scored correctly on scenarios only if they answered both questions correctly (for scenarios with two questions).

Table 4.1 Theory of Mind Battery Scenarios, adapted from Wellman & Liu (2004)

Mental state	Scenario	Question(s)
Real vs. apparent emotions	Steve doesn't want his aunt to know he doesn't like the present she got him.	How does Steve look on his face? Happy or sad? How does Steve feel inside? Happy or sad?
Not-own desire	Which is your favorite, a carrot or a cookie? Well, Mr. Jones likes (whichever the child did not choose), he doesn't like (participant's choice). Now it's snack time!	What will Mr. Jones pick? A carrot or a cookie?

(*continued*)

Table 4.1 (*continued*)

Mental state	Scenario	Question(s)
Not-own belief	Captain Blue's cat is missing. Where do you think it is hiding, in the bushes or in the garage? Captain Blue thinks his cat is in the (whichever the child did not choose).	Where will Captain Blue look for his cat? In the bushes or in the garage?
Belief vs. emotion	Here is a Mac – n'cheese box! (Experimenter shows child that the box contains sticks, not Mac – n'cheese). Now here comes Teddy, who has never seen inside the box. Teddy loves Mac – n'cheese.	How does Teddy feel when he first gets the box? Happy or sad? How does Teddy feel after he looks in the box? Happy or sad?
Contents false belief	Here is a Band-Aid box. What do you think is inside? Well, actually it's a toy dog! Now here comes Paul, who has never seen inside the box.	What does Paul think is in this box? Band-Aids or a dog? Did Paul see inside box? Yes or no?
Explicit false belief	Scott is looking for his mittens. Scott's mittens are really in his backpack, but he thinks they – re in the closet.	Where will Scott look for his mittens? In his backpack or in the closet? Where are his mittens really? In his backpack or in the closet?
Knowledge access	Here is a plain small box. [Experimenter shows child a small toy chair inside the box, then closes the box again]. Here comes Farmer Brown, who has never ever seen inside.	Does Farmer Brown know what's in the box? Yes or no? Did Farmer Brown see inside the box? Yes or no?

3.2 Results

We found that both 4- and 5-year-olds in Experiment 2 performed better overall on the *Ad hoc Implicature* task than on the *Ignorance Implicature* task. To assess this, we used a mixed effects logistic regression to predict correct responses using Task (*Ignorance* Implicature vs. *Ad hoc* Implicature), Trial Type (Critical vs. Control), and Age (4- vs. 5-years-old) as fixed effects and subject as a random effect.[3] This model included a random intercept for subject and a random Task by-subject slope.

The model found significant main effects of Task ($\beta = -.90$, $SE = 0.38$, $p = .017$, $OR = .41$, 95% $CI = [.19, .85]$) and Trial Type ($\beta = -1.01$, $SE = .36$, $p = .0056$,

3. The model containing Theory of Mind as a fourth fixed effect did not converge even with the minimal random effects structure, suggesting that the data do not support a 4-way interaction. Since our question was whether 4-year-olds fail to compute ignorance implicatures because of lower theory of mind abilities than 5-year-olds, we analyzed the effect of Theory of Mind in a separate analysis that included data from both age groups, but without Age as a factor.

OR = .36, 95% *CI* = [.17, .74]), indicating that subjects performed better overall on the *Ad hoc* Implicature task than on the *Ignorance* Implicature task and better overall on the control trials than on the critical trials in both tasks (Figure 3.2). Exactly as in Experiment 1, performance was significantly above chance on both control and critical trials for the Ad Hoc Implicature task (*p's* < .0001) and on control trials for the Ignorance Implicature task (*p* < .0001), but not critical trials. The model found no significant main effect of Age (*p* = .35). However, it did find marginally significant interactions involving Task and Age (β = 1.05, *SE* = .62, *p* = .087, *OR* = 2.87, 95% *CI* = [.85, 10.14]) and involving Task, Trial Type, and Age (β = −1.45, *SE* = .77, *p* = .062, *OR* = .24, 95% *CI* = [.05, 1.08]).

To explore these interactions, we next looked at the effects of Task and Trial Type on performance separately for each Age group. Figure 3.2 shows the mean performance for each age group by Task and Trial Type. For 4-year-olds, the best fitting model included Task and Trial Type as fixed effects, a random intercept for subject, and a Task by-subjects random slope. This model found significant main effects for Task (β = −.88, *SE* = .37, *p* = .017, *OR* = .42, 95% *CI* = [.20, .85]) and Trial Type (β = −.99, *SE* = .36, *p* = .006, *OR* = .37, 95% *CI* = [.18, .75]) but no significant interaction (*p* = .94). For 5-year-olds, the best fitting model included Task and Trial Type as fixed effects, a random intercept for subject, and random by-subject slopes for Task, Trial Type, and Task by Trial Type interaction. This model found a significant interaction of Task and Trial type consistent with better performance on the *Ad hoc* task than on the *Ignorance* task (β = −1.48, *SE* = .67, *p* = .027, *OR* = .23, 95% *CI* = [.06, .83]) but no main significant main effects of Task (p=.78) or Trial Type (*p* = .28).[4]

Finally, we analyzed the effect of Theory of Mind scores on children's judgements by conducting a mixed-effects logistic regression containing Task, Trial Type, and Theory of Mind scores as fixed effects, and Subject as a random effect. The final model included a random intercept for Subject and a random by-subject slope of Task. This model found a significant main effect of Theory of Mind (β = .52, *SE* = .17, *p* = .002, *OR* = 1.69, 95% *CI* = [1.21, 2.35]), suggesting that subjects with better Theory of Mind scores performed better overall on the implicature tasks. A significant interaction of Task, Trial Type, and Theory of Mind was also found, (β = −.93, *SE* = .26, *p* = .0003, *OR* = .39, 95% *CI* = [.24, .66]). As shown in Figure 3.3, these interactions were attributable to correlations between Theory of Mind and all trial types *except* critical ignorance implicature trials. This included Ignorance Control trials, suggesting that, although Theory of Mind ability was

4. To compare our *Ignorance* Implicature results to those in Hochstein et. al (2016), we conducted separate ANOVAS for each age group, using Trial Type (Control or Critical) and Experiment as factors. For both age groups we found an effect of Trial Type, but no effect of Experiment, or Trial Type by Experiment interaction. Thus, despite small differences in means across studies, our Ignorance Implicature results did not differ significantly from those in Hochstein et al.

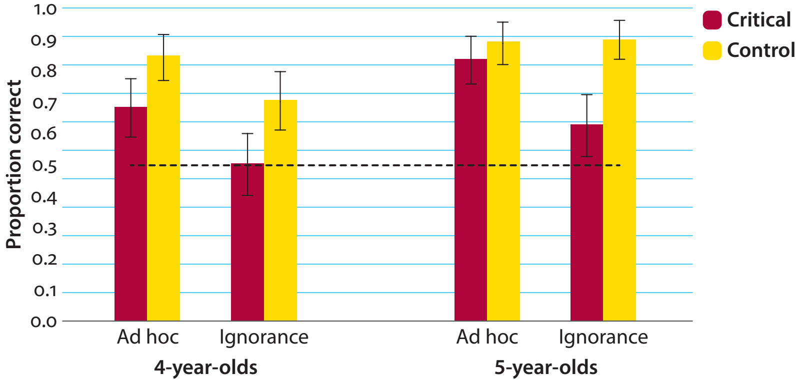


Figure 3.2 Proportion of correct responses on critical vs. control trials in the *Ad hoc* and *Ignorance* Implicature tasks in Experiment 2. Error bars represent standard error of the mean

generally related to inference making ability in this task, it could not fully explain children's difficulties with ignorance implicature computation. We interpret this result to indicate that, ignorance implicature, while surely dependent upon Theory of Mind, also depends upon additional abilities that may take longer to acquire – e.g., knowledge of which particular utterance types imply ignorance. Thus, Theory of Mind abilities may be necessary but not sufficient for ignorance implicature, thus explaining the lack of correlation we see here. Whatever additional ability ignorance implicature involves has not yet emerged at this age.

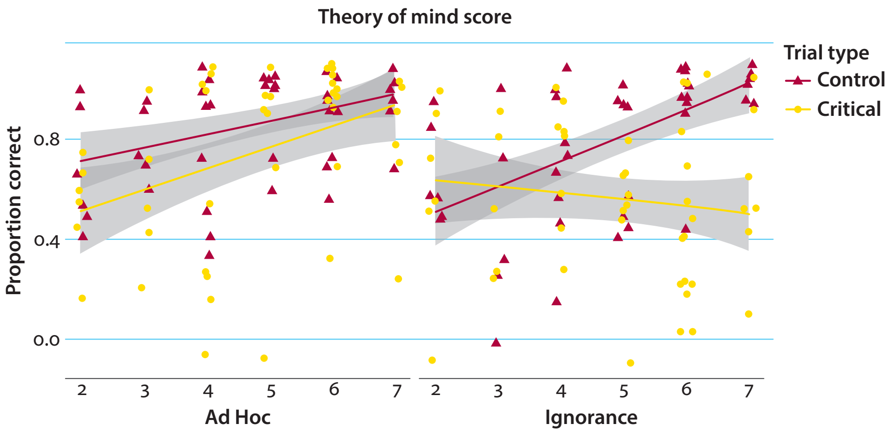


Figure 3.3 Correlation of Theory of Mind scores and proportion of correct responses on critical vs. control trials in the *Ad hoc* and *Ignorance* Implicature tasks in Experiment 2. Error regions are 95% confidence intervals

4. General discussion

In two experiments, we found that many children performed well on tests of *ad hoc* implicature despite failing to compute ignorance implicatures, suggesting that epistemic inference may not be constitutive of scalar implicature, contrary to neo-Gricean models of communicative inference. In Experiment 1, we found that one group of 4-year-old children performed well on a test of *ad hoc* implicature whereas a second group of 4-year-olds struggled to compute ignorance implicatures. In Experiment 2, we replicated this finding using a within-subjects paradigm that also included 5-year-olds. Overall, children again performed well on the *ad hoc* implicature task, but did not differ from chance on critical ignorance implicature trials. Finally, we showed that although Theory of Mind is generally predictive of children's inference making ability – including ignorance implicature control trials, which also depended on Theory of Mind – it was not related to ignorance implicature critical trials. This suggests that, while ignorance implicature clearly involves Theory of Mind – enough to restrict performance even on control trials – Theory of Mind ability is not sufficient for ignorance implicature. In addition to being able to infer a speaker's mental states, children must also recognize how particular types of utterances correspond to states of knowledge. Being able to identify when an individual is ignorant is necessary for this task, but in addition the child must recognize that a disjunctive statement could only be spoken by such a person. Taken together, these results suggest that, for children, Gricean epistemic reasoning is not strictly necessary for computing implicatures. Below, we explore this conclusion and consider how *ad hoc* implicature might be computed in absence of Gricean epistemic reasoning.

By most analyses, computing an *ad hoc* implicature involves three main steps. First, the hearer must identify a contextually specified scale or substitution class (see Hirschberg, 1985). In our experiment, the contextual scale included the two nouns that best described the two objects presented on the screen (e.g., {orange, banana}).[5,6] Second, upon hearing a statement, the hearer must generate alternatives

---

5. According to some accounts, such scales would also include the conjoined noun phrase, e.g., *banana and orange*. However, it is unlikely that this could provide a viable analysis of children's abilities since they are known to experience difficulty computing implicatures based on conjunctive alternatives (see Chierchia et al., 2001; Gualmini et al., 2001). For this reason, we focus our discussion on alternative sets that does not include conjunction.

6. To simplify our discussion we will treat the *ad hoc* scale as a substitution class containing lexical items, although this should be understood as a proxy for a scale containing nominal meanings (i.e., denotations) rather than actual lexical items, more akin to the types of alternatives discussed in Rooth (1992), and elsewhere.

using the contextually specified scale. These alternatives are constructed by substituting members of the scale for one another in the original utterance (Horn, 1979), or by deleting certain words or phrases in the original utterance (as discussed in Katzir, 2007). Note that deriving alternatives by deletion is not a necessary step for the *ad hoc* implicatures in our experiments, although it is a critical step for computing ignorance from disjunctions – a point we return to shortly. Whether the alternatives are derived by substitution or deletion, critically the alternative set only contains statements that are not entailed by the original utterance. For example, if the contextual scale were {orange, banana} and the original utterance was, "*The bear took a banana*", then the set of alternatives would contain the statement, "*The bear took an orange*" (substitution of *orange* for *banana*), which is not entailed by the original utterance. Having constructed a set of alternatives, the hearer must then compute a strengthened meaning that implies the falsity of every sentence in the set of alternatives that is not entailed by the original utterance. There are two ways this strengthened meaning could be computed:

(1) The hearer engages in Gricean reasoning and derives the falsity of the alternatives by taking the epistemic step (as hypothesized in Sauerland, 2004; Geurts, 2009, 2010; among others). In other words, the hearer assumes that the speaker knows whether each of the alternatives is true or false and that for any alternative *p*, the speaker would have uttered a sentence that entails *p* if the speaker believed that *p* was true;
(2) The hearer strengthens the original statement either by default (Levinson, 2000) or via a silent grammatical operator akin to *only* in English (as hypothesized by Chierchia et al., 2012; Chierchia, 2006; Fox, 2007; among others). On the grammatical view, for example, the sentence *The bear took a banana* would receive an interpretation similar to *The bear took only a banana*, which would entail the falsity of each non-weaker alternative in the alternative set (e.g., the falsity of *The bear took an orange*).

Critically, although both classes of theory acknowledge that epistemic reasoning plays a role in licensing implicatures, only on the first class of theories are the computations themselves couched in the language of epistemic reasoning. In contrast, both classes of account acknowledge that ignorance implicature constitutes a form of epistemic reasoning (since after all, the computations are made in order to infer ignorance or knowledge).[7] Given these two basic approaches, it is possible

7. A notable exception is Meyer (2013), who derives ignorance implicatures with a silent operator in a similar way that Chierchia et al. (2012) derives scalar implicatures. Such a theory lies beyond the present discussion.

to evaluate the developmental data regarding ignorance and *ad hoc* implicatures in relation to existing theoretical frameworks. We see three ways in which our data might be interpreted.

Non-Gricean scalar implicature: One possible account of our experimental results is that when children ascribe underinformative utterances to the silly puppet in the *ad hoc* task they do so by computing a scalar implicature. For example, upon seeing the bear take a banana and an orange and then hearing, "Someone says 'The bear took a banana.' Who said that?" children might compute a strengthened meaning for *The bear took a banana*, and thus ascribe it to the silly puppet because it implied a false proposition – i.e., that the bear didn't take an orange. However, it is important to note that this is a suitable explanation only if the *ad hoc* implicatures were computed in a non-Gricean way – whether via a grammatical operator or otherwise. We return to the difficulties these latter accounts might encounter below.

Non-Gricean informativeness inference: Another possible explanation of our data is that children succeeded at the *ad hoc* task not because they computed *ad hoc* implicatures, but instead because the utterance did not fully answer the question under discussion, and thus was not maximally informative (for discussion, see Katsos & Bishop, 2011). For example, on the trials featuring bananas and oranges, the puppets were asked, "What did the bear take?" and then were told that someone said, "The bear took a banana". Since the bear also took an orange, this response does not fully answer the question, and thus might be more compatible with a less competent speaker i.e., the silly puppet. A problem with this account, however, is that it fails to account for children's performance on the ignorance implicature task. Recall that in this task, children were told that someone said, "The bear took an orange or a banana" in a context in which the bear took both. As with the *ad hoc* implicature task, this statement was underinformative in this context (and in fact is even less informative than the statement, "*The bear took a banana*"). Critically, children did not ascribe this statement consistently to the less competent speaker, i.e., the blindfolded puppet, suggesting that an appeal to informativeness cannot alone explain our data.

Ignorance of ignorance alternatives: A third possible account of 4-year-olds' better performance with *ad hoc* implicatures is that relevant *ad hoc* alternatives may be easier to generate than alternatives required for ignorance implicature. To compute the *ad hoc* implicatures in our study, children needed only to access a single lexical substitution (e.g., *The bear took an orange*) by substituting *orange* for *banana* in the utterance *The bear took a banana*. By some accounts (e.g., Katzir, 2007), the alternatives needed to compute an ignorance implicature are generated via deletion rather than substitution; for example, to generate *The bear took a banana* from *The* *bear took an orange or a banana*, it is necessary to delete “an orange or”.[8] This, of course, is unlike *ad hoc* implicature, where substitution is required (i.e., a conjoined pair of NPs replaces a single NP). While it is possible that deletion poses a greater challenge to children than substitution, this account seems implausible without further modification, since substitution actually requires deletion (i.e., substitution of *orange* for *banana* requires first deleting *banana* and then inserting *orange*).

Given these considerations, we believe that the best account of children’s behavior is that early in development they compute implicatures in absence of rich epistemic reasoning. When scalar alternatives are supplied contextually, as they were in our *ad hoc* implicature task, 4-year-olds are able to compute scalar implicatures, consistent with various other recent reports which indicate forms of success with scalar implicature at ages as young as 3 years of age (Stiller, Goodman, & Frank, 2011; Miller et al., 2005; Skordos & Papafragou, 2016; Tantalou & Papafragou, 2003). Although we know that even infants demonstrate sensitivity to speaker knowledge and intentions in some contexts (Baldwin, 1991; Tomasello & Barton, 1994; Sabbagh & Baldwin, 2001; Liszkowski et al., 2008), our study suggests that they may not yet be able to deploy this existing Theory of Mind ability to compute linguistic implicatures in the service of inferring speaker states. Specifically, children appear able to recognize that other people’s beliefs may differ from their own, and be contingent on the evidence they have witnessed, but not yet understand how choosing underinformative utterances indicates ignorance about other, alternative statements that might have been made.

One question that our study does not fully resolve is when children first make *ad hoc* implicatures, and to what extent this ability is related to other forms of linguistic inference children make early in development. Besides the evidence presented here and elsewhere regarding *ad hoc* implicature, other studies have shown that children as young as 2 years of age can compute mutual exclusivity inferences: When a speaker says, “I want a blicket”, children infer that they do not want, e.g., an apple, since if they had wanted an apple then they would have said so, by saying, “I want an apple” (Diesendruck & Markson, 2001; Markman et al., 2003). While Clark (1990) has argued that such inferences require Gricean reasoning, the actual computations posited by others, like Markman et al. (2003) are much simpler (for discussion, see Barner & Bachrach, 2010; Bale & Barner, 2013) and can explain children’s behaviors without appeal to epistemic reasoning. Minimally, what is needed is the ability to generate a set of relevant alternatives, and the ability to negate alternatives in order to exhaustify the original utterance. While mutual exclusivity

8. Sauerland (2004) proposes that the alternatives vital to the computation of ignorance implicatures involve the same type of substitution. We do not discuss this theory here since it is unable to explain the difference between the two types of implicature in this study.

differs from implicature with respect to the types of alternatives involved, children's ability to compute these inferences clearly shows that inferences similar to implicature can be generated by children in absence of conventionalized scales, and well before Theory of Mind development is fully complete. What is unknown is whether implicatures like those described in our study actually deploy the same type of mechanism as mutual exclusivity, or whether a qualitatively different mechanism is required. Future studies should explore this question by probing the conditions under which children compute mutual exclusivity inferences and *ad hoc* implicature.

To summarize, in this study we explored children's ability to compute *ad hoc* scalar implicatures and ignorance implicatures. To do this, we conducted the first direct comparison of these inferences using similar tasks, and found that 4-year-olds successfully compute *ad hoc* implicatures despite failing to compute ignorance implicatures. We conclude that children can compute *ad hoc* implicatures in absence of Gricean epistemic reasoning about speaker knowledge.

Acknowledgement

Many thanks to Allison Myers, Junyi Chu, Katie Kimura, Laura Taouk, Kaitlyn Peck, Breanna Lengyel, Tilly Sirota, Danny Fox, members of the UCSD Developmental Brown Bag, the Language and Development Lab, and AndyLab, and to all the families that participated in this study.

References

Baillargeon, R., Scott, R. M., & He, Z. (2010). False-belief understanding in infants. *Trends in Cognitive Sciences*, 14, 110–118. https://doi.org/10.1016/j.tics.2009.12.006

Baldwin, D. A. (1991). Infants' contribution to the achievement of joint reference. *Child Development*, 62, 874–890. https://doi.org/10.1111/j.1467-8624.1991.tb01577.x

Bale, A., & Barner, D. (2013). Grammatical alternatives and pragmatic development. In A. Falaus (Ed.), *Alternatives in semantics*. Houndmills: Palgrave.

Barner, D., & Bachrach, A. (2010). Inference and exact numerical representation in early language development. *Cognitive Psychology*, 60, 40–62. https://doi.org/10.1016/j.cogpsych.2009.06.002

Barner, D., Brooks, N., & Bale, A. (2011). Accessing the unsaid: The role of scalar alternatives in children's pragmatic inference. *Cognition*, 188, 87–96.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2014). lme4: Linear mixed-effects models using Eigen and S4. *R package version*, 1(7), 1–23.

Bates, D., Kliegl, R., Vasishth, S., & Baayen, H. (2015). Parsimonious mixed models. arXiv preprint arXiv:1506.04967.

Bates, D., Kliegl, R., Vasishth, S., & Baayen, R. H. (2015a). Parsimonious mixed models. arXiv:1506.04967.

Bates, D., Maechler, M., Bolker, B., & Walker, S (2015b). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Behne, T., Carpenter, M., Call, J., & Tomasello, M. (2005). Unwilling versus unable: infants' understanding of intentional action. *Developmental Psychology*, 41, 328–337. https://doi.org/10.1037/0012-1649.41.2.328
Bloom, P. (2000). *How children learn the meanings of words*. Cambridge, MA: The MIT Press.
Braine, M. D. S., & Rumain, B. (1983). Logical reasoning. In P. H. Mussen (Ed.), *Handbook of child psychology, Vol. III: Cognitive development* (4th ed., J. H. Flavell & E. M. Markman (Vol. Eds.), pp. 263–340). New York, NY: John Wiley.
Carpenter, M., Akhtar, N., & Tomasello, M. (1998). Fourteen- to 18-month-old infants differentially imitate intentional and accidental actions. *Infant Behavior and Development*, 21, 315–330. https://doi.org/10.1016/S0163-6383(98)90009-1
Chierchia, G. (2004). Scalar implicatures, polarity phenomena, and the syntax/pragmatics interface. In A. Belletti (Ed.), *Structures and beyond* (pp. 39–103). Oxford: Oxford University Press.
Chierchia, G. (2006). Broaden your views: Implicatures of domain widening and the "locality" of language. *Linguistic Inquiry*, 37, 535–590. https://doi.org/10.1162/ling.2006.37.4.535
Chierchia, G., Crain, S., Guasti, M. T., Gualmini, A., & Meroni, L. (2001). The acquisition of disjunction: Evidence for a grammatical view of scalar implicatures. In A. H-J. Do, L. Domínguez, & A. Johansen (Eds.), *Proceedings of the 25th annual Boston University Conference on Language Development* (pp. 157–168). Somerville, MA: Cascadilla Press.
Chierchia, G., Fox, D., & Spector, B. (2012). The grammatical view of scalar implicatures and the relationship between syntax and semantics. In C. Maienborn, K. von Heusinger, & P. Portner. (Eds.), *Semantics: An international handbook of natural language meaning* (Vol. 1). Berlin: Mouton de Gruyter.
Clark, E. (1990). On the pragmatics of contrast. *Journal of Child Language*, 17, 417–431. https://doi.org/10.1017/S0305000900013842
de Marchena, A., Eigsti, I. M., Worek, A., Ono, K. E., & Snedeker, J. (2011). Mutual exclusivity in autism spectrum disorders: Testing the pragmatic hypothesis. *Cognition*, 119, 96–113. https://doi.org/10.1016/j.cognition.2010.12.011
Diesendruck, G., & Markson, L. (2001). Children's avoidance of lexical overlap: A pragmatic account. *Developmental Psychology*, 37, 630–641. https://doi.org/10.1037/0012-1649.37.5.630
Foppolo, F., Guasti, M. T., & Chierchia, G. (2012). Scalar implicatures in child language: Give children a chance. *Language Learning and Development*, 8, 365–394. https://doi.org/10.1080/15475441.2011.626386
Fox, D. (2007). Free choice and the theory of scalar implicatures. In U. Sauerland & P. Stateva (Eds.), *Presupposition and implicature in compositional semantics*. Houndmills: Palgrave Macmillan.
Fox, D. & Hackl, M. (2006). The universal density of measurement. *Linguistics and Philosophy*, 29, 537–586. https://doi.org/10.1007/s10988-006-9004-4
Gathercole, V. C. (1989). Contrast: A semantic constraint? *Journal of Child Language*, 16, 685–702. https://doi.org/10.1017/S0305000900010795
Gazdar, G. (1979). *Pragmatics: Implicature, presupposition, and logical form*. New York, NY: Academic Press.
Geurts, B. (2009). Scalar implicature and local pragmatics. *Mind and Language*, 24, 51–79. https://doi.org/10.1111/j.1468-0017.2008.01353.x
Grice, P. (1975). Logic and conversation. In P. Cole & J. Morgan (Eds.), *Syntax and Semantics* (Vol. 3). New York, NY: Academic Press.

Gualmini, A., Crain, S., Meroni, L., Chierchia, G., & Guasti, M. T. (2001). At the semantics/pragmatics interface in child language. In R. Hastings, B. Jackson, & Z. Zvolenszky (Eds.), *Proceedings of Semantics and Linguistic Theory XI*. Ithaca, NY: CLC Publications.
Hirschberg, J. (1985). A theory of scalar implicature (Unpublished doctoral dissertation). University of Pennsylvania.
Hochstein, L., Bale, A., Fox, D., & Barner, D. (2016). Ignorance and inference: Do problems with Gricean epistemic reasoning explain children's difficulty with scalar implicature? *Journal of Semantics*, 33(1), 107–135.
Horn, L. (1972). On the semantic properties of logical operators in English (Unpublished doctoral dissertation). University of California, Los Angeles, CA.
Katsos, N., & Bishop, D. (2011). Pragmatic tolerance: Implications for the acquisition of informativeness and implicature. *Cognition*, 120, 67–81. https://doi.org/10.1016/j.cognition.2011.02.015
Katzir, R. (2007). Structurally-defined alternatives. *Linguistics and Philosophy*, 30, 669–690. https://doi.org/10.1007/s10988-008-9029-y
Landman, F. (1998). Plurals and maximalization. In S. Rothstein (Ed.), *Events and grammar* (pp. 237–271). Dordrecht: Kluwer. https://doi.org/10.1007/978-94-011-3969-4_10
Levinson, S. (2000). *Presumptive meaning*. Cambridge MA: The MIT Press.
Liszkowski, U., Carpenter, M., & Tomasello, M. (2008). Twelve-month-olds communicate helpfully and appropriately for knowledgeable and ignorant partners. *Cognition*, 108, 732–739. https://doi.org/10.1016/j.cognition.2008.06.013
Markman, E. M., & Wachtel, G. F. (1988). Children's use of mutual exclusivity to constrain the meanings of words. *Cognitive Psychology*, 20, 121–157. https://doi.org/10.1016/0010-0285(88)90017-5
Markman, E. M., Wasow, J. L., & Hansen, M. B. (2003). Use of the mutual exclusivity assumption by young word learners. *Cognitive psychology*, 47(3), 241–275.
Meyer, M. (2013). Ignorance and grammar (Unpublished doctoral dissertation). MIT.
Miller, K., Schmitt, C., Chang, H., & Munn, A. (2005). Young children understand some implicatures. In A. Brugos, M. Clark-Cotton, & S. Ha (Eds.), *Proceedings of the 29th annual Boston University Conference on Language Development* (pp. 389–400). Somerville, MA: Cascadilla Press.
Noveck, I. (2001). When children are more logical than adults: Experimental investigations of scalar implicature. *Cognition*, 78, 165–188. https://doi.org/10.1016/S0010-0277(00)00114-1
O'Neill, D. K. (1996). Two-year-old children's sensitivity to a parent's knowledge state when making requests. *Child Development*, 67, 659–677. https://doi.org/10.2307/1131839
Onishi, K. H., & Baillargeon, R. (2005). Do 15-month-old infants understand false beliefs? *Science*, 308, 255–258. https://doi.org/10.1126/science.1107621
Papafragou, A., & Musolino, J. (2003). Scalar implicatures: Experiments at the semantics-pragmatics interface. *Cognition*, 86, 253–282. https://doi.org/10.1016/S0010-0277(02)00179-8
Papafragou, A., & Tantalou, N. (2004). Children's computation of implicatures. *Language Acquisition*, 12, 71–82. https://doi.org/10.1207/s15327817la1201_3
Paris, S. (1973). Comprehension of language connectives and propositional logical relationships. *Journal of Experimental Child Psychology*, 16, 278–291. https://doi.org/10.1016/0022-0965(73)90167-7
Pouscoulous, N., Noveck, I., Politzer, G., & Bastide, A. (2007). A developmental investigation of processing costs in implicature production. *Language Acquisition*, 14, 347–375. https://doi.org/10.1080/10489220701600457

Preissler, M. A., & Carey, S. (2005). The role of inferences about referential intent in word learning: Evidence from autism. *Cognition*, 97, B13–B23. https://doi.org/10.1016/j.cognition.2005.01.008

R Core Team (2014). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. <http://www.R-project.org/>

Reinhart, T. (2004). The processing cost of reference-set computation: Acquisition of stress shift and focus. *Language Acquisition*, 12, 109–155. https://doi.org/10.1207/s15327817la1202_1

Rooth, M. (1992). A theory of focus interpretation. *Natural Language Semantics*, 1(1), 117–121. https://doi.org/10.1007/BF02342617

Russell, B. (2011). Topics in the computation of scalar implicatures. PhD diss., Brown University. http://semanticsarchive. net/Archive/WY1YTRhM/russell2012probabilistic. pdf.

Russell, B. (2011). Topics in the computation of scalar implicatures (Unpublished doctoral dissertation). Brown University. Posted on Semantics Archive: <http://semanticsarchive.net/Archive/WY1YTRhM/russell2012probabilistic.pdf>

Sabbagh, M. A., & Baldwin, D. A. (2001). Learning words from knowledgeable versus ignorant speakers: Links between preschoolers' theory of mind and semantic development. *Child Development*, 72, 1054–1070. https://doi.org/10.1111/1467-8624.00334

Sauerland, U. (2004). Scalar implicatures in complex sentences. *Linguistics and Philosophy*, 27, 367–391. https://doi.org/10.1023/B:LING.0000023378.71748.db

Skordos, D., & Papafragou, A. (2016). Children's derivation of scalar implicatures: Alternatives and relevance. *Cognition*, 153, 6–18.

Smith, C. (1980). Quantifiers and question answering in young children. *Journal of Experimental Child Psychology*, 30, 191–205. https://doi.org/10.1016/0022-0965(80)90057-0

Southgate, V., Senju, A., & Csibra, G. (2007). Action anticipation through attribution of false belief by 2-year-olds. *Psychological Science*, 18, 587–592. https://doi.org/10.1111/j.1467-9280.2007.01944.x

Stiller, A., Goodman, N. D., & Frank, M. C. (2011). Ad-hoc scalar implicature in adults and children. In L. Carlson (Ed.), *Proceedings of the 33rd annual Meeting of the Cognitive Science Society*, Boston, July.

Stiller, A. J., Goodman, N. D., & Frank, M. C. (2015). Ad-hoc implicature in preschool children. *Language Learning and Development*, 11(2), 176–190.

Sullivan, J., Davidson, K., & Barner, D. (2011). Children's conversational implicatures. Paper presented at the 36th annual Boston University Conference on Language Development, Boston, MA.

Sullivan, J.; Davidson, K.; Wade, S.; and Barner, D. (under review). Differentiating scalar implicature from mutual exclusivity in language acquisition.

Surian, L., Caldi, S., & Sperber, D. (2007). Attribution of beliefs by 13-month-old infants. *Psychological Science*, 18, 580–586. https://doi.org/10.1111/j.1467-9280.2007.01943.x

Syrett, K., & Arunachalam, S. (2016). Young children's developing expectations about the language of events. In J. A. Scott & D. Waughtal (Eds.), *Proceedings of the 40th annual Boston University Conference on Language Development* (pp. 375–390). Somerville, MA: Cascadilla Press.

Tomasello, M., & Barton, M. E. (1994). Learning words in nonostensive contexts. *Developmental Psychology*, 30, 639–650. https://doi.org/10.1037/0012-1649.30.5.639

Wellman, H. M., & Liu, D. (2004). Scaling of Theory-of-Mind tasks. *Child Development*, 75, 523–541. https://doi.org/10.1111/j.1467-8624.2004.00691.x

Woodward, A. L., & Markman, E. M. (1998). Early word learning. In W. Damon, D. Kuhn, & R. Siegler (Eds.), *Handbook of child psychology, Vol. 2: Cognition, perception, and language* (pp. 371–420). New York, NY: John Wiley & Sons.

Wynn, K. (1992). Children's acquisition of the number words and the counting system. *Cognitive psychology*, 24(2), 220–251.

Correspondence information

Correspondence regarding this article should be addressed to
David Barner, Department of Psychology, UCSD
9500 Gilman Drive, La Jolla, CA 92093–0109
Email: barner@ucsd.edu

Funding information

Statement: Funding for this project was provided through a fellowship to LH from the Jacob K Javits Foundation and a grant to DB from the James S. McDonnell Foundation.

CHAPTER 15

The acquisition path of near-reflexivity

Valentina Brunetto and Tom Roeper
University of Leeds / University of Massachusetts, Amherst

Classical statue sentences ('Ringo hit himself' meaning 'Ringo hit his statue') are a long-standing puzzle for binding theories. We enrich the Partition theory (Schwarzschild, 1996) to allow semantic partitions (based on contextual contrasts) to explain acquisition experiments. The semantic partitions, in turn, correspond to a syntactic analysis of bimorphemic versus monomorphic reflexives. Only morphologically complex anaphors allow near-reflexive reference to a statue. Two experiments on the acquisition of near-reflexivity in Italian and English show that this innate interface is present very early. Results from yes/no questions-after-stories given to children 4;0–6;0 years (Italian: $N = 29$; English: $N = 36$) and adult controls (Italian: $N = 30$; English: $N = 72$) supported our prediction: English children allowed near-reflexivity with *herself*, Italian children blocked near-reflexivity with *se*.

Keywords: reflexivity, near-reflexivity, proxy, partitions, binding, clitics, anaphora, acquisition, Italian

1. Introduction

1.1 Innate interfaces and theory within experimentation

Acquisition research has often focused on rare and complex expressions because they are where innate claims about grammar seem to be virtually inevitable. We explore in this chapter the acquisition of near-reflexivity (Jackendoff, 1992; Lidz, 2001). In the classical case of *statue sentences* by Jackendoff (1992), the predicate is intrinsically ambiguous between a true reflexive reading and a statue reading (henceforth, NR):

(1) Ringo walked into the Mme Tussauds museum and saw a statue of him.
Ringo hit *himself* (*himself* = Ringo <u>or</u> Statue of Ringo)

https://doi.org/10.1075/tilar.24.15bru

Near-reflexivity is not the only case of non-coextensional or *proxy* interpretation: other contexts include metonymy (part-whole relations) and guise contexts, which we argue to be less constrained than near-reflexivity. Because of their rarity, near-reflexive contexts qualify as a testing ground for the 'poverty of the stimulus' argument: properties of grammar emerge in the child from innate UG and not from any specific input.

Where phenomena are rare, experiments can work like a prism to isolate a hidden innate interface between grammar, semantics and pragmatics. This chapter in part is also aimed at uncovering interface effects between syntax and pragmatics that we argue play a critical role in UG as well. The underlying hypothesis is that there exist *strict innate interfaces* between syntax, semantics, and pragmatics (Roeper, 2014).

Why does the assumption of interfaces increase the need for innate connections? Traditional learnability arguments were one reason for the "autonomy of syntax". If each module is autonomous, the set of possible grammars is smaller. If they are connected, then the set increases exponentially. If the highly constraining connections are innate, however, the set of possible grammars does not increase. Many innate biological interfaces exist, such as eye-hand coordination: the ability to organize hand muscles online in terms of visual information, as in reading music while playing the piano.

Experimentation into interface phenomena usually requires close attention to the contextual pragmatics of stories to isolate the grammar. Examples include: long-distance movement, paired *wh*-, inverse scope, binding, exhaustivity. In our stories, we manipulated context to invite reference to a statue in reflexive constructions with English SELF anaphors and Italian SE clitics. In doing so, our goal was to explore how morphology and pragmatics conspire to *block* near-reflexivity. In brief, our experiments show that Italian-speaking children with mono-morphemic reflexive *se* know from the outset that near-reflexivity is impossible, while English-speaking children with bi-morphemic *himself* allow it immediately. We argue that an innate interface between pragmatics, semantics, and syntax is needed to explain this result.

2. Metonymy, proxy readings and near-reflexives

2.1 Morphological constraints on near-reflexivity

In wax museum scenarios, the following near-reflexive interpretation is acceptable, though marginal (Jackendoff, 1992):

(2) Ringo Starr put a hat on *himself*. (Interpretation: Ringo put a hat on his statue)

The semantic representation of this sentence requires that two *distinct* arguments (Ringo in the flesh and Ringo's statue) are projected in the syntax. In fact, dropping the reflexive results in an intransitive predicate, blocking the near-reflexive interpretation, and one must choose the real Ringo even if the contextual focus is on the statue:

(2′) Ringo put a hat on (Interpretation: *Ringo put a hat on his statue)

A critically important observation is that that the near-reflexive interpretation also does not hold with *mono-morphemic* reflexives in Italian, German, Dutch and other languages (Reuland, 1997; Rooryck & Vanden Wyngaerd, 1998; Lidz, 2001; Safir, 2004):[1]

(3) Ringo *si* mette il cappello (Italian)
Ringo SE_{cl} puts the hat
'Ringo puts a hat on himself' (Ringo/*Statue of Ringo)

In each of these languages, Ringo puts a hat on the real Ringo, just like with the intransitive English. There seems to be a complex syntactic, morphological, pragmatic interface here: only bimorphemic reflexives allow near-reflexivity.

Is there a direct connection between pragmatics and morphology? It seems far-fetched, but not inconceivable. We will argue that the connection becomes more natural when argument structure serves as a semantic bridge to near-reflexivity.

It is commonly assumed that reflexive verbs that have obligatory morphology (SE) are single argument verbs (Grimshaw, 1990; Marantz, 1984): (4a) is not parallel to (4b), but to intransitive forms like (4c):

(4) a. He SE washed
b. He washed SELF
c. He washed

1. The clitic cases are disputed by some (see Labelle's (2008) judgements for French and Marelj & Reuland (2013) for Serbo Croatian). The experimental results we report, including adults, shows unmistakably that bi-morphemic reflexives allow near-reflexivity where it is blocked for mono-morphemic cases. We think therefore that the linguists' acceptability judgements are strained and do not reflect the essential contrast, which the literature cited maintains.

An anonymous reviewer also notes that a near-reflexive reading for a sentence in (3), with a dative clitic, may not be completely impossible, whereas the ungrammaticality is clearer with verbs like *wash* or *shave*. We agree with these intuitions. The fact that the constraint is ameliorated with benefactive objects points to a difference in the argument structure (if the benefactive clitic is introduced by an applicative head as in Pylkkänen's (2008) analysis).

We claim that bimorphemic reflexives must map onto independent arguments of the verb and therefore constructions must carry both Agent and Theme.[2] If they have two arguments, then a dependency relation between individual and extension can be construed. If there is only one argument, no option for independent reference is possible, therefore reflexive clitic constructions are semantically intransitive.

A complication arising from evaluating this proposal is that referential substitution is not unique to near-reflexives. In fact, it occurs effortlessly in language:

(5) He painted the girl with blue eyes with brown eyes (Jackendoff, 1975)

Here some girl-entity remains the same despite a physical shift. However, we can even ignore bigger aspects of identity:

(6) He portrayed the German educator as an American educator.

There has been a renewed interest in the ability of a noun phrase (R-expression, pronoun, reflexive) to function as a *proxy*, namely, as an extension of an individual x that is interpreted as f(x) (see Reuland, 2011; Safir, 2004):

(7) a. *Ringo* looks real (statue/picture)
b. *Plato* is next to *Sappho* (book/work of art)
c. Mary rolled doubles and got out of jail free (game/vehicle)

The examples in (7) are classical cases of *metonymy*, which is pervasive in language. A principled account of near-reflexivity must therefore separate constructions like (3) from broader cases of metonymy. If we define near-reflexivity as a dependency relation between an Agent and a Theme that does not result in identity, then we must exclude: (i) metonymy in the broad sense (reference shifters); (ii) mental images in belief contexts.

It can be shown that near-reflexives differ from metonymical NPs, a classic case of 'reference shifters', in crucial respects, as Jackendoff (1992) points out:

(8) a. [Pointing to Michelangelo's David] That's *Michelangelo*
b. [Michelangelo falls on David] *Michelangelo fell on *himself*

---

2. We are treating the Italian case (*SE* + STESSO) and the English case (pronoun + SELF) on a par. To the extent that both Italian and English adults accept statue reference with these anaphors, their assumption of transitivity is satisfied. We do not assume that the two anaphor types are lexically and referentially equivalent, and perhaps languages with more complex paradigms such as Dutch, which has both *SE* + SELF and pronoun + SELF, may reveal subtler interpretive differences. Safir's (2004) competition scale, to which we return in Section 2.2, explicitly makes this prediction.

Unlike the proxy in (8a), the near-reflexive in (8b) bears a special relation with the subject that must be encoded in the syntactic representation. Referential expressions and pronouns can refer independently to a proxy, but anaphors cannot. Near-reflexives are not logophors or exempt anaphors, and do not require a special perspective:

(9) Walking through the wax museum, Ringo glanced at *himself*, without recognizing it was *him* the statue portrayed.

Pronouns allow metonymy, and do not bear a special syntactic relation with their antecedent (in fact they do not need an antecedent at all). Proxy pronouns are typically illustrated in the literature by examples like (10a). The problem, however, is that sentences involving propositional attitudes obscure the difference between *he* as the statue (related to but *extensionally distinct* from Ringo), and the image of Ringo portrayed in the statue (which is still part of Ringo). In (10b), where the pronoun is necessarily unbound, the proxy interpretation (*him* = statue of Ringo) sounds less natural, because the predicate would require the statue (Theme) to be introduced by an inanimate pronoun (*it*):

(10) a. Ringo thought *he* looked goofy
b. Ringo dared George to punch *it* / *?him*

Reflexive clitics resist near-reflexivity, but can participate in structures where the subject NP has been transposed by metonymy:

(11)

Mary	ha	tirato	i	dadi	e	*si*	è	mossa	di	tre	posti
Mary	aux	thrown	the	dice	and	SE_{cl}	aux	moved	of	three	spaces

'Mary threw the dice and moved herself three spaces'

Although the subject in (11) can be interpreted as Mary's extension in the board game, the clitic is interpreted as indistinct from its antecedent, and therefore the dependency relation does not involve near-reflexivity (as (7c) confirms under conjunction: *Mary rolled the dice and __ got out of jail free*. The real person rolls dice and the figurine is removed from jail.)

The boundaries between near-reflexivity and metonymy become muddled when nonagentive predicates (*see, think, believe*) are considered. Romance clitic *se* is not incompatible with mirror images and mental images:

(12)

Gianni	*si*	guardò	allo	specchio / nella	foto
Gianni	SE_{cl}	saw	at.the	mirror/in.the	picture

'Gianni saw himself in the mirror / in the picture'

(13) Davanti alla statua che lo ritraeva, Nerone *si* ammirava
In front of the statue that him$_{cl}$ portrayed, Nero SE$_{cl}$ admired
compiaciuto
satisfied'
'In front of the statue portraying him, Nero admired himself with satisfaction'

Examples like these have sometimes been claimed as evidence that reflexive clitics can introduce proxies (see Marelj & Reuland, 2013; Labelle, 2008). Labelle (2008: 856, Example 63) adduces an equivalent example to (13) as evidence that French reflexive clitics are acceptable in statue contexts. Both Labelle (2008) and Marelj and Reuland (2013) interpret this acceptability as evidence that reflexive clitic constructions are semantically transitive.

There are, however, semantic constraints in the availability of mental image readings. Pica and Snyder (1997) have pointed out that reflexive clitics in psychological predicates resist Experiencer readings.[3] They attribute this semantic constraint to the clitic's inability to introduce a *psychological partition* between the subject as 'knower' and object of knowledge. Without a modifier (*compiaciuto* in 13 or *beaucoup* in 14a) the predicate is degraded because the clitic alone is unable to imply conscious knowledge or awareness on part of the subject:

(14) a. Jean s'apprécie beaucoup
Jean SE$_{cl}$ likes very.much
b. ??Jean s'apprécie
Jean SE$_{cl}$ likes

In Marelj and Reuland's (2013) proposal, reflexive clitic constructions retain two thematic roles; however, if the clitic has X° status, it is argued to form a uniform chain with the subject which is interpreted as one argument. This is the case of Russian -*sja*, or Icelandic -*st*, which indeed block near-reflexivity.

(15) Marko zakryl**sja** na fotografii (Russian – Marelj & Reuland 2013: 77)
Marko covered.SE$_{cl}$ on photograph
(Impossible meaning: Marko covered his image in the photo; only possible meaning: a picture of Marko covering himself).

3. The following example illustrates the contrast between physical and psychological predicates in Italian: a verb like *colpire* (strike) ambiguous between a physical interpretation (to hit) and a psychological one (to impress) blocks the reflexive clitic in the psychological case:

(i) a. Gianni si è colpito con un bastone Belletti & Rizzi (1988: 299)
'Gianni struck himself with a stick'
b. *Gianni si è colpito per la sua prontezza
'Gianni struck himself by virtue of his quickness'

On the other hand, clitics can also appear in Exceptional Case Marked complements of perception verbs and belief reports. The critical factor here is that ECM constructions involve two clauses and access to two thematic role projections. If we consider a sentence like (16), quite paradoxically, the reflexive appears to pick an individual who is 'less identical' to its antecedent than the pronoun in the embedded clause.

(16) Vincent *si* ritraeva più alto di quanto non fosse
Vincent SE_{cl} portrayed more tall than not was:subjunctive
'$Vincent_i$ portrayed $himself_i$ taller than he_i (really) was'

Notice that the form of the anaphor constrains this possibility: for instance, in Dutch, we find, as expected, that only the bimorphemic anaphor results in a possible meaning:

(17) a. Vincent schilderde *zichzelf* langer af dan *hij* eigenlijk is
Vincent portrays SELF taller as than he really is
'Vincent portrays himself as taller than he really is.'
b. *Vincent schilderde *zich* langer af dan *hij* eigenlijk is
Vincent portrays SE taller as than he really is

The source of the grammaticality in the clitic sentences cannot be attributed to the clitic's capacity to introduce a distinct thematic role. For example, Icelandic clitic -*st* can participate in ECM structures (18).[4] These facts warrant an analysis where the second thematic role is shared between the surface subject and its lower copy.

(18) $Hann_i$ þiki**st** [t_i vera snillingur] (Andrews, 1990, p. 201)
$Hann_i$ se.thinks [t_i to.be genius]
'Hann thinks he is a genius.'

Although we are not committing to any specific syntactic analysis of reflexive clitics here, we suggest that the contrast between (17b) and (18) follows naturally if the clitic, unlike *zich*, is not an argument of the small clause, compatible with movement analyses of reflexives (e.g. Kayne, 1989; McGinnis, 1999) where the surface subject is a derived subject. Since *zich* cannot be construed as extensionally distinct from its antecedent, (17b) is impossible, whereas reflexive clitic constructions undergo a different derivation.

In sum, we argue that a morphological interface blocks near-reflexivity with both *zich* anaphors (long distance SE) and clitic anaphors. Once we exclude

4. Icelandic –*st* differs from Romance *se* because its use is not productive and is in fact highly idiosyncratic. Nevertheless, the availability of these structures shows that near-reflexives and mental images should not be assimilated.

perception and psychological predicates, it can be shown that mental images and non-coextensional proxies (such as statues) fall into different classes. Clitics and long distance (*zich*) anaphors pattern together in forcing indistinctness with their antecedent.

2.2 The syntax and semantics of near-reflexivity

We define near-reflexive predicates as predicates with two semantic roles (Agent, Theme). This difference at the level of argument structure predicts that anaphoric relations with bimorphemic and mono-morphemic reflexives must have different semantic representations.

Reinhart and Reuland's (1993) binding theory, which distinguishes between simple and complex anaphors, fails to adequately capture the semantics of this distinction because its only semantic relation is coargumenthood, leaving everything else outside the core of the binding principles. Since reflexivity in syntax requires co-indexation of two arguments, they argue that bimorphemic reflexives are the only anaphor type to form a reflexive predicate, whereas mono-morphemic long-distance anaphors are treated like pronominals. The near-reflexive facts suggest that, in fact, the account should be reversed: bimorphemic reflexives can behave like pronouns, but mono-morphemic reflexives do not. The core observation in (19) is that *Ringo shaved* (19b) cannot mean that he shaved the statue because it is intransitive and the same holds for Italian *se* (19c). So these contrasts arise:

(19) a. Ringo shaved *himself* (Ringo/Statue of Ringo) (two θ roles)
b. Ringo shaved Ø (Ringo/*Statue of Ringo) (intransitive: one θ role)
c. Ringo si è rasato (Ringo/*Statue of Ringo) (clitic: one θ role)
Ringo SE_{c}l aux shaved

Our hypothesis is presented in (20):

(20) All morphologically deficient anaphors block near-reflexivity.

Moreover, the difference between strict identity and near-reflexivity cannot be reduced to a difference between bound variable and (accidental) coreference: for instance, the availability of NR in the following sentence shows that near-reflexives are not free variables, in fact they can be bound to a quantified subject and still yield nonidentity:

(21) Every Beatle started undressing *himself*.

These facts suggest that the source of near-reflexivity should *not* be located in a richer referential capacity of the *self* morpheme.

Inherent reflexive verbs like *shave* and *undress* illustrate the fact that a complex reflexive (*him+self*) adds a theme in the theta grid, creating the potential for a near-reflexive relation. This potential interpretation, we argue, must be encoded in the argument structure, not in the pragmatics or in the lexicon.

According to Reinhart and Reuland's (1993) proposal, all reflexive predicates should give rise to the same semantic representation, namely, λx P(x), an intransitive with a single thematic role. However, the data in (19) show that this claim is not borne out (as noted by Lidz 2001). Bimorphemic reflexives do not give rise to such a representation, hence their ability to engage in near-reflexive ambiguities. An alternative view, proposed by Safir (2004) capitalizes on the fact that SELF, unlike the mono-morphemic form, cannot represent readings of indistinctness and should therefore be considered the marked form. In his competition model, the availability of a more dependent anaphor causes the SELF anaphor to be outcompeted due to its inability to represent obligatory indistinctness. In inherent reflexive predicates, the availability of a null form (always a mono-morphemic anaphor in the other Germanic languages), allows the contrast to reemerge. This is compatible with our proposal that SELF *never* forms a one-place predicate, which accounts for the subtle semantic differences between (19a) and (19b).[5]

Safir (2004) also notes that R&R's (1993) notion of coargumenthood, being symmetrical, does not account for the fact (first noted by Jackendoff, 1992 – see also 8b) that the two referential options never attain complete equivalence:

(22) Context: the statue portraying Ringo falls on Ringo
$\nrightarrow$ *Ringo fell on himself.

It thus appears that near-reflexivity cannot work on a chain [NP … SELF] where the NP has been already transposed by metonymy. The near-reflexivity is thus an *asymmetrical* relation from an individual to his representation. Although it is not clear how one would capture this formally,[6] we think that, while there is a role for metonymy in allowing the statue to be called Ringo, the ambiguity of the reflexive has a distinct syntactic source.

5. Near-reflexivity is one semantic reflex of this difference in argument structure between *John shaves* and *John shaves himself*, but not the only one. For example, as Safir (2004: 123) notes, adjoining a manner adverb to the intransitive can only describe the way in which one shaves but not the style of one's beard. This prediction seems to be borne out:

(i) John shaved himself crookedly/??John shaved crookedly.

6. To test for this claim, we could imagine the following (thought) experiment:

Picture 1: Ringo is leaning on his statue; Picture 2: The statue of Ringo is leaning on Ringo. Question: *where is Ringo leaning on himself*? Our intuitions quite strongly go for Picture 1. We return to this issue in our partition discussion.

3. Mapping reflexives to partitions

Other existing accounts of near-reflexivity have focused on the referential capacity of reflexives (Lidz, 2001; Zuckerman, Avrutin, & Vasič, 2002; Reuland & Winter, 2009). It is the mono-morphemic form that is too 'deficient' to introduce new referents or *guises*, whereas the bimorphemic anaphor is lexically specified for richer referential capacity.

Our account is different because we treat this relation as quantificational in the way that reciprocals like *each other* require a quantificational analysis. Our starting observation is that interpreting a near-reflexive involves a mapping between two sets of individuals – real people, statues – and an asymmetric relation between them. While these anaphors defy a traditional definition of bound variable under which we would expect two theta roles to be interpreted as one, the relation of near-identity between the two NPs is not 'accidental'. The working hypothesis is that near-reflexivity should coincide with other semantic properties. Therefore, it should not need an *ad hoc* semantic representation. Instead, we make the following claim:

(23) The relation between an NP and a bimorphemic anaphor calls for a *partition* of that NP into subsets, and a relation between those sets.

This proposal builds on Schwarzschild's (1996) analysis of partition effects in plural predicates. Against a tradition where rich semantic representations (for instance, Link's (1983) D-operator) are posited to account for distributive/collective ambiguities, Schwarzschild argues that there is a complex cognitive and pragmatic dimension that needs to be granted a more prominent place. The analysis expands Higginbotham's (1981) idea that the interpretation of a plural predicate involves a partition of the plurality into subsets:

(24) [$_S$ NP$_{plural}$ VP] is true iff there is a partition C of the plurality denoted by NP such that VP is true for every element in C (Higginbotham 1981, p. 100)

(25) [Scenario: three boys playing baseball, two boys watching, three girls playing basketball]
  a. Are the boys playing?
  b. Are the boys playing baseball?

The principle in (24) simply captures the intuition that different partitions of a plurality can shift the truth conditions of a plural predicate. For instance, most speakers would judge (25a) false because the question requires us to partition the boys into players and non-players. However, (25b) introduces a more refined partition (including players only) which makes the sentence true if *baseball* is contrasted with *basketball*. Schwarzschild (1996) further shows that a partition should be intended as a pragmatic 'free variable', susceptible to both linguistic and nonlinguistic sets.

Our proposal is that there are also partition effects triggered by complex anaphoric expressions, both reflexives and reciprocals. These involve a pragmatic partition of the set denoted by the NP into relevant subsets and a quantificational relation between these subsets. To illustrate, we start by noticing how ambiguities are triggered by the possibility of partitioning the entities in the anaphoric relation into different subsets.

(26) a. The princess and the frog kissed *themselves*
b. The princess and the frog kissed *each other*
c. The princesses and the frogs kissed *each other*

When a plurality is made of two singleton sets, as in (26a) and (26b), the simplest partition allows for an unambiguous relation between the two sets, a self-kissing event in (26a) and a reciprocal one in (26b). In contrast, (26c) is more complex. It could be that there is a reciprocal pairing between a set of princesses and a set of frogs, or that there are two sets of events, one involving only princesses and one involving only frogs.

In the mono-morphemic case, like German *sich* and Italian *si*, although the reciprocal reading is favored, the reflexive reading is allowed (27). In fact, plural predicates construed with monomorphemic reflexives are consistently ambiguous between a reciprocal and a reflexive reading:

(27) Sie küssen *sich* (reflexive/ reciprocal)
they kissed SE
'They kissed themselves' or 'they kissed each other.'

This is not the end of the story. If one creates a contrastive context in English, then the reciprocal interpretation can emerge even with the reflexive:

(28) Scenario: a group of princesses asks a group of frogs to kiss them
The frogs were shy. The frogs kissed themselves.

This could, under a partition between groups, allow the reciprocal reading where one frog kisses another frog (but crucially, not the princesses). We show that it is this kind of partition, which is contextually required, that plays an important role in the child's acquisition path.[7]

---

7. Acquisition evidence points to the existence of partition effects in quantificational contexts. For example, Avrutin and Thornton (1994) report that children overwhelmingly allowed Principle B violations (intrasentential coreference between an NP and a pronoun) in plural sentences when the NP could be interpreted collectively as in (i):

(i) *[Big Bird and the clown]$_i$ dried them$_i$ [Collective: together, with a big towel]
(ii) *[[Big Bird and the clown]$_i$ D$_j$]dried them$_j$ [Distributive: in two different events]

This shows that children distinguish quantificational relations from coreference relations, but also suggests that a partition involving first order sets ([A+B] rather than [A] + [B]) is preferred whenever the context supports it (see also Section 6).

4. The acquisition challenge

What is the acquisition path for near-reflexivity? If we exclude metonymy, then a first hypothesis would be that it is very late. Nevertheless, we might make a different guess if we make an easy pragmatic assumption: there is a *pragmatic* extension from *de se* reflexivity (actual self) to a *de re* extension (another version of self).

Evidence that children master mirror images early has been robustly found by studies on the acquisition of binding in Romance small clauses with perception verbs (Baauw et al. 1997; Brunetto 2012) in mirror scenarios like (29):

(29) *La bambina si vede ballare (allo specchio).* (> 90% accurate)
the girl SE_{cl} sees dance (at.the mirror)
'The girl sees herself dance (in the mirror).'

Are children aware of the distinction between metonymy and near-reflexivity? A study on Dutch children's sensitivity to morphological form in near-reflexive scenarios with videos and statues (Zuckerman & Vlasveld, 2004) suggests that children are aware of the different referential capacity of *zich* and *zichzelf*. Moreover, children pay attention to the nature of the near-reflexive relation, rejecting what the authors call 'bad guises', for example a character doing something to another character who is dressed in a costume portraying him.

In our hypothesis, we argue for a richer pragmatic interface. If near-reflexivity involves a partition, that is, a pairing between sets, its challenge goes beyond the ability to recognize what referent can stand as a proxy. It calls into question an ability to exclude the reflexive event from the set of alternatives the sentence calls for.

The relation between English intransitives, Romance clitics and near-reflexivity has never been investigated. Our exploration will therefore address two kinds of questions in the grammar of English and Italian children:

(30) a. Is NR constrained by the form of the anaphor?
b. Is NR subject to partition effects? Are children and adults alike?

We believe that this is a case where, as a general assumption, we should assume zero input because we also assume that near-reflexive situations, unlike metonymy with mirrors, almost never occur for children. Notice, moreover, that the traditional idea that complex anaphors are lexically specified to allow reference to a statue comes with the corollary that lexical information is acquired through exposure, hence it should be delayed. If the child only adds near-reflexivity to *himself* when he hears it, then it will be very delayed. If it follows from the semantic analysis of a bi-morphemic anaphor, then it is invoked as soon as the child sees that *himself* involves two morphemes. That is, if it comes directly from UG, then no delay based on exposure is predicted.

Where could early acquisition come from? Our strategy is traditional, so let us maximize what UG could supply a child. A morphological assumption could be that the concept of strict identity is achieved by monomorphemic elements (all clitics), not in argument structure but in a non-argument position. In English, this could be the abstract clitic position postulated by Keyser and Roeper (1992).[8] That is:

(31) John stood up/played dumb/ repaid the loan

Allows a complex verbal head:

(32) V*
V particle
re-
idiomatic adj

Reflexive compounds in English are one construction where mono-morphemic reflexives arise, suggesting that is indeed in this non-argumental position that monomorphemic clitics are generated: *self-serving ideas/*himself-serving ideas*.

Monomorphemic reflexives in compounds exactly exclude NR readings. One could imagine the following scenario, in which Ringo's statue falls on the real Ringo and injures him: it would not be called a "self-inflicted would". The restriction of monomorphemic reflexives to a non-argument position means that near-reflexivity can only be construed from an argument position. Clitic reflexives are restricted to strict identity.

English *himself* which survives in inherent reflexive constructions is also a non-argument:

(33) a. They behaved *themselves*.
b. *It was *themselves*, that they behaved.
c. *It was up that John stood.
[compare: "it was *himself* that John blamed"]

Similarly, himself in adjunct position blocks near-reflexivity.

(34) Ringo *himself* shaved (Ringo/*The statue of Ringo).

Naturalistic evidence shows that English-speaking children master adjunct reflexives very early, and possibly use them earlier than argument reflexives.

8. Keyser and Roeper (1992) show that, because abstract dative markers, bare nouns, idiomatic nouns and adjectives, prefix –*re* and particles all occupy the same position, they are mutually exclusive (*He re-left a note; *He re-lost face; *He re-picked up the ball; *He re-played me dumb).

(35) a. I do this myself. (2;1.21)[9]
b. Let me play myself. (2;2.0)
c. Butch bite myself. (2;4.15)
d. You dress myself. (2;9.20)

If reflexive binding means that there is no additional argument, then the fact that a near-reflexive is a separable object becomes a problem with the higher interface with pragmatics or cognition.

4.1 Homogeneous and heterogeneous partitions

Based on the hypothesis that the interpretation of near-reflexives relies on partitioning the domain into subsets (here, statues as opposed to people in the flesh), we address the question in (30b) by distinguishing two pragmatic contexts:

a. a *heterogeneous partition* (containing members of heterogeneous sets, i.e. Real and Statue);
b. a *homogeneous partition* (mapping to members of a homogeneous set, i.e. the statues).

The idea is that a contrast in set types (real person, statue) or non-contrast (statue, statue) can shift the interpretation towards near-reflexivity. This can be seen in the difference between the following, which we shall show becomes much stronger with a distinctive context:

(36) [context – Only two options: Ringo, his statue → Ringo undresses his statue]
Is Ringo undressing himself? (expected answer: no)

(37) [context – Only two options: Statue of Ringo and Statue of George → Ringo undresses the statue of Ringo]
Is Ringo undressing himself? (expected answer: yes)

The intuition that near-reflexivity is more likely if the statue of Ringo is contrasted only with another statue may reflect an important fact. It suggests that the ambiguity does not boil down to the notion of *saliency* that is usually invoked in cases of referential ambiguity (think of: *Is Ringo undressing him*? where the pronoun gets assigned to the most salient character in context). Why would Ringo's statue be more salient in (37)? It seems that if the context provides a contrast between statues, then the mapping from *self* to a statue becomes acceptable. To demonstrate this difference, we sought experimental evidence.

9. Data from the CHILDES database. (35a): manchester/aran/aran07b.cha:1247; (35b): brown/eve/eve17.cha:2948; (35c): bloom70/peter/peter12.cha:3502; (35d): brown/sarah/sarah030.cha:865).

5. Experimental studies

5.1 Experiment 1: Near-reflexives in child Italian

Participants – We start by reporting the final design of a series of experiments. Twenty-nine Italian children were tested, between the ages of 4;0 and 5;11 (4-year-olds: $N = 17$, mean age 4;5; 5-year-olds: $N = 12$, mean age 5;8). In addition, 30 adults were tested (age range 18 to 53 years, mean age 34 years). Participants were told short stories followed by yes/no questions. The theme of the stories was a statue contest involving the Disney princesses. Each princess had a statue of herself (an identical figurine painted in white) to take to an exhibit and various events preceding the exhibit were presented on screen. Four test sentences involved *near-reflexive events*, i.e. the princess doing something to her statue. The verbs used were open to both transitive and reflexive uses: *liberare* 'get out', *vestire* 'dress', *coprire* 'cover' and *coricare* 'lay down'. In addition, five fillers were used, three of which contained a simple transitive predicate and two of which contained object clitics in bound variable Principle B contexts.[10] The filler stories did not involve statues and ensured that children were paying attention to the task and could recognize the princesses.

Design – The design included two fully crossed conditions, with factors: morphology (*se stessa*, *si*); pragmatics (Real/Statue, Two Statues). Children were randomly assigned to one of two lists, where they heard a story like (38a) or (38b), followed by a question containing the clitic *si* (39a) or the complex reflexive *se stessa* (39b).

10. The transitive fillers elicited yes/no answers:

(i) a. Si è messa il vestito di Aurora?
'Did she put on Aurora's dress?'
b. Ha dato il cerchio a Biancaneve?
'Did she give the hula hoop to Snow White?'
c. Ha messo il cappello a Cenerentola?
'Did she put the hat on Cinderella?'

Two *wh*- questions contained accusative clitics and elicited full NPs:

(ii) a. Chi l'ha asciugata?
'Who dried her?'
b. Chi la considera la più bella?
'Who considers her the most beautiful?'

(38) Intro: This is a story about princesses and their statues. The princesses are going to bring their statues to a show and will decide which is the best statue. Before they arrive at the show, some things happen to them.
  a. real/statue: Tiana has made a statue that looks exactly like her. But on the way to the show she thinks: "What if my statue breaks?" So she has an idea. She takes a scarf. Let's see what she's going to do with the scarf.
  b. two-statue: Tiana has two statues: one represents her and one is Belle. But on the way to the show she thinks: "What if my statue breaks?". So she takes a scarf. But there's a problem! She has only one scarf. Let's see what she decides to do with the scarf.

(39) Picture: Tiana puts the scarf on her statue (see Figure 1)
  a. clitic: *Che cosa ha fatto Tiana con la sciarpa? <u>Si</u> è coperta*?
     What did Tiana do with the scarf? Did she cover SE?
  b. se stessa: *Che cosa ha fatto Tiana con la sciarpa? Ha coperto <u>se stessa</u>*?
     What did Tiana do with the scarf? Did she cover SELF?

Figure 1. Left: real/statue (heterogeneous) scenario; Right: two-statue (homogeneous) scenario

Each list included the same four test questions, counterbalanced by pragmatic factor (invited partition: between two statues or between the real person and the statue):

(40) a. Ha vestito *se stessa*? (List A: two-statue; List B: person-statue)
     Has dressed SELF?
  b. Ha coperto *se stessa*? (List A: person-statue; List B: two-statue)
     Has covered SELF?
  c. *Si* è coricata? (List A: two-statue; List B: person-statue)
     SE lied down?
  d. *Si* è liberata? (List A: person-statue; List B: two-statue)
     SE got out?

Results – The percentage of near-reflexive interpretations is summarized in Figure 2. No adult accepted near-reflexive interpretations in the clitic sentences, and children did so only 10% of the time. In contrast, both children and adults allowed near-reflexive interpretations with the complex anaphor *se stessa* (46.5% yes answers overall for children, 18% for adults). We compared the percentage of near-reflexive answers for each age group in a 2 × 3 mixed ANOVA with within-subject factor anaphor type (*se stessa*, clitic *si*) and between-subject factor age group (four-year olds, five-year olds, adults). There was a main effect of anaphor type ($F(1, 56) = 44.40$, $p < .001$), reflecting that there were more near-reflexive interpretations in *se stessa* sentences for all groups. In addition, there was a main effect of age ($F(2, 56) = 5.57$, $p = .006$), reflecting that there were more near-reflexive answers overall for younger relative to older age groups. Acceptance of statue readings in the clitic sentences was very low in the 5-year-old group (4% yes, 1/24) whereas the rate of acceptance of near-reflexive interpretations for the complex reflexive *se stessa* was the same for the younger and the older children. The interaction between anaphor and age was not statistically significant, $F(2, 56) = 2.43$, $p = .097$, suggesting that all age groups were sensitive to the form of the anaphor.

In these stories, relevance is overwhelmingly biased towards the statues and protecting them. In fact, the whole motif is about winning the competition with the best looking statue. Nevertheless, children's interpretation of the two types of anaphors revealed a striking effect of morphology constraining the range of meanings allowed by the sentences in (40).

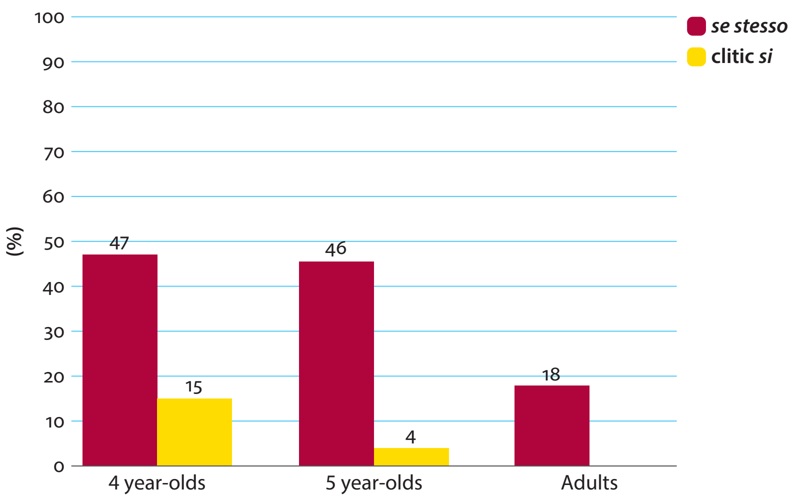


Figure 2. Percentage of near-reflexive answers according to morphology

Our environments involved two kinds of partition: homogeneous and heterogeneous. In the *se stessa* sentences, we found that both children and adults allowed near-reflexive readings more (children: 70–75%; adults: 27%) with the homogeneous case (only statues differentiated – Scenario 2). Paired sample t-tests showed that the rate of near-reflexive interpretations was significantly higher in the two-statue scenario for both children and adults ($t(28) = 5.477, p < .001, t(29) = 2.408, p = .02$).

Table 1. Percentage of NR responses in Italian according to the pragmatics of the scenarios

	Scenario 1 (REAL/STATUE)			**Scenario 2 (TWO-STATUE)**		
	4 y.o.	**5 y.o.**	**adults**	**4 y.o.**	**5 y.o.**	**adults**
SE STESSA	23.53%	16.67%	10%	**70.59%**	**75.00%**	**26.66%**
CLITIC SE	11.76%	0.00%	0.00%	17.65%	8.33%	0.00%

Children's individual profiles confirm these patterns. The majority of children allowed near-reflexivity, and were sensitive to the form of the anaphor (Type II). Of those who accepted near-reflexivity with both types of anaphors (Type III), two children answered 'pragmatically', accepting near-reflexivity in two-statues scenarios only, and two children answered "yes" to all the test questions.

Table 2. Subject profiles in Italian experiment

TYPE I: Always rejects near-reflexivity	7% (2/29)
TYPE II: Allows near-reflexivity, but only with SELF	62% (18/29)
TYPE III: Allows near-reflexivity with both SE and SELF	14% (4/29)

Discussion – With respect to our first experimental question, the result is clear-cut and supports the Dutch data in Zuckerman and Vlasveld (2004), who reported a significantly lower acceptance of proxy readings in sentences with *zich* as opposed to *zichzelf*.[11] Romance provides an interesting angle since, unlike Dutch *zich*, the clitic is the unmarked form and causes virtually no trouble in comprehension (see Zesiger et al., 2010 for a review). Italian *se stesso* is much rarer than *si*, nevertheless children made a clear association between the complex morphology and near-reflexivity.

The strength of the constraint warrants a further observation: the clitic sentences pursued an anti-pragmatic logic. If children make a choice against the thrust of the story, then they are obeying just the grammar. In this instance, clitics were rejected even in contexts where a reflexive outcome was the opposite of what the preceding discourse suggested. Such a sophisticated capacity to ignore context when UG is at work should not be underestimated, and raises questions about the methodological

11. Acceptance of *zich* in 'bad guise' sentences reported by Zuckerman and Vlasveld (2004) was low but not null (31% on average, ranging between 50% and 10%) and decreasing in the age groups tested (3- to 5-year-olds). The 'guises' in the experiment were of different kinds (statues, costumes, mirror images, video recordings) so a direct comparison with our figures would not be helpful, but the ability to discriminate between the two anaphor types is suggested by both experiments.

ideal of meeting felicity conditions in grammaticality judgment tasks. We do not find, for example, evidence to support Conroy et al.'s (2009: 454) concern that "if an interpretation is too strongly biased in the scenario provided, then a child may say 'yes' owing to contextual coercion of an ungrammatical interpretation".

With respect to our second question, namely, whether near-reflexivity can be invited by a homogeneous partition, the answer is also positive. Manipulating the event alternatives in the stories resulted in children overwhelmingly accepting a near-reflexive reading of complex anaphors in two-statue scenarios (both age groups more than 70% of the time, see Table 1), where a contrast was set out between the princess's own statue and another statue.

Near-reflexive interpretations dropped to less than a quarter when the scenario was changed. In this case, the only variable manipulated was the presence of a contrast set of statues excluding the real princess. The presence of a homogeneous set of individuals (statues), to which the partition triggered by the anaphor applied, was critical to creating the ideal conditions for near-reflexivity.

We manipulated context as the experimentation evolved, trying out different possibilities which made near-reflexivity emerge clearly. Careful analysis of the results shows that a completely new dimension of computation emerged which was to some degree more interesting, namely, a sophisticated theory of partitions. We discovered that such pragmatic dimension was an integral part of the computation required to interpret a near-reflexive, and not a mere task effect. Therefore it leads to an enrichment of the theory of near-reflexivity in pragmatic terms.

5.2 Experiment 2: Pilot English data

Participants and stimuli – At the outset of this study we tested 13 English-speaking children, aged 4;2 to 6;8 (mean age 5;7). The verbs used were *dry off, lay down, bundle up, get out, turn around, brush off, dress up*, and *sit down*. Children were shown short stories in a PowerPoint presentation followed by 8 yes/no questions. Each story was displayed as a series of still images and narrated by the experimenter. The stimuli were the same for all children.

Design – The design of the study included three experimental conditions (A, B, C) and one control condition (D). Two sentence types were tested: with a complex reflexive and with a null anaphor (41a,b). The conditions differed minimally for one factor: (i) pragmatics (A vs. B); (ii) morphology (B vs. C). The fourth condition (D) included scenarios in which the outcome was neither reflexive nor near-reflexive (for example, Cinderella dressed up another princess' statue). This condition was included to control for possible generic interpretations, under which the predicate could be construed as involving a null object (Cinderella dressed up *someone*) – although null objects are impossible in English particle constructions.

(41) a. Transitive, with direct object *SELF* (Condition A and B)
Did Cinderella dress herself up? → Yes: near-reflexive; No: reflexive
b. Intransitive, with inherent reflexive verb (Condition C and D)
Did Cinderella turn around? >→→ *Yes: near-reflexive; No: reflexive

Table 3. Design in the pilot English experiment

Pragmatics	Morphology	
	SELF	NULL
REAL/STATUE	Condition A	–
TWO- STATUES	Condition B	Condition C
TWO STATUES	–	Condition D

As in Experiment 1, we expected the inherent reflexive construction to block near-reflexivity and the complex anaphor to allow reference to a statue. In order for our questions to satisfy ideals of plausible dissent (Crain et al., 1996), namely, that both a reflexive and a near-reflexive outcome be salient enough to be negated, we first piloted our stories around two events, one reflexive and one near-reflexive:

(42) Tiana goes to bed and tells Sleeping Beauty she should go to bed too. Sleeping Beauty wants to watch the cartoons so she puts her statue on the bed.
Picture: Sleeping Beauty lays her statue on the bed
Question: *Did Sleeping Beauty lay herself down*?

This story yielded only one near-reflexive 'yes' answer. Indeed, children's answers made it clear that the event in question was negated *because* it was contrasted with the first reflexive event in the story:

(43) a. No. She put <u>her statue</u> in bed (M, 4.8)
b. No, she laid <u>her statue</u> down (A., 4.2)

Results – Overall, acceptance of near-reflexive readings was low, even when two statues were present in the stories (Table 4). Even more surprisingly, children provided more 'yes' answers in the intransitive condition (C). When we followed up their answers, it became clear that children allowed the subject NP to be interpreted as referring to the statue as well as the real princess:

(44) Did Belle turn around?
Yeah but this [points to the statue], not this [points to the real princess] (S., 5;1)

(45) *Yes, <u>one</u> turned around.*
The real Belle?
No, the statue of Belle. (J., 5;4)

Table 4. Percentage of yes answers (statue readings) in pilot experiment

Condition	Anaphor	Acceptance of statue readings
A (Real/Statue)	Self	26% (7/26)
B (Two-Statue)	Self	35% (9/26)
C (Two-Statue)	null	46% (12/26)
D (Generic)	null	11% (3/26)

Discussion – Satisfying the ideal of plausible dissent resulted in a much narrower possibility for near-reflexivity. Why does the availability of a salient true reflexive outcome block the availability of statue readings? We think an explanation in terms of partitions among events, not entities would explain these results, although we will only sketch our theory at the moment. In brief, we hypothesize that the implicit contrast between a reflexive and a near-reflexive event in the stories blocked near-reflexive reference to a statue – eliminating an ambiguity that the morphology would otherwise allow. Computing the sequence itself as a partition of events (or propositions sharing the same 'subject matter') might overrule the partitions in terms of entities if it appears to be the Question Under Discussion (Roberts, 1996).

In a sense that also deserves exploration, the experiment itself can be seen in terms of partitions. Subjects are aware that they hear many reflexives, and may seek to build sets of interpretations. In the last version of our English experiment (which the Italian experiment followed with minor improvements), we found a significantly higher proportion of near-reflexive interpretations for the first item. We opened the experiment with a self reflexive in a person/statue story (Condition A). Overall, this condition yielded significantly fewer NR answers than the two-statue condition (B). However, when the complex reflexive was presented at the start of the experiment, reference to the statue was allowed more than half of the time. Children appeared to become more 'strict' in the course of the experiment, as if the whole sequence of questions built up a metalinguistic partition which progressively marginalized the near-reflexive meaning. Crucially, however, this tendency was noticed only in person-statue scenarios and not in two-statue ones[12] – which remains the primary domain for near-reflexivity.

---

12. Repetition conspires to create this effect. Legal cases are a clear example:

Speaker A: Did you see that Bill took the money?
Speaker B: Yes, I went to the room just after him and the money was gone [to *see* = to understand]
Speaker A: Did you SEE Bill take the money? [to *see* = to look]

In effect, a repeated question creates the opposite of a 'yes bias', forcing to reconsider the meanings of the word 'see' to narrow down the core meaning.

To maximize the availability of the near-reflexive interpretation, we modified the TWO-STATUE scenarios to invite a contrast between a near-reflexive outcome (statue of self) and a non-reflexive outcome (statue of other), and conducted a second experiment.

5.3 Near-reflexives in English

Participants and stimuli – Thirty-six children who were native speakers of English were tested in a primary school in Northampton, Massachusetts, aged between 4;2 and 6;0 (4 year-olds: $N = 18$, mean age 4;7; 5 year-olds: $N = 18$, mean age 5;8). In addition, we recruited 72 adults, all native speakers of English, who submitted their judgments through a Mechanical Turk survey. The test included the same verbs as in Experiment 2. Each verb was administered in two minimally different conditions, and two lists were created such that each item would appear once in each list. Children were randomly assigned to one of two lists. The stories were revised to eliminate the saliency of a true reflexive outcome, as in Experiment 1. In the REAL-STATUE scenario, the story preceding the target question underspecified the expected outcome (e.g. *let's see what X is going to do with the scarf*). In the TWO-STATUE scenario, in contrast, there was an explicit expectation that the princess would do something to the statue of herself or to the other statue.

Design – The design was the same as in Experiment 2. We added a fifth condition in which the reflexive was in adjunct position to the subject NP (46):

(46) Cinderella is getting ready for the show. She has a very long dress but doesn't know how to wear it. So Snow White helps her put it on.
Did Cinderella herself dress up?
Expected answer: no (Snow White dressed Cinderella)

We included this condition to test children's ability to compute more complex alternatives over events. Differently from argument SELF, the adjunct here requires alternatives to be computed over the agent of the event.

Results – Both children and adults allowed near-reflexivity with SELF anaphors (see Table 5). As in our pilot study, the proportion of 'yes' answers (potentially near-reflexive) in intransitive sentences (Condition C) was higher than we would expect. Since near-reflexive interpretations are incompatible with inherent reflexive verbs, the comparable rate of yes answers among children and adults seems to point to an issue in identifying the intended referent for the subject NP.[13] In other words,

13. We tried to solve this problem by giving the statues different names, e.g. Little Cinderella, Little Snow White, and so on. Although this reduced the number of statue readings from Experiment 2 to Experiment 3, it did not solve the issue entirely.

Table 5. Proportion of 'yes' answers with SELF and null anaphors in English

	SELF (A+B)	NULL (C)	GENERIC (D)
4-year-olds	46% (33/72)	36% (13/36)	22% (8/36)
5-year-olds	36% (26/72)	27% (10/36)	8% (3/36)
Adults	25% (18/72)	27% (10/36)	2% (1/36)

this would qualify as a case of metonymy (extending NP Cinderella to the statue of Cinderella) rather than near-reflexivity.

A second trend we observed, consistent with the data reported in Zuckerman and Vlasveld (2004), is that near-reflexivity declines with age (Figure 3).

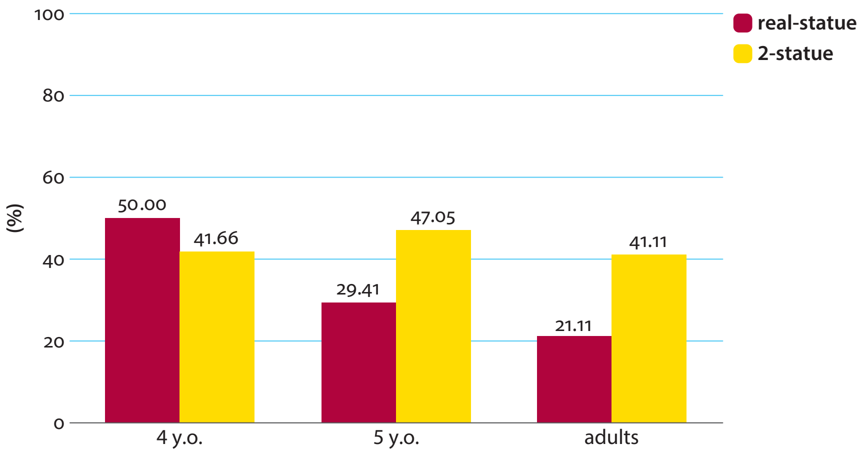


Figure 3. Proportion of yes answers with SELF anaphors in different pragmatic scenarios

Given that NR answers are always 'yes' answers, Zuckerman and Vlasveld (2004) suggest that this trend should be attributed to a 'yes bias', which is known to decline with age. However, we find that older children, and even adults, do not reduce their tolerance to near-reflexivity in all contexts. A mixed 2 (pragmatic scenario: 1 statue/2 statues) x 3 (Group: 4 year olds, 5 year olds, adults) ANOVA revealed that there was a significant interaction between pragmatics and age ($F(2,50) = 3.991$, $p = .02$). Unlike the children, the adults were significantly more likely to allow a near-reflexive interpretation when the pragmatics invited it ($t(27) = -2.915$, $p = .01$). The difference in rate of near-reflexive interpretations in the two scenarios was marginally significant in the 5 year-old group ($t(16) = -2.073$, $p = .05$). This effect of context supports the view that near-reflexivity is triggered by partition effects, and that sensitivity to this subtle contextual variable develops with age.

6. Discussion

The 4-year-olds differed from the 5-year-olds, who displayed a near-adultlike profile, in over-accepting near-reflexives in Real/Statue scenarios (A), where the partition involved a relation between heterogeneous sets of individuals (see Table 3). We need a more sophisticated notion of partition to capture the intuition that some representations are more difficult than others: why do near-reflexive representations become more effortful (hence marginal) in absence of a contrast between statues? We think – although more research is needed – that younger children have difficulty computing multiple partitions, namely, contrasting both individuals and events. To explore this question, we integrate our original hypothesis with a richer theory of partition complexity:

(47) a. FIRST ORDER partition: a partition into homogeneous sets of individuals.
b. SECOND ORDER partition: a relation between heterogeneous sets of individuals and events.

(48) Hypothesis: there is an inherent difficulty in making a second order partition. In child grammar, this ability may develop with time.

We maintain that there is a linguistic interface between complex anaphors and complex partitions, which we can illustrate with a further example that does not involve near-reflexivity. We had stories like (46) where the reflexive was an adjunct. It is claimed that the adjunct SELF creates contrast sets, and sometimes has an 'assistive-exclusive' presupposition (Eckardt, 2001). The contrast is all about the agent role, the presupposition being: *who dressed Cinderella*? In order to negate this sentence, a partition over multiple events must be computed.

In our stories, the assistive-exclusive reading was false (the princess was not an agent) but 4-year-olds crucially said 'yes' 47% of the time, 5-year-olds only 27% (and adults 6% of the time). These numbers might indicate that partition complexity is not structure-specific, and therefore is not comparable across structures in a straightforward way. Clifton and Frazier (2013) found that adults strongly prefer partitions over individuals to partitions over multiple times when interpreting plural predicates. For example, they interpret a sentence like (49) as 'some students go to the gym, some to the cafeteria', instead of 'all the students go to the gym one day, to the cafeteria the next':

(49) Every day the students go to the gym and the cafeteria for the first period
(Clifton and Frazier 2013: 620)

Syntactic mechanisms create structure-specific complexity: in the case of near-reflexivity, the simplest partition is one over individuals only (statues). In contrast, the adjunct *self*, by creating an open proposition, imposes a partition over both individuals and events. This speculation is post-hoc and our goal is simply to suggest that other dimensions of partitioning could be relevant.

7. Conclusion

Our core results are very simple although we have a taken a circuitous path engaging many examples to put them into theoretical perspective:

(i) Young children with no reliable exposure can comprehend near-reflexivity.
(ii) Near-reflexivity reflects the subtle morphological distinction between mono-morphemic and bi-morphemic reflexives.
(iii) Children can make these distinctions in anti-pragmatic environments. They prefer strict interpretation for monomorphemic reflexives even when the story promotes near-reflexivity.
(iv) We enriched the theory of pragmatic partitions to explain children's and adults' behavior across a variety of contexts. That is, we articulated a notion of partitions for homogeneous sets (among statues) and heterogeneous sets (between real and statue).

Guiding this argumentation is the claim that projection between the syntax of morphology to argument structure to the partitions available in context is essentially a set of strict interfaces stipulated by UG. The untutored contrast between English and Italian in identical experiments presented is, on our view, stronger evidence for the UG nature of the contrast than intuitional adult judgments which are open to a variety of obscuring contemplative factors.

In essence we have extended the traditional argumentation of UG from syntax to a much larger realm which becomes then a crucial ingredient in explaining the familiar, but still astonishing, rapidity and efficiency of acquisition. In addition, we have demonstrated that acquisition data can provide the most direct evidence for a subtle linguistic contrast.

Our analysis follows the classic logic of linguistics, now extended to an interface between syntax (bi-morphemic structure), pragmatics (partitions), and semantics (extension of reflexivity to include near-reflexives). These connections across formal structures reflect innate biological dispositions, not semantically or pragmatically necessary connections. Nevertheless 'technical' solutions are often the first visible form of 'leading ideas' and yet deeper explanations for such connections may yet emerge.

Acknowledgement

We thank Jeremy Hartman, Ray Jackendoff, Lyn Frazier, Jeffrey Lidz, Barbara Pearson, Roger Schwarzschild and the members of the Language Acquisition Research Group at UMass, William Snyder and the members of UUSLAW, and anonymous reviewers for various large and small comments. Errors of interpretation are ours, not theirs. Barbara Pearson provided not only intellectual commentary, but crucial assistance in design and execution of the experiment.

References

Andrews, A. D. (1990). Case structures and control in Modern Icelandic. In J. Maling & A. Zaenen (Eds.), *Syntax and semantics, Vol. 24: Modern Icelandic syntax*. San Diego, CA: Academic Press.

Avrutin, S., & Thornton, R. (1994). Distributivity and binding in child grammar. *Linguistic Inquiry*, 25, 165–171.

Baauw, S., Escobar, L., & Philip, W. (1997). A delay of Principle B effect in Spanish speaking children: The role of lexical feature acquisition. In A. Sorace, C. Heycock, & R. Shillcock (Eds.), *Proceedings of the GALA 97 conference on language acquisition* (pp. 16–21). Edinburgh: Edinburgh University Press.

Belletti, A., & Rizzi, L. (1988). Psych-verbs and θ-theory. *Natural Language & Linguistic Theory*, 6, 291–352. https://doi.org/10.1007/BF00133902

Brunetto, V. (2012). The pronoun interpretation problem in Italian complex predicates (Unpublished doctoral dissertation). University of Leeds.

Clifton Jr., C. & Frazier, L. (2013). Partition if you must: Evidence for a no extra times principle. *Discourse Processes*, 50, 616–630. https://doi.org/10.1080/0163853X.2013.850604

Conroy, A., Takahashi, E., Lidz, J., & Phillips, C. (2009). Equal treatment for all antecedents: How children succeed with Principle B. *Linguistic Inquiry*, 40, 446–486. https://doi.org/10.1162/ling.2009.40.3.446

Crain, S., Thornton, R., Boster, C., Conway, L., Lillo-Martin, D., & Woodams, E. (1996). Quantification without qualification. *Language Acquisition*, 5, 83–153. https://doi.org/10.1207/s15327817la0502_2

Eckardt, R. (2001). Reanalysing selbst. *Natural Language Semantics*, 9, 371–412. https://doi.org/10.1023/A:1014875209883

Grimshaw, J. (1990). *Argument structure*. Cambridge, MA: The MIT Press.

Higginbotham, J. (1981). Reciprocal interpretation. *Journal of Linguistic Research*, 1, 97–117.

Jackendoff, R. (1975). On belief-contexts. *Linguistic Inquiry*, 6, 53–93.

Jackendoff, R. (1992). Mme. Tussaud meets the binding theory. *Natural Language & Linguistic Theory*, 10, 1–31. https://doi.org/10.1007/BF00135357

Kayne, Richard. 1989. Facets of Romance past participle agreement. In P. Beninc (Ed.), *Dialect variation and the theory of grammar* (pp. 85–103). Dordrecht: Foris.

Keyser, S. J., & Roeper, T. (1992). Re: The abstract clitic hypothesis. *Linguistic Inquiry*, 23, 89–125.

Labelle, M. (2008). The French reflexive and reciprocal se. *Natural Language & Linguistic Theory*, 26, 833–876. https://doi.org/10.1007/s11049-008-9053-1

Lidz, J. (2001). Condition R. *Linguistic Inquiry*, 32, 123–140. https://doi.org/10.1162/002438901554603

Link, G. (1983). The logical analysis of plurals and mass terms: A lattice-theoretical approach. In R. Bauerle, C. Schwarze, & A. von Stechow (Eds.), *Meaning, use, and the interpretation of language* (pp. 127–146). Berlin: De Gruyter. https://doi.org/10.1515/9783110852820.302
Marantz, A. (1984). *On the nature of grammatical relations*. Cambridge, MA: The MIT Press.
Marelj, M., & Reuland, Eric. (2013). Clitic SE in Romance and Slavonic revisited. In I. Kor Chahine (Ed.), *Current studies in Slavic linguistics* (pp. 75–88). Amsterdam: John Benjamins. https://doi.org/10.1075/slcs.146.05mar
McGinnis, M. (1999). Reflexive clitics and the specifiers of vP. *MIT Working Papers in Linguistics*, 35, 137–160.
Pica, P., & Snyder, W. (1997). On the syntax and semantics of local anaphors in French and English. In A. M. Di Sciullo (Ed.), *Projections and interface conditions: Essays on modularity* (pp. 235–250). Oxford: Oxford University Press.
Pylkkänen, L. (2008). *Introducing arguments*. Cambridge, MA: The MIT Press. https://doi.org/10.7551/mitpress/9780262162548.001.0001
Reinhart, T., & Reuland, E. (1993). Reflexivity. *Linguistic Inquiry*, 24, 657–720.
Reuland, E. (1997). Logophoricity as orientation (Unpublished manuscript). Utrecht University.
Reuland, E. J. (2011). *Anaphora and language design*. Cambridge, MA: The MIT Press.
Reuland, E., & Winter, Y. (2009). Binding without identity: Towards a unified semantics for bound and exempt anaphors. In S. L. Devi & A. B. R. Mitkov (Eds.), *Anaphora processing and applications* (pp. 69–79). Berlin: Springer. https://doi.org/10.1007/978-3-642-04975-0_6
Roberts, C. (1996) Information Structure: Towards an integrated formal theory of pragmatics. In Jae Hak Yoon and Andreas Kathol (eds.) OSUWPL Volume 49: Papers in Semantics, 1996. The Ohio State University Department of Linguistics.
Roeper, T. (2014). Strict interface principles and the acquisition engine: From unlabeled to labeled and minimal modular contact. *Language Sciences*, 46, 115–132. https://doi.org/10.1016/j.langsci.2014.05.003
Rooryck, J., & Vanden Wyngaerd, G. (1998). The self as other: A minimalist approach to zich and zichzelf in Dutch. *Proceedings of the North Eastern Linguistic Society*, 28, 359–373.
Safir, K. J. (2004). *The syntax of anaphora*. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195166132.001.0001
Schwarzschild, R. (1996). *Pluralities*: Dordrecht: Kluwer. https://doi.org/10.1007/978-94-017-2704-4
Zesiger, P., Zesiger, L. C., Arabatzi, M., Baranzini, L., Cronel-Ohayon, S., Franck, J., Frauenfelder, U. H., Hamann, C., & Rizzi, L. (2010). The acquisition of pronouns by French children: A parallel study of production and comprehension. *Applied Psycholinguistics*, 31, 571–603. https://doi.org/10.1017/S0142716410000147
Zuckerman, S., Avrutin, S., & Vasić, N. (2002). The syntax-discourse interface and the interpretation of pronominals by Dutch-speaking children. In S. Fish & A. H-J. Do (Eds.), *Proceedings of the 26th annual Boston University Conference on Language Development* (pp. 781–792). Somerville, MA: Cascadilla Press.
Zuckerman, S., & Vlasveld, I. (2004). Reference to a 'guise' in child language. In A. Brugos, L. Micciulla, & C. E. Smith (Eds.), *Proceedings of the 28th annual Boston University Conference on Language Development* (pp. 681–688). Somerville, MA: Cascadilla Press.

Funding information

Statement: This research was supported by the Fondazione Marica De Vincenzi, which is hereby gratefully acknowledged.

Subject index

Language index

Author index

This volume presents the state of the art of recent research on the acquisition of semantics. Covering topics ranging from infants' initial acquisition of word meaning to the more sophisticated mapping between structure and meaning in the syntax-semantics interface, and the relation between logical content and inferences on language meaning (semantics and pragmatics), the papers in this volume introduce the reader to the variety of ways in which children come to realize that semantic content is encoded in word meaning (for example, in the event semantics of the verbal domain or the scope of logical operators), and at the level of the sentence, which requires the composition of semantic meaning. The authors represent some of the most established and promising researchers in this domain, demonstrating collective expertise in a range of methodologies and topics relevant to the acquisition of semantics. This volume will serve as a valuable resource for students and faculty, and junior and seasoned researchers alike.

ISBN 978 90 272 0137 9

9 789027 201379

JOHN BENJAMINS PUBLISHING COMPANY