

## Language in Interaction

# *Trends in Language Acquisition Research*

*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>

## **Series Editors**

Shanley Allen  
University of Kaiserslautern  
allen@sowi.uni-kl.de

Caroline F. Rowland  
University of Liverpool  
crowland@liverpool.ac.uk

## **Editorial Board**

Ruth A. Berman  
Tel Aviv University  
  
Morten H. Christiansen  
Cornell University  
  
Jean Berko Gleason  
Boston University  
  
Nancy Budwig  
Clark University  
  
Ewa Dąbrowska  
University of Sheffield  
  
Philip S. Dale  
University of New Mexico

Paul Fletcher  
University College Cork  
  
Steven Gillis  
University of Antwerp  
  
Annick De Houwer  
University of Erfurt  
  
Elena Lieven  
Max Planck Institute for Evolutionary  
Anthropology, Leipzig  
  
Brian MacWhinney  
Carnegie Mellon University  
  
Marilyn Vihman  
University of York

## **Volume 12**

Language in Interaction. Studies in honor of Eve V. Clark  
Edited by Inbal Arnon, Bruno Estigarribia, Marisa Casillas and Chigusa Kurumada

# **Language in Interaction**

Studies in honor of Eve V. Clark

*Edited by*

**Inbal Arnon**

Hebrew University, Jerusalem

**Bruno Estigarribia**

University of North Carolina at Chapel Hill

**Marisa Casillas**

Max Planck Institute for Psycholinguistics, Nijmegen

**Chigusa Kurumada**

University of Rochester

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.

**Library of Congress Cataloging-in-Publication Data**

Language in Interaction : Studies in honor of Eve V. Clark / Edited by Inbal Arnon, Bruno Estigarribia, Marisa Casillas and Chigusa Kurumada.

p. cm. (Trends in Language Acquisition Research, ISSN 1569-0644 ; v. 12)

Includes bibliographical references and index.

1. @@@@-@@@. I. @@@@-@@@-@@@-@@@.

@@@-@@@ @@@@ 2014

@@@-dc22

@@@

ISBN 978 90 272 4401 7 (Hb ; alk. paper)

ISBN 978 90 272 7030 6 (Eb)

© 2014 – John Benjamins B.V.

No part of this book may be reproduced in any form, by print, photoprint, microfilm, or any other means, without written permission from the publisher.

John Benjamins Publishing Co. · P.O. Box 36224 · 1020 ME Amsterdam · The Netherlands  
John Benjamins North America · P.O. Box 27519 · Philadelphia PA 19118-0519 · USA

# Table of contents

Acknowledgements	vii
List of contributors	ix
Introduction: Language acquisition in interaction <i>Chigusa Kurumada and Inbal Arnon</i>	1
<b>PART 1. The social and interactional nature of language input (five papers)</b>	
Conversational input to bilingual children <i>Susan Ervin-Tripp</i>	13
Social environments shape children's language experiences, strengthening language processing and building vocabulary <i>Adriana Weisleder and Anne Fernald</i>	29
The interactional context of language learning in Tzeltal <i>Penelope Brown</i>	51
Conversation and language acquisition: Unique properties and effects <i>Edy Veneziano</i>	83
Taking the floor on time: Delay and deferral in children's turn taking <i>Marisa Casillas</i>	101
<b>PART 2. The role of paralinguistic information in language learning (three papers)</b>	
Temporal synchrony in early multi-modal communication* <i>Barbara F. Kelly</i>	117
Shared attention, gaze and pointing gestures in hearing and deaf children <i>Aliyah Morgenstern</i>	139

- How gesture helps children learn language 157  
*Susan Goldin-Meadow*

**PART 3. Pragmatic forces in language learning (six papers)**

- Referential pacts in child language development 175  
*Gemma Stephens and Danielle Matthews*

- “We call it as puppy”: Pragmatic factors in bilingual language choice 191  
*Medha Tare and Susan A. Gelman*

- Learning words through probabilistic inferences about speakers’  
communicative intentions 207  
*Michael C. Frank*

- Word order as a structural cue and word reordering as an interactional  
process in early language acquisition 231  
*Aylin C. Küntay and Duygu Özge*

- The discourse basis of the Korean copula construction in acquisition 251  
*Patricia M. Clancy*

- Emergent clause-combining in adult-child interactional contexts 281  
*Ruth A. Berman and Lyle Lustigman*

**PART 4. Interactional effects on language structure and use (three papers)**

- Analytic and holistic processing in the development of constructions 303  
*Joan Bybee*

- From speech with others to speech for self: A case study  
of “externalized drama” 315  
*Dan I. Slobin*

- How to talk with children 333  
*Herbert H. Clark*

## Acknowledgements

We would like to acknowledge the tremendous support given by the TiLAR series editors, Shanley Allen and Caroline Rowland, who read all the chapters and provided us with insightful comments. We would also like to thank all the external reviewers who participated in the peer-review process: David Barner, Heike Behrens, Adele Goldberg, Susanne Grassmann, Petra Hendriks, Erika Hoff, Evan Kidd, Mardi Kidwell, Dave Kleinschmidt, Elena Lieven, Katherine Nelson, Elinor Ochs, Asli Özyürek, Felicia Roberts, Julie Sedivy, Ludovica Serratrice, and Susan Wagner-Cook.



## List of contributors

### Editors:

Inbal Arnon  
(inbal.arnon@gmail.com)  
Psychology department, Hebrew  
University  
Mount Scopus, Jerusalem  
  
Marisa Casillas  
(Marisa.Casillas@mpi.nl)  
Max Planck Institute for  
Psycholinguistics  
Postbus 310  
6500 AH Nijmegen, The Netherlands

Chigusa Kurumada  
(ckurumada@bcs.rochester.edu)  
Department of Brain and Cognitive  
Sciences  
University of Rochester  
Meliora Hall, Rochester, NY 14627

Bruno Estigarribia  
(estigarr@email.unc.edu)  
Dey Hall, Room 332, CB# 3170  
University of North Carolina  
at Chapel Hill

### Authors:

Ruth Berman  
(rberman@post.tau.ac.il)  
Linguistics Department, Tel-Aviv  
University  
Tel-Aviv, Israel

Penelope Brown  
(penelope.brown@mpi.nl)  
Max Planck Institute for  
Psycholinguistics  
PO Box 310  
6500 AH Nijmegen  
The Netherlands

Joan Bybee  
(joan.bybee@gmail.com)  
Professor Emerita  
Department of Linguistics  
University of New Mexico  
Albuquerque, NM 87131-0001  
USA

Herbert H. Clark  
(clark@stanford.edu)  
Jordan Hall, Room 302  
450 Serra Mall  
Stanford, CA 94305-2150

Patricia Clancy  
(pclancy@linguistics.ucsb.edu)  
Department of Linguistics, South Hall  
3607  
University of California, Santa Barbara  
Santa Barbara, CA 93106-3100

Susan Ervin-Tripp  
(ervintripp@berkeley.edu)  
Department of Psychology  
3210 Tolman Hall,  
University of California  
Berkeley CA 94720-1650

Anne Fernald  
(afernald@stanford.edu)  
Department of Psychology  
Stanford University  
Palo Alto, 94305, USA

Michael C. Frank  
(mcfrank@stanford.edu)  
Jordan Hall, Room 420-278  
450 Serra Mall  
Stanford, CA 94305-2150

Susan Gelman  
(gelman@umich.edu)  
Department of Psychology  
University of Michigan  
530 Church St.  
Ann Arbor, MI 48109-1043

Susan Goldin-Meadow  
(sgm@uchicago.edu)  
University of Chicago  
Department of Psychology  
5848 South University Avenue  
Chicago, IL 60637

Barbara Kelly  
(b.kelly@unimelb.edu.au)  
School of Languages and Linguistics,  
University of Melbourne,  
Australia

Aylin Küntay  
(akuntay@ku.edu.tr)  
Department of Psychology  
Koç University  
Rumeli Feneri Yolu  
Sariyer 34450, Istanbul

Lyle Lustigman  
Department of Linguistics  
Tel-Aviv University, Webb Building,  
Tel-Aviv, Israel

Danielle Matthews  
(danielle.matthews@sheffield.ac.uk)  
Department of Psychology  
University of Sheffield  
Western Bank  
Sheffield S10 2TP, UK

Aliyah Morgenstern  
(aliyah.morgenstern@univ-paris3.fr)  
Department of Linguistics,  
Université Sorbonne Nouvelle – Paris 3  
Paris, France

Duygu Özge  
(dozge@ku.edu.tr)  
Department of Psychology  
Koç University  
Rumeli Feneri Yolu  
Sariyer 34450, Istanbul

Dan Slobin  
(slobin@berkeley.edu)  
Professor Emerita  
Department of Psychology  
University of California, Berkeley  
USA

Gemma Stephens  
(gstephens1@sheffield.ac.uk)  
Department of Psychology  
University of Sheffield  
Western Bank  
Sheffield S10 2TP, UK

Medha Tare  
(mtare@umd.edu)  
University of Maryland  
7005 52nd Avenue  
College Park, MD 20742

Edy Veneziano  
(edy.veneziano@parisdescartes.fr)  
LaPsyDe, CNRS FRE 3521  
46 Rue St. Jacques  
75005 Paris, France

Adriana Weisleder  
(adriana.weisleder@gmail.com)  
New York University School  
of Medicine



## INTRODUCTION

# Language acquisition in interaction

Chigusa Kurumada and Inbal Arnon  
University of Rochester and Hebrew University

For more than 40 years, Professor Eve V. Clark has been teaching Introduction to Language Acquisition in the Department of Linguistics at Stanford University. It is known as one of the most demanding and work-intensive classes in the department. Students (both undergraduate and graduate) go to a nursery school every week to work on their class assignments. They collect conversational data and transcribe it, run small experiments, and write essays addressing theoretical questions. The format of the class reflects one of Eve's foundational contributions to the field of language acquisition – an emphasis on interpersonal interaction as a driving force in how children learn to talk. Eve's work has consistently highlighted the importance of looking at conversational exchanges between children and adults, and of studying language acquisition as it occurs within those conversations.

An emphasis on the role of interaction in learning was rare when Eve started out her academic career in the 1960's. In the context of Universal Grammar – with its focus on innate and abstract linguistic knowledge – few studies examined the ways in which interaction and communication shape the acquisition process. But Eve's stance on the essential role of interaction set her apart, impacting the research questions she asked and the methods she used to study them. In her seminal work, Eve showed how pragmatic forces affect children's assignment of meaning. Existing theories suggested that innate constraints guide how children learn unfamiliar words. Meaning is considered to be assigned through the application of constraints like the Whole-object assumption (Markman & Wachtel, 1988) or the Basic-level assumption (Mervis, 1984). Eve proposed instead that two pragmatic principles guides children's acquisition of meaning: The principle of Contrast – different forms have different meanings – and the more general principle of Conventionality – learning language involves learning the linguistic conventions of your speech community (Clark, 1983; Clark, 1987; Clark, 1990a). Together, these principles provided a general theory of lexical development that goes beyond the acquisition of concrete nouns. The principles make concrete predictions about how children acquire various linguistic elements (nouns, verbs, adjectives,

particles, constructions) by analysing their attribution of meaning in different interactional contexts.

Among the many domains Eve examined are relational terms such as *before* and *after*, and locative and orientation terms such as *on* vs. *under*, the meanings of which are not easily constrained by perceptible features of events or states (e.g., Clark, 1971; 1972; 1973; Clark & Garnica, 1974). Her work on this topic carefully outlined the gradual steps young learners take towards constructing a complex semantic network. Children first latch onto frequently-used and conceptually more accessible terms, and then generalize the use of these terms unless they receive evidence of competing, and more conventionalized forms to express a subset of the meaning (e.g., Clark, 1971; 1973; 1977; 1983; 1987; 1990a; 1993).

A crucial assumption behind this proposal is that lexical knowledge does not only consist of simple associations of word forms and their denotations (e.g., objects and actions). By choosing a word, the speaker chooses a perspective that highlights properties pertinent to the goal of the discourse (e.g., *before* vs. *after*, *front* vs. *back*, *come* vs. *go*, *the dog* vs. *my friend* when referring to the same pet dog). Word learning, therefore, inevitably includes acquiring knowledge about how a given word uniquely specifies the speaker's perspective, as well as pragmatic reasoning as to why the speaker picked the word in the context at hand. Eve's work has illuminated the multitude of ways in which children and adults collaboratively embark on this task of word learning, which leads to the acquisition of paradigmatic relationships between words (e.g., *dog* vs. *cat*) as well as taxonomic relationships (e.g., *dog* vs. *pet*) or context-dependent choices of referential expressions (e.g., *dog* vs. *he*) (e.g., Clark, 1990b, 1997).

Eve's work provides many examples from spontaneous conversational exchanges to show that children and adults go through countless sequences of questions, clarifications, and ratification, all gravitating towards negotiation of mutual understanding. Her examples also illustrate how children actively use their current word-forming resources to coin terms for things they have no word for yet (e.g., *plate-egg* for a fried egg; Clark, Gelman & Lane, 1985; Clark & Berman, 1984; Clark & Hecht, 1982; see also Clark & Clark, 1979 for examples of adults' creative uses of nouns as verbs). Likewise, in comprehension, children assign novel forms they hear to fill gaps in their lexical knowledge (e.g., Clark, 1987; Clark, 1990a). Communication thus guides children to innovatively expand their lexicon while also motivating them to give up idiosyncratic word uses.

Over the years, more work in linguistics and psychology has acknowledged the importance of communication in language learning and change. This shift in perspective was impacted by the development of new theoretical paradigms that emphasized the links between language use and language structure (e.g., Barlow & Kemmer, 2000; Bates & McWhinney, 1982; DuBois, Kumpf, & Ashby, 2003; Fox,

Jurafsky, & McDaniel, 1999; Goldberg, 1995; Tomasello, 1998, 2003a). In particular, the development of usage-based approaches to language learning (e.g., Tomasello, 2003b), and the expansion of tools for studying children in interactional contexts have led to an increased interest in the role of interaction in language learning. The development of large-scale corpora collections like CHILDES (MacWhinney, 2000) has allowed researchers to examine children's conversational settings in detail, tracking the way interactional settings and input statistics impact learning trajectories (e.g., Clark & Kelly, 2006; Goldberg, 2006; Tomasello, 2003b).

In recent years, the role of interaction and communication has been investigated in domains as diverse as word learning (e.g., Frank, Goodman, & Tenenbaum, 2009), statistical learning (e.g., Lieven, 2010), lexical and grammatical development (e.g., Berman & Slobin, 1994; Goldberg 2006), and the content of speech directed toward children (e.g., Chouinard & Clark, 2003; Clark & Bernicot, 2008). The idea that communicative needs and pressures play an important role in shaping how we learn and use language has been applied not only in first language acquisition, but also in the study of cross-linguistic or cross-situational variation (e.g., Arnon & Clark, 2011; Evans & Levinson, 2009); the relation between learnability and the typological distribution of word order and case-marking (e.g., Culbertson et al., 2012; Fedzechkina et al., 2012), the cultural evolution of language (e.g., Griffiths, Kalish & Lewandowsky, 2008; Kirby, 2011; Scott-Phillips & Kirby, 2010) and the relation between human's socio-pragmatic skills and their unique language capacity (e.g., Herrmann et al., 2007).

Despite recent heightened interest, however, there are many open questions about the way interaction impacts language learning. One challenge is to turn our knowledge about the effect of *input* on language learning into a theory on *interaction*, which captures reciprocal and spatio-temporally coordinated nature of adult-child conversations. In many studies, looking at input means pooling together types and tokens of particular sounds, words or constructions. While informative, a great deal of information is lost in this process of abstraction about why and how each piece of linguistic data was given at each point of time. Another challenge has to do with the definition of the interactional context: to study the effect of interaction we need to clearly define what the relevant context is and how it may differ across speakers and communicative goals. A further challenge is to examine how interactional patterns differ between populations (e.g., monolinguals vs. bilinguals, Western vs. non-Western societies) and learning situations. As in many domains, a lot of our knowledge about what interactional contexts comes from a rather restricted pool of participants. A theory of interaction in language acquisition must account for both the consistency and variability across different learners and learning situations.

With various collaborators, Eve has touched upon many of these issues. She provided detailed analyses of how adults talk to young children at different stages of language development. She asked questions such as: How do adults maintain children's attention and invite them into interactions (Estigarribia & Clark, 2007)? How do they introduce new words and new information in common ground (Clark & Wong, 2002; Clark & Amaral, 2010)? How do they ratify children's utterances and recast incorrect usages (Chouinard & Clark, 2003; Clark, 2004, 2007; Clark & Bernicot, 2008; Clark, 2010, Clark & de Marneffe, 2012)? Eve and her colleagues have also looked at the paralinguistic cues used to mark communicative intent in conversations with children, such as prosody, gestures, gaze, and facial expressions (Clark, 1978; 1980; 2001; Clark & Estigarribia, 2011).

In this book, we hope to both pay tribute to Eve's long-standing and significant contributions to the field of language acquisition while also presenting a timely response to the field's renewed interest in the social and interactional aspects of language learning. We focused on several challenges facing the study of interactional effects on language learning. In the first part, we ask what constitutes an interactional context and how this differs across languages and speakers (e.g., monolingual and bilingual, high and low socio-economic status). In the second part, we examine the range of paralinguistic cues available to children (joint attention, gaze, gesture) and ask how these cues affect language learning, and how the effects may differ for different populations (e.g., hearing vs. deaf children). In the third part, we ask how pragmatic forces impact the course of language learning in several domains (bilingual language choice, word learning, construction learning). In the fourth part, we consider interaction more broadly to ask what role it plays in adult language use and in language change. Together, these four parts provide us with a comprehensive view of the role interaction in language acquisition while also highlighting the challenges of conducting research on such a multi-dimensional aspect of human behavior.

The five chapters in Part 1: *The social and interactional nature of language input* examine the differences and similarities in the interactional contexts of different children. **Ervin-Tripp** provides a theoretical overview of the factors shaping children's bilingual development. While presenting the unique nature of bilingual acquisition, Ervin-Tripp draws our attention to commonalities in the effects of communication on monolingual and bilingual development. This chapter shows how inter-personal communication and language acquisition are influenced by multiple variables including institutional and socio-economical status (SES) of languages and their speakers. **Weisleder and Fernald** further highlight the effect of SES on language development. They review new research showing that the amount and quality of child-directed speech in infancy contributes to the development of language processing skills, which in turn facilitate

vocabulary growth. **Brown** looks at caregiver-child interactions in a Tzeltal Mayan community to examine the effect of culture on interaction. The chapter raises questions about what we think of as a stereotypical form of parental speech. Brown provides a number of examples to illustrate that, compared to a western standard, adult Tzeltal speakers are far less likely to direct their speech to young children. The examples are then used to discuss how these interactions differ from those reported in much of the literature on language learning, much of which is based on the experiences of Western children. **Veneziano** explores what it is that makes conversation such a prime location for learning in general, and language learning in particular. She focuses on the aspects of exchange and accommodation inherent in conversation that make successful coordination and collaboration possible. **Casillas** examines how children can hold the floor to maximize their contribution to an interaction. This way, they manage themselves the language input they hear and can modulate the feedback they receive from adults. In particular, the author shows that differentiation of two common delay markers, *uh/er* and *um/erm*, happens before the age of four, with *um* signaling longer delays just as in the adult language.

The second part: *The role of paralinguistic information in language learning*, contains three chapters looking at the effect of gaze, gesture and attention on language learning. **Kelly** provides evidence for the role of gesture-speech timing as a motor of children's transition towards multiword combinations. Kelly argues that children's acquisition and automatization of the synchronous use of gesture and speech is a necessary cognitive precursor for gesture and speech to be used synchronously to designate different meaning, prefiguring two-word combinations. In so doing, she points out how caregivers' interpretations of gestures can both help determine referents and paves the way from co-referential to non-co-referential speech-gesture combinations. **Morgenstern** explores how deaf infants learn to achieve joint attention through the exclusive use of the visual modality (as opposed to visual and auditory modality combination), and how signing children learn to grammaticalize gesture and gaze. The author shows that gaze is recruited by signing children more often as a means to check and manipulate the caregiver's attention. Moreover, pointing gestures are gradually replaced by deictics in the hearing child's speech, but they appear grammaticalized earlier in the deaf signing child. **Goldin-Meadow** similarly shows how gestures provide non-verbal practice in producing conversational contributions, hence predicting the onset of sentences and more complex constructions in later speech. Importantly, this happens in a context where parents use information gleaned from child gestures. They adjust their language level and provide the input necessary for the acquisition of the linguistic feature prefigured by the gesture.

The five chapters in the third part: *Pragmatic forces in language learning* illustrate the effect of pragmatic information on learning of various linguistic domains. **Stephens and Matthews** review the literature on adults' referential pacts and link it to child experimental studies in the same vein. They link these results to Eve's Principles of Convention and Contrast, and draw an elegant link between adult and child behavior on the same tasks showing how children are sensitive to these principles in establishing reference. **Tare and Gelman** take us into the bilingual realm, looking at how children acquire the skill to identify and address speakers in the appropriate language across different interactional tasks. They discuss the roles of external (types of interactions) and internal (Theory of Mind, language ability) skill sets that may enable children to do this successfully. **Frank** provides a formal account of what pragmatic forces might look like for computational models of word learning. He discusses how statistical learning over the course of time can be integrated with within-context pragmatic inferences to inform children's learning of word meanings. **Küntay and Özge** bring together two apparently disparate views – *language-as-product* and *language-as-action* – on children's acquisition of flexible word order. They conclude that experimental and naturalistic complement one another in revealing how interaction and linguistic skills play out together in children's structural and pragmatic interpretations of word order. **Clancy** provides a longitudinal analysis of the copula construction produced by two young Korean-speaking children and their mothers. The primary goal of the chapter is to explore the discourse basis of the construction over developmental time. Because of its structural simplicity and its identifying and naming functions, the copula construction can appear in child's utterances from early on. However, the full functionality and appropriate encoding of information structure and the speaker's perspectives are acquired only gradually through repetitive uses and negotiation with conversational partners. **Berman and Lustigman** conduct a detailed corpus study of early clause combination in the speech of three Hebrew-acquiring girls in self-initiated autonomous constructions compared to three types of interlocutor-supported contexts. They document the gradual increase in the number and complexity of clause combinations and suggest that interactive contexts play a role in advancing the consolidation of early complex syntax.

The final part of the book: *Interactional effects on language structure and use*, expands the scope of inquiry to adult language, language change, and the development of non-linguistic skills. **Bybee** asks how interaction impacts the process of language change for constructions. She presents a corpus analysis over several decades of American English to show how contexts of use influence the development of the minor construction from an idiom: *not have two Xs to rub together*. **Slobin** presents case study of a particular kind of speech-for-self produced by a preschool-aged girl, characterized as "externalized dramas." This unique conversational setting

– where a conversation takes place between two voices of the same participant – is seen as a platform for practicing and refining pragmatic devices while dealing with extra-linguistic concerns such as emotional states and other minds. Using examples from spontaneous conversation, **Herb Clark** illustrates how grounding facilitates communication between parents and their young children, and how it affects the forms of child-directed speech, as well as adult conversation.

The papers in this volume do not only pay tributes to Eve's work, but also critically reexamine some of its assumptions or suggest alternative views. We feel this is a true reflection of Eve's intellectual voice and her commitment to critical and rigorous scientific investigation. In keeping with her scholarship, we attempted to provide an overview of various approaches without concealing or smoothing out any disagreements among them. With our respect to Eve's contribution to the field as a common thread, the papers in this volume will provide a point of departure for our future investigations.

We would like to end this introduction with a note on Eve's influence as a teacher, advisor, and mentor. When you interact with Eve even for a short while, it becomes immediately clear that she is an extremely charming person with a warm heart (spiced up with her lovely British accent). At the same time, Eve is an honest and fierce discussion partner who will not let any inaccurate comments slide, or accept any conjectures that are not attested in the data. In this, she embodies a great interactional partner – one who listens, understands, recasts, rephrases, and encourages us to try again. With you, Eve, “language acquisition in interaction” is not a mere theory or a hypothesis: We all learned to speak our academic language through interactions with you. This volume is our thank-you note, and an attempt to open up our conversations with you – and the deep insights they generate – to a wider audience.

## References

- Arnon, I., & Clark, E.V. (Eds.). (2011). *Experience, variation, and generalization: Learning a first language*. Amsterdam: John Benjamins. DOI: 10.1075/tilar.7
- Barlow, M., & Kemmer, S. (2000). *Usage based models of language*. Stanford, CA: CSLI.
- Berman, R.A., & Slobin, D.I. (Eds.). (1994). *Relating events in narrative: A crosslinguistic developmental study, Vol. 1*. Hillsdale, NJ: Laurence Erlbaum Associates.
- Bates, E., & MacWhinney, B. (1982). Functional approaches to grammar. In E. Wanner, & L. Gleitman (Eds.), *Language acquisition: The state of the art* (pp. 173–218). Cambridge: CUP.
- Chouinard, M.M., & Clark, E.V. (2003). Adult reformulations of child errors as negative evidence. *Journal of Child Language*, 30, 637–669. DOI: 10.1017/S0305000903005701
- Clark, E.V. (1971). On the acquisition of the meaning of before and after. *Journal of Verbal Learning & Verbal Behavior*, 10, 266–275. DOI: 10.1016/S0022-5371(71)80054-3

- Clark, E.V. (1972). On the child's acquisition of antonyms in two semantic fields. *Journal of Verbal Learning & Verbal Behavior*, 11, 750–758. DOI: 10.1016/S0022-5371(72)80009-4
- Clark, E.V. (1973). How children describe time and order. In C.A. Ferguson & D.I. Slobin (Eds.), *Studies of child language development* (pp. 585–606). New York, NY: Holt, Rinehart & Winston.
- Clark, E.V. (1977). Strategies and the mapping problem in first language acquisition. In J. Macnamara (Ed.), *Language learning and thought* (pp. 147–168). New York, NY: Academic Press.
- Clark, E.V. (1978). From gesture to word: On the natural history of deixis in language acquisition. In J.S. Bruner & A. Garton (Eds.), *Human growth and development: Wolfson College lectures 1976* (pp. 85–120). Oxford: OUP.
- Clark, E.V. (1980). Here's the "Top": Nonlinguistic strategies in the acquisition of orientational terms. *Child Development*, 51, 329–338.
- Clark, E.V. (1983). Convention and contrast in acquiring the lexicon. In T.B. Seiler & W. Wannenmacher (Eds.), *Concept development and the development of word meaning* (pp. 67–89). Berlin: Springer. DOI: 10.1007/978-3-642-69000-6\_5
- 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. (1990a). On the pragmatics of contrast. *Journal of Child Language*, 17(2), 417–431. DOI: 10.1017/S0305000900013842
- Clark, E.V. (1990b). Speaker perspective in language acquisition. *Linguistics* 28, 1201–1220. DOI: 10.1515/linc.1990.28.6.1201
- Clark, E.V. (1993). *The lexicon in acquisition*. Cambridge: CUP. Reprinted 1994; paperback edition, 1/1995. DOI: 10.1017/CBO9780511554377
- Clark, E.V. (1997). Conceptual perspective and lexical choice in acquisition. *Cognition*, 64(1), 1–37. DOI: 10.1016/S0010-0277(97)00010-3
- Clark, E.V. (2001). Grounding and attention in the acquisition of language. In M. Andronis, C. Ball, H. Elston & S. Neuvel (Eds.), *Papers from the 37th meeting of the Chicago Linguistic Society, Vol. 1* (pp. 95–116). Chicago, IL: Chicago Linguistic Society.
- Clark, E.V. (2004). Pragmatics and language acquisition. In L.R. Horn & G. Ward (Eds.), *Handbook of pragmatics* (pp. 562–577). Oxford: Blackwell.
- Clark, E.V. (2007). Young children's uptake of new words in conversation. *Language in Society* 36, 157–182. DOI: 10.1017/S0047404507070091
- Clark, E.V. (2010). Adult offer, word-class, and child uptake in early lexical acquisition. *First Language*, 30 (3–4), 250–269. DOI: 10.1177/0142723710370537
- Clark, E.V., & Amaral, P.M. (2010). Children build on pragmatic information in language acquisition. *Language & Linguistics Compass* 4(7), 445–457. DOI: 10.1111/j.1749-818X.2010.00214.x
- Clark, E.V., & Berman, R.A. (1984). Structure and use in the acquisition of word-formation. *Language*, 60, 542–590. DOI: 10.2307/413991
- Clark, E.V., & Bernicot, J. (2008). Repetition as ratification: How parents and children place information in common ground. *Journal of Child Language*, 35, 349–371. DOI: 10.1017/S0305000907008537
- Clark, E.V., & Clark, H.H. (1979). When nouns surface as verbs. *Language*, 55, 767–811. DOI: 10.2307/412745

- Clark, E.V., & de Marneffe, M.-C. (2012). Constructing verb paradigms in French: Adult construals and emerging grammatical contrasts. *Morphology*, 22, 89–120. DOI: 10.1007/s11525-011-9193-6
- Clark, E.V., & Estigarribia, B. (2011). Using speech and gesture to inform young children about unfamiliar word meanings. *Gesture*, 11(1), 1–23. DOI: 10.1075/gest.11.1.01cla
- Clark, E.V., & Garnica, O.K. (1974). Is he coming or going? On the acquisition of deictic verbs. *Journal of Verbal Learning & Verbal Behavior*, 13, 559–572. DOI: 10.1016/S0022-5371(74)80009-5
- Clark, E.V., Gelman, S.A., & Lane, N.M. (1985). Noun compounds and category structure in young children. *Child Development*, 56, 84–94. DOI: 10.2307/1130176
- Clark, E.V., & Hecht, B.F. (1982). Learning to coin agent and instrument nouns. *Cognition*, 12, 1–24. DOI: 10.1016/0010-0277(82)90027-0
- Clark, E.V., & Kelly, B.F. (2006). Constructions and acquisition. In E.V. Clark & B.F. Kelly (Eds.), *Constructions in acquisition* (pp. 1–14). Stanford, CA: CSLI.
- Clark, E.V., & Wong, A.D-W. (2002). Pragmatic directions about language use: Words and word meanings. *Language in Society*, 31, 181–212.
- Culbertson, J., Smolensky, P., & Legendre, G. (2012). Learning biases predict a word order universal. *Cognition*, 122(3), 306–329. DOI: 10.1016/j.cognition.2011.10.017
- Du Bois, J.W., Kumpf, L.E., & Ashby, W.J. (Eds.). (2003). *Preferred argument structure: Grammar as architecture for function*. Amsterdam: John Benjamins. DOI: 10.1075/sidag.14
- Estigarribia, B., & Clark, E.V. (2007). Getting and maintaining attention in talk to young children. *Journal of Child Language*, 34, 799–814. DOI: 10.1017/S0305000907008161
- Evans, N., & Levinson, S.C. (2009). The myth of language universals: Language diversity and its importance for cognitive science. *Behavioral and Brain Sciences*, 32(5), 429–492. DOI: 10.1017/S0140525X0999094X
- Fedzechkina, M., Jaeger, T.F., & Newport, E.L. (2012). Language learners restructure their input to facilitate efficient communication. *Proceedings of the National Academy of Sciences*, 109(44), 17897–17902.
- Fox, B.A., Jurafsky, D., & Michaelis, L.A. (Eds.). (1999). *Cognition and function in language*. Stanford, CA: CSLI.
- Frank, M.C., Goodman, N.D., & Tenenbaum, J.B. (2009). Using speakers' referential intentions to model early cross-situational word learning. *Psychological Science*, 20, 578–585. DOI: 10.1111/j.1467-9280.2009.02335.x
- Goldberg, A.E. (1995). *Constructions: A construction grammar approach to argument structure*. Chicago, IL: University of Chicago Press.
- Goldberg, A.E. (2006). *Constructions at work: The nature of generalization in language*. Oxford: OUP.
- Griffiths, T.L., Kalish, M.L., & Lewandowsky, S. (2008). Theoretical and experimental evidence for the impact of inductive biases on cultural evolution. *Philosophical Transactions of the Royal Society*, 363, 3503–3514. DOI: 10.1098/rstb.2008.0146
- Herrmann, E., Call, J., Hernandez-Llorada, M.V., Hare, B., & Tomasello, M. (2007). Humans have evolved special skills of social cognition: The cultural intelligence hypothesis, *Science*, 317, 1360–1366. DOI: 10.1126/science.1146282
- Kirby, S. (2011). Language as an adaptive system: The role of cultural evolution in the origins of structure. In M. Tallerman & K. Gibson (Eds.), *Oxford handbook of language evolution*. Oxford: OUP.

- Lieven, E. (2010). Input and first language acquisition: Evaluating the role of frequency. *Lingua*, 120, 2546–2556. DOI: 10.1016/j.lingua.2010.06.005
- MacWhinney, B. (2000). *The CHILDES project: Tools for analyzing talk*, 3rd Edition. Mahwah, NJ: Lawrence Erlbaum Associates.
- Markman, E.M., & Wachtel, G.F. (1988). Children's use of mutual exclusivity to constrain the meanings of words. *Cognitive Psychology*, 20(2), 121–157. DOI: 10.1016/0010-0285(88)90017-5
- Mervis, Carolyn B. 1984. Early lexical development: Contributions of mother and child. In C. Sophian (Ed.), *Origins of cognitive skills* (pp. 339–370). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Scott-Phillips, T.C., & Kirby, S. (2010). Language evolution in the laboratory. *Trends in Cognitive Sciences*, 14, 411–417. DOI: 10.1016/j.tics.2010.06.006
- Tomasello, M. (Ed.). (1998). *The new psychology of language: Cognitive and functional approaches to language structure*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Tomasello, M. (Ed.). (2003a). *The new psychology of language, Volume 2: Cognitive and functional approaches to language Structure*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Tomasello, M. (2003b). *Constructing a language: A usage-based theory of language acquisition*. Cambridge, MA: Harvard University Press.

PART 1

**The social and interactional nature  
of language input (five papers)**



# Conversational input to bilingual children

Susan Ervin-Tripp  
University of California, Berkeley

All children learn from conversational input, including children who have simultaneous or successive acquisition of several languages in childhood. Partners who stimulate speech by asking questions, responding, and expanding a child's utterances can accelerate learning at all levels of language, in addition to scaffolding the development of speech acts, speech events, social styles and appropriate code-switching. In family bilingualism with a person dichotomy, the frequency of conversation and the instructional style of each parent can affect the rate of acquisition. Bilingualism in minority languages is most successful when parents maintain a language different from the community, so there is a place dichotomy. Children whose second language is acquired in school can become fully bilingual but there is wide variation in acquisition time.

*The consistent patterns of adult offer and child uptake support the view that lexical acquisition takes place in the course of conversation.*

Eve V. Clark (1999: 1)

## Conditions of child first language acquisition

Children begin learning about language in the womb; from birth they react to features that were in the auditory input (Fernald, 2006). Clark takes a conversational view of acquisition, "Children – like adults – are sensitive to certain basic pragmatic conditions, namely, that (a) the speaker and addressee share a joint focus of attention during any conversational exchange; (b) the speaker and addressee make use of physical co-presence in identifying referents; and (c) the speaker and addressee make use of linguistic co-presence in identifying referents" (Clark, 1999:4). Speech acts too are learned from interaction. Children can use gaze, pointing and reaching to assert and request (Clark & Amaral, 2010).

### *First language success*

We know that there is considerable variation in the level of vocabulary and syntactic complexity achieved by child monolinguals of the same age. Gordon Wells (1985), who may have been the first to study the natural context of acquisition by random audio recording in homes, noted that if interactional variables are entered into prediction equations first, 95% of the variation is accounted for. The interactional variables Wells found to be predictors were speech to the child, encouraging the child's exploration and questioning, reading to the child, and speech exchanges such as questioning and responding to the child's initiatives.

In a pioneering observational study starting at 2 1/2, Hart and Risley (1992) identified in a multiple regression three factors in early parent communicative behavior. These we could call *participation* by joining in and talking often and with many different words, *accommodation* by taking turns and responding in comprehensible style, and *constructive involvement* by asking questions, repeating or expanding the child's utterance, and encouraging rather than prohibiting exploration. These studies showed that early parent interaction was related to higher vocabulary and MLU and reading scores in children in the first years of elementary school (Walker et al., 1994).

One test of this need for cooperative interaction is interference with learning. Using a cross-lagged longitudinal design, Camaioni and Longobardi (1994) showed that if mothers interrupted toddler's activities by overlapping the child's speech, ignoring topic initiations, or changing the topic, they significantly delayed the development of vocabulary and fluency four months later. If mothers supported children's own initiatives by repeating, expanding, reformulating, paraphrasing, referring to shared knowledge, replying, and using activity or game routines, their children were linguistically precocious. They found that replies to child-initiated discourse are crucial in accelerating or retarding language development. Mother-initiated talk did not have these effects.

In the next sections we will apply these acquisition results to various kinds of child bilingualism, looking at factors in success, stages, types of home bilingualism and home instructional exchanges, learning at school, and learning pragmatic skills such as speech act exchanges, codeswitching, and speech event structure.

### **Bilingual development**

Some children hear more than one language from birth; others move to a different country, or in going to school and the playground they hear a second language later. These have been called *simultaneous* vs. *successive* acquisition. De Houwer

(2009) has distinguished bilingual first language acquisition (BFLA) from child second language acquisition (CSLA). Meisel (2004) suggests CSLA is learning of the second language between five and ten. Because of changes in a bilingual child's interactional milieu, we will see there are shifts between (a) balanced speech in several languages, (b) speech dominance in one of the languages, (c) comprehension only in one without productive speech, or (d) language loss. Even siblings could differ in language history. These shifts may be why Pearson et al. (1993) found that receptive, but not productive, vocabulary of bilingual children was comparable to monolingual norms.

There may be different optimal ages for acquiring different components of language skill. Hyltenstam and Abrahamsson (2003) suggest that the optimal age for phonology precedes syntax, and that learning specific phonetic components may optimize at different ages up to five. Ervin-Tripp's (1974) findings differed.

### *Factors in successful child bilingualism*

Eilers et al., (2006) found that the pressures to become a monolingual English speaker are strong in Miami, Florida. Children were most likely to become and remain Spanish-English bilinguals if their families consistently spoke only Spanish at home and were of low socio-economic level, and the children went to a two-way immersion school program providing a bilingual day. Gathercole and Thomas (2009) confirm this contrast in Wales where English is dominant; only Welsh use at home led to bilingualism. De Houwer (2009) found in Flanders that the maximum bilingualism – 96% – occurs when both parents speak only the minority language at home, and only one speaks the community language. In families that used the traditional one person/one language, only 74% produced bilingual children. In Wales, in Flanders, in English-Japanese families in Japan, and in Miami, the maximum bilingual outcome was from home use of only the minority language (De Houwer, 2009:113).

These studies have shown that if there is a choice, it may be a mistake for immigrant or diplomatic families to switch from their home language. Child performance in schools is better when the parents use their best language to the children, perhaps because they provide more linguistic stimulation then (King et al., 2008).

Successful CSLA at school, according to studies by Wong Fillmore (1991) depends on both the cognitive and social abilities of the learner, and the way the learner is taught. Children high in sociability profit from social settings where there are many L2 speakers who can act freely with them and want to. L2 learners who are segregated together do not have peer models. If learners or their peers are shy, reclusive, laconic, or socially inept there is less learning from peers. But in tightly structured teacher-centered school contexts, cognitive abilities matter

more, such as attentiveness, verbal memory, auditory memory, and pattern recognition. Because of these individual differences, there were contrasts as great as five years in the time it takes a schoolchild to get a working command of a second language from school (Wong Fillmore, 1991: 61).

### *Age and stage in BFLA*

Crystal's (1976) description of monolingual development is of five stages: I (1 word), II (2 words), III (3 elements), IV (simple sentences), V (clauses and co-ordinates). Most case histories of BFLA show differences in developmental rate in the languages for dichotomy-speaker families.

An illustration of these stages can be seen in the study reported by Juan-Garau and Pérez-Vidal (2001) of a Barcelona multilingual family in which the mother spoke Catalan and the father English to the child. At 1;11 Andreu was at Stage I in both languages, at 2;2 at stage II in Catalan, at 2;8 at III in Catalan and II in English. At this point there was a visit to English grandparents, and at 3;2 both languages are at Stage II. By 3;5 Catalan is at IV, English at III, and at 3;8 Catalan is at IV/V, English at early IV.

Lyon (1996) did a similar analysis of three Welsh-English children who became bilingual in a Welsh community where English dominates outside the home, but the mothers spoke Welsh. In Lyon's table (1996: 139) we see that as the Welsh of each child grew from stage I to V, the English did not reach stage III until the Welsh was at IV in one child and V in two children. The mother's variety was always acquired faster. In the case where the father spoke Welsh, the children did not become bilingual but spoke only English.

De Houwer (2009:45) described productive dominance. She found in Flanders that by four there was great variation between BFLA children in productive speech choices, between language A, language B, and code-mixing in an utterance. The range of child usage patterns went from speech only in the child's dominant language, to half in each, or half speech in the dominant language and half mixed, with some patterns between these extremes. Her book explores bilingual development in each aspect of language in detail.

Fernald (2006) points out that children with bilingual input are slower to learn the acquired phonological equivalences of monolinguals. Accented speech in parents could delay a child's ability to distinguish lexicon by input phonology. With respect to competence in each variety, Meisel (2004) argues that by the second year in child bilingual input cases, there are differentiated syntactic systems, not a single system. Even in mixed utterances, children switch at appropriate places.

## Contexts for learning

Interactional contexts have two functions in learning. One is that some contexts are more likely to facilitate learning of all levels of language knowledge and skill. The other is that pragmatic knowledge is part of what is to be learned in both languages. Hymes (1974) and Schiffrin (1987) have proposed schemas for the pragmatic structure of interaction, which have influenced Table 1 (Ervin-Tripp 1995).

**Table 1.** Levels in Interaction.

	Plane	Units	Examples
Context	Personnel	Status	Mother
	Setting	Place; equipment	Playground
	Activity	Phase	Role-playing
Verbal level of linguistic choice	Speech event	Episode; Local discourse; Genre	Instructional exchange
	Participation in episode	Turn	<i>Hello</i>
	Action	Exchange of speech acts	Request/refusal
	Ideation	Topic; Propositions	<i>Who is that?</i>
	Social meaning	Social markers; Key	<i>Please, Mr. Reilly</i>
	Text	Language; register	Code-switching

### *Instructional exchanges*

The optimal situation for language learning, judging from the monolingual evidence of Wells (1985), Hart and Risley (1992) and Walker et al. (1994) is a high frequency of Instructional Exchanges with the child's parents or peers.

In her Australian study of German-English bilingual families, Döpke (1992) distinguished **insisting strategies** that required no response from the child, and **response-eliciting strategies**. Because of the strong one person-one language rule in the Norwegian-English families that Lanza (1997:273) studied, she determined that children who mixed were signaling a lexical gap. The examples below were maternal (anglophone) moves to fill the lexical gaps when the child mixed or used Norwegian to the mother, using increasingly permissive response strategies described in Döpke's categories:

(1) **Minimal grasp/request clarification**

Mo: The cow! What's the cow doing right there?

Si 2;7: *Spis/*

> Mo: Hm?

S: Eat

(2) **Expressed guess or clarification question**

Mo: What happened?  
Si(2;1): *Miste/misse/* [lose/drop]  
> Mo: *Miste?* Did she lose her – uh –her purse?  
Siri: Yeah  
Mo: She dropped her purse. And look! All of the food came out.  
Yeah. Mm.

(3) **Insisting: Native speaker repetition or translation (most frequent)**

Siri(1;11): *Ben^ ben^* [leg]  
> Mo: His leg! It's broken. He has a cast on his leg.

(4) **Move-on strategy**

Si 2;2: *Takke mat* [thank food]  
God [good]  
Mo: Hm?  
Siri: *Takk* [thanks]  
> Mo: You're welcome

(5) **Code-switching**

Mo: (reading “Pancakes for Breakfast”}  
And she didn't have any eggs, so what did she do?  
To 2:0: *hane* [young rooster]  
> Mo: She went out to the *hØne*, to the hens. And it's the  
hens that make the eggs.  
Tomas: *der war hens/*

We see similar instructional strategies in the Juan-Garau & Pérez Vidal (2001) data on a Catalan-English speaking family. A special method the English-speaking father used was the addition of an English-speaking puppet that he made speak with a high-pitched voice. The addition of a puppet voice allowed more insistence, since the puppet could not speak Catalan. The puppet is using a Minimal Grasp strategy that elicits a translation.

(6) **Minimal grasp by puppet**

An 3;2: look!  
Fa: what?  
An: *un forat* [hole] *un forat* up there.  
>Pu: what's that, *un forat?* [high voice]  
An: a hole  
Pu: a hole, yes [high voice] Juan-Garau & Pérez-Vidal:77

When Andreu was 3;2, the puppet used Minimal Grasp three times, indicating that he didn't understand what Andreu said in Catalan. The father found a way to

supply an English Repetition-translation via the puppet, and then in his own voice made a Repetition for pronunciation that the child repeated.

(7) Minimal grasp>Translation>Repetition

- An: *va buscar, vol menjar peixos* [he's looking, he wants to eat fish]
- >Pu: what's that? [high voice]
- An: *vol menjar peixos*
- >Fa: well he doesn't understand that, Sooty.
- An: *vol menjar peixos*
- >Fa: hey.
- Fa then ah, well no, I'll talk to Sooty [whispers to puppet]
- >Pu: he wants to eat fishies! [high voice]
- An: to-to eat pishies
- >Fa: fishies!
- An: *ui ui ui*
- Fa: (ah!)
- An: (x) fishies!
- Pu: he wants to eat fishies. [high voice]

- Juan-Garau & Pérez-Vidal:77-78

### *Bilingualism in the home*

The first contrast in Table 1 is the level of setting, personnel and activity. Schmidt-Mackey (1977) points out that there are different types of bilingual contexts. Within the household, *dichotomy-person* is the one person, one language family. Children in immigrant families may have a *dichotomy-place* system, using the community language outside the household in nursery school and playground, and another language inside the home. A *dichotomy-time, topic, or activity* allows for many contexts of alternation, such as a mealtime or storybook switch in a household. In dual-immersion schools, any of these may occur. And then of course there is alternation of languages, or code-switching by topic, key, or genre, which normally is a speaker's option in bilingual conversation.

The first detailed case studies of developing bilinguals were in the home setting, of the children of linguists Ronjat (1913) and Leopold (1939-49), who used the principle of separation by speaker, commonly referred to as "One Person One Language." In their families, each parent regularly used only one language to the child, though the parents varied in what they spoke to each other and to others.

In the one person-one language arrangement, how much stimulation in each language the child gets depends both on time with each parent, and the parent's interactional activities. In a study of German-English bilingualism in Australia,

Döpke (1992) found that fathers and mothers might have such different interactional styles that a child learned one language faster than the other in dichotomy-person families. In the Welsh study by Lyon (1996) of mixed families, the children learned the Welsh of the at-home mothers.

Households, of course, can include more than parents. Monolingual grandparents in immigrant households can be crucial in retention of the family language; Juan-Garau and Pérez-Vidal (2001) reported that a visit by Andreu to his English grandparents reduced his level of Catalan. Household workers and visitors may use another language in the presence of the child. Older siblings bring home the school language, which often turns out to be the language preferred between siblings. For this reason, the oldest child usually has the best knowledge of the family language.

Even when the mothers were at home with the child during the day, Döpke (1992) found that the fathers who used frequent *instructional strategies* had children who learned language faster than they learned the mother's variety. Thus one can view the one person/one language family as a natural experiment if the parents differ in conversational style with the child. Goodz (1994) found that conversations matter more for acquisition than precise language separation.

In Döpke's (1992) study of Australian German-English families there were stay-at-home full time mothers; one parent spoke German and the other English. What turned out to be important in these families was conversational style. Only two children became bilingual and used German spontaneously; only in these families did the German-speaking parent use instructional strategies more than the English-speaking parent. In some families, the English-speaking parent used them more, and in one, neither; in that family the child was the slowest language learner. The instructional conversational exchanges of the German-speaking parent were crucial to the child's bilingualism, since there was no community support.

In each of these families, the frequency of strategies changed with the level of the child's knowledge. Döpke (1992:252) commented regarding Australian German-English families that "Only those children who acquired an *active command* of German were met with high-constraint insisting strategies such as unspecified **clarification requests** and **requests for translation**." Juan-Garau and Pérez-Vidal and Lanza noticed that the parent who spoke the minority language strove the hardest, by using clarification requests, because the threat of language loss is greatest. The parent speaking the community language might even accept code-switches.

### *Contrasts of setting*

As we have seen, the most effective strategy for speakers of minority languages is to learn Language A from parents and Language B outside the home in a different

setting. A child who is very shy and does not readily either initiate interaction or join the activities of other children has less opportunity to learn a second language from them conversationally. Also, some activities are more verbal, and result in a higher variety of language than others. Thus whether the child learns at home, at school or on the playground, an active conversational partner may be a requirement, or a kind of activity that pressures both to participate.

## Pragmatic skills

### *Learning conversational discourse*

Conversational skills develop significantly between two and five, though even two-year-olds know how to greet, to request, to attract attention, establish a topic, respond to check questions, and state what is wanted. Child second-language learners use calls or attention-getters within the first month, for instance address terms, or “Lookit.” Though monolingual children might not initiate repair requests until near the end of the third year, five-year-old second-language learners use them in the first month of a new language, e.g. the learner might say “Huh? What? What you say?” Questions and replies also occur within the first month for five-year-old second language learners in Wong Fillmore’s study.

In first language research, replies to *where*, *who*, and *what* questions precede replies to *when*, *how*, and *why* questions. But there is no such difference in CSLA comprehension, since the children already understand these concepts (Lightbrown, 1978).

Because conceptual development may be more advanced than lexical resources, we find, in the first months, surprising work-arounds, like “Hey what a your name?” “What time you my house,” and conversation facilitators like “What happened?” and “Whatsa matter?” A seven-year-old Anglophone, after two months speaking French, asked a child’s name this way: “C’est quoi, le nom à toi?” [it’s *what, the name to you?*] (Ervin-Tripp, 1981:42).

### *Speech acts*

The development of speech acts in first language learners between two and two and a half is rapid. These acts also occur when the children use another language, so what we find is the same acts as in monolinguals, but perhaps awkward expressions for them because of primitive lexical repertoire.

The first learning from peers we see is at the act exchange level.

Drury (2007) reveals through the transcripts of Pahari speaking children in England the order in which English speech acts appeared in the CSLA transcripts. Earliest to appear were asserting possession, practicing and rehearsing. Some children asked questions and engaged in language play and some did not.

### *Meaningful code-switching*

Situational switching shows what children notice: people, place, activity, and genre. Person switching is robust very early. Ronjat's child was indignant if the wrong language was used in his one-person-one-language family. Lanvers (2001) found self-corrective switching to parents by 1;7 and 1;10. Eventually situational person switching can even generalize to strangers similar in appearance (Fantini, 1985; McClure, 1981).

Surprisingly early, children learn which partners are monolingual. Genesee et al., (1996) found that bilingual Canadians, at a very early stage (MLU in French 1.37 and in English 1.50) could even use the appropriate language with unfamiliar conversational partners who pretended to be monolingual. Comeau et al. (2003) found that very young Canadian bilinguals were sensitive even to the language-mixing rate of bilingual conversational partners.

Kwan-Terry (1992) identified an activity-based situational contrast in her son. He insisted that his mother's narratives should always occur in Cantonese, but in fantasy play the animals and play characters were anglophones.

In families with a dichotomy-person rule, mixing can simply at first be due to lexical gaps. But other meanings accrue to switching. Fantini (1985) says his sons used the wrong language deliberately to amuse or startle relatives. Lanvers (2001) found that even before age two, German-American code-switching was used for crutching, appeal, and emphasis. In bilingual communities, children learn the meanings of code-switching early (Reyes & Ervin-Tripp, 2010).

Role play is a rich source of data on the social meaning of languages to children. During role play children may use the dominant societal language for negotiating play and making commentary, but the other language for enacting roles. Paugh (2005) gives us an elaborate analysis of how bilingual children in Dominica use code-switching between English- and French-based Patwa (patois) to distinguish roles. Kwan-Terry (1992), Halmari and Smith (1994), and Paugh illustrate these contrasts. In addition in everyday interaction with peers, Dominican children use Patwa for "directing, evaluating, and criticizing one another's actions, speech, and demeanor" (Paugh, 2005, p. 67):

## (8) English/Patwa child switching.

Reiston and Sherona (both 3 years) are playing in a small field of bananas maintained by Sherona's father near her home. At one point Reiston stops and bends over, pulling on weeds as if clearing the field:

Reis: (looking down, pulling weeds) *Mwen ka twaway*. [I'm working]

Sher: (watching her old sister) A jumby will go with you though. An evil spirit will take you (i.e., if she continues walking near the edge of the banana field)

Reis: (looking up, concerned) Shero let's go.

Sher: (looks back to the banana field and begins stomping down the hill again, speaking fast in a gruff voice))

*Mwen ka alé. Mwen ka alé.* [[I'm going. I'm going.]

Reis: (following Sherona, stomping with arms flailing)

Paugh (2005:73)

But code-switching has other uses in peer talk. In peer groups, Harrison & Piette (1980) reported switches to establish group membership and influence peers. Those who couldn't switch were excluded, so switching was an important ability in group practice.

*Speech event structure*

Many bilingual children interact with native speakers in nursery school or in playgrounds where the children's activity context affects learning. A good example of a speech event is the telephone conversation, in which there are normal exchange segments such as greetings, entry and reply, core message, and farewell. The following is a play phone example by CSLA Spanish speaker Rosa after 7 months in English:

## (9) Telephone sequence

Rosa: Hello

E: What's you doin?

Rosa: Fine.

E: My mommy told me to go to school.

Rosa: Me too.

E: Ok bye. I'll call you back tomorrow.

Rosa: OK Bye.

Ervin-Tripp 1986: 340

What we see here is that Rosa has already heard some standard telephone replies. She uses the reply to a telephone question "Fine." "Me too" has many uses, and she can use repetitions like "OK Bye" to make a plausible conversation.

In another sequence, there was a correction by the native speaker child about mistakes in telephone procedures by Nora, a Spanish speaker:

(10) Correcting speech event

- Nora: Hello. Come to my house, please.  
E: Who are you?  
Nora: Nora.  
E: Nora, you've got to say "What are you doing?"  
Nora: What are you doing?  
E: Making cookies. What are you doing?  
Nora: Making cookies, too.  
E: Ok, bye.  
Nora: Bye.

Wong Fillmore (1976)

Language use varies considerably in different child activities, a contrast which could be used deliberately to teach second languages. Talk in activities may include essential, social-expressive, dispute talk and secondary registers as in sports-caster talk or roles in role-play.

The activity structure factors that seem to facilitate language learning are redundancy with action that makes situated formulae meaningful, availability of simple, repetitive, and salient models to imitate, predictable verbal routines, and sufficient abstract knowledge of the activity structure to make contextual inference about meanings possible. In some activities, successful participation is necessary, so partners are facilitative. A cooperative attitude by the partner leads to willingness to gesture, interpret the learner's tries, probe for appropriate moves, or give language instruction in the form of corrective modeling (Ervin-Tripp, 1986: 455).

Examples of activity types for child second language learners included detailed analysis of phone conversations, arguments, soccer games, board and card games, and role-playing of classrooms.

## Summary

Interaction is required for most language learning. Learners acquire linguistic forms in adjacency pair discourse formats, social routines, speech act routines, and recognizable speech events. Children who grow up in families where both parents speak an immigrant or minority language have the strongest chance to become active bilinguals. In cases where each parent speaks a different language to the child, the frequency and quality of the parent's instructional interaction is crucial to acquisition and maintenance.

Sociability and knowledge of social routines, or effective classroom structure can make it possible for learners to enter into interaction outside the home and become child second language learners, though there is wide variation in rates of acquisition. Whether at home, school or in the playground, learners need cooperative adult or peer partners who will help to clarify, remedy misunderstandings, instruct and elicit appropriate language use. Access to such conversational partners from both communities is required for active bilingualism to be developed and maintained. In addition to teaching phonology, syntax and lexicon, these contexts provide children with skill in the organization of contextually appropriate speech acts, speech events, and meaningful code-switching.

## References

- Camaioni, L., & Longobardi, E. (1994). A longitudinal examination of the relationships between input and child language acquisition. First Lisbon Meeting on Child Language with Special Reference to Romance Languages. Lisbon, Portugal.
- Clark, E.V. (1999). Acquisition in the course of conversation. *Studies in the Linguistic Sciences (Forum Lectures from the 1999 Linguistic Institute)*, 29(2), 1–18.
- Clark, E.V., & Amaral, P.M. (2010). Children build on pragmatic information in language acquisition. *Language & Linguistic Compass*, 4, 445–457. DOI: 10.1111/j.1749-818X.2010.00214.x
- Comeau, L., Genesee F., & Lapaquette, L. (2003). The modeling hypothesis and child bilingual code-mixing. *International Journal of Bilingualism*, 7, 113–126. DOI: 10.1177/13670069030070020101
- Crystal, D. (1976). *Child language: Learning and linguistics*. London: Arnold.
- De Houwer, A. (2009). *Bilingual first language acquisition*. Bristol: Multilingual Matters.
- Döpke, S. (1992). *One parent-one language: An interactional approach*. Amsterdam: John Benjamins. DOI: 10.1075/sibil.3
- Drury, R. (2007). *Young bilingual learners at home and school: Researching multilingual voices*. Stoke on Trent: Trentham Books.
- Eilers, R.E., Pearson, B.Z., & Cobo-Lewis, A.B. (2006). Social factors in bilingual development: The Miami experience. In P. McCardle & E. Hoff (Eds.), *Childhood bilingualism: Research on infancy through school age* (pp. 68–90). Clevedon: Multilingual Matters.
- Ervin-Tripp, S.M. (1974). Is second language learning like the first? *TESOL Quarterly*, 8, 111–127. DOI: 10.2307/3585535
- Ervin-Tripp, S.M. (1981). Social process in first and second language learning. In Winitz, H. (Ed.), *Native language and foreign language acquisition (Annals of the NY Academy of Science)*, Vol. 379, pp. 33–47). New York, NY: New York Academy of Science.
- Ervin-Tripp, S.M. (1986). Activity structure as scaffolding for children's second language learning. In J. Cook-Gumperz, W.A. Corsaro & J. Streeck (Eds.), *Children's Worlds and Children's Language* (pp. 327–358). Berlin: Mouton de Gruyter.
- Ervin-Tripp, S.M. (1995). Impact of the interactional setting on the acquisition of syntax. *Acquisition et Interaction en Langue Étrangère (AILE)*, 4, 53–88.

- Fantini, A. (1985). Language acquisition of a bilingual child: A socio-linguistic perspective. Clevedon: Multilingual Matters.
- Fernald, A. (2006). When infants hear two languages: Interpreting research on early speech perception by bilingual children. In P. McCardle & E. Hoff (Eds.), *Childhood bilingualism: Research on infancy through school age* (pp. 19–29). Clevedon: Multilingual Matters.
- Gathercole, S., & Thomas, E.M. (2009). Bilingual first language development: Dominant language takeover, threatened minority language takeup. *Bilingualism: Language and Cognition*, 12, 213–237. DOI: 10.1017/S1366728909004015
- Genesee, F., Boivin, I., & Nicoladis, E. (1996). Talking with strangers: A study of bilingual children's communicative competence. *Applied Psycholinguistics*, 17, 427–442. DOI: 10.1017/S0142716400008183
- Goodz, N. (1994). Interactions between parents and children in bilingual families. In F. Genesee (Ed.), *Educating second language children* (pp. 61–81). Cambridge: CUP.
- Halmari, H., & Smith, W. (1994). Code-switching and register shift: Evidence from Finnish-English child bilingual conversation. *Journal of Pragmatics*, 21, 427–445. DOI: 10.1016/0378-2166(94)90013-2
- Harrison, G., & Piette, A.B. (1980). Young bilingual children's language selection. *Journal of Multilingual and Multicultural Development*, 3, 217–230. DOI: 10.1080/01434632.1980.9994016
- Hart, B. & Risley, T.R. (1992). American parenting of language-learning children: Persisting differences in family-child interactions observed in natural home environments. *Developmental Psychology*, 26, 1096–1105. DOI: 10.1037/0012-1649.28.6.1096
- Hyltenstam, K., & Abrahamsson, N. (2003). Maturational constraints in SLA. In C.J. Doughty & M.H. Long (Eds.), *The handbook of second language acquisition* (pp. 538–588). Oxford: Blackwell. DOI: 10.1002/9780470756492.ch17
- Hymes, Dell. (1974). *Foundations of sociolinguistics: An ethnographic approach*. Philadelphia, PA: University of Pennsylvania Press.
- Juan-Garau, M., & Pérez-Vidal, C. (2001). Mixing and pragmatic parental strategies in early bilingual acquisition. *Journal of Child Language*, 28, 59–86. DOI: 10.1017/S030500090004591
- King, K.A., Fogle, L., & Logan-Terry, A. (2008). Family language policy. *Language and Linguistics Compass*, 2, 907–922. DOI: 10.1111/j.1749-818X.2008.00076.x
- Kwan-Terry, A. (1992). Code-switching and code-mixing: The case of a child learning English and Chinese simultaneously. *Journal of Multilingual and Multicultural Development*, 13, 243–259. DOI: 10.1080/01434632.1992.9994494
- Lanvers, U. (2001). Language alternation in infant bilinguals: A developmental approach to codeswitching. *International Journal of Bilingualism*, 5, 437–464. DOI: 10.1177/13670069010050040301
- Lanza, E. (1997). *Language mixing in infant bilingualism: A sociolinguistic perspective*. Oxford: Clarendon Press.
- Leopold, W.F. (1939–1949). *Speech development of a bilingual child: A linguist's record*, Vols 1–4. Evanston, IL: Northwestern University Press/New York, NY: AMS Press, 1970.
- Lightbown, P.M. (1978). Question form and question function in the speech of young French L2 learners. In M. Paradis (Ed.), *Aspects of bilingualism* (pp. 25–42). Columbia, SC: Hornbeam Press.
- Lyon, J. (1996). *Becoming bilingual: Language acquisition in a bilingual community*. Clevedon: Multilingual Matters.

- McClure, E. (1981). Formal and functional aspects of the code switched discourse of bilingual children. In R.P. Duran (Ed.), *Latino language and communicative behavior* (pp. 69–94). Norwood, NJ: Ablex.
- Meisel, J. (2004). The bilingual child. In T.K. Bhatia & W.C. Ritchie (Eds.), *The handbook of bilingualism*: Blackwell handbooks in linguistics (pp. 91–113). Oxford: Blackwell.
- Paugh, A.L. (2005). Bilingual play: Children's code-switching, role play, and agency in Dominica, West Indies. *Language in Society*, 34, 63–86. DOI: 10.1017/S00474045050037
- Pearson, B.Z., Fernández, S.C., & Oller, D.K. (1993). Lexical development in bilingual infants and toddlers: Comparison to monolingual norms. *Language Learning*, 43, 93–120. DOI: 10.1111/j.1467-1770.1993.tb00174.x
- Reyes, I., & Ervin-Tripp, S.M. (2010). Language choice and competence: Codeswitching and issues of social identity in young bilingual children. In M. Shatz & L. Wilkinson (Eds.), *The education of English language learners: Research to practice* (pp. 67–84). New York, NY: Guilford Press.
- Ronjat, J. (1913). *Le développement du langage observé chez un enfant bilingue [The development of language in a bilingual child]*. Paris: Librairie Ancienne H. Champion.
- Schiffrin, D. (1987). *Discourse Markers*. Cambridge: CUP. DOI: 10.1017/CBO9780511611841
- Schmidt-Mackey, I. (1977). Language strategies of the bilingual family. In W. Mackey & T. Andersson (Eds.), *Bilingualism in early childhood* (pp. 132–146). Rowley, MA: Newbury House.
- Walker, D., Greenwood, C., Hart, B., & Carta, J. (1994). Prediction of school outcomes based on early language production and socioeconomic factors. *Child Development*, 65, 606–621. DOI: 10.2307/1131404
- Wells, G. (1985). *Language Development in the Preschool Years*. Cambridge: CUP.
- Wong Fillmore, L. (1976). *The second time around: Cognitive and social strategies in second language acquisition*. Unpublished doctoral dissertation. Stanford University, CA.
- Wong Fillmore, L. (1991). Second-language learning in children: A model of language learning in a social context. In E. Bialystok (Ed.), *Language processing in bilingual children* (pp. 49–69). Cambridge: CUP. DOI: 10.1017/CBO9780511620652.005



# **Social environments shape children's language experiences, strengthening language processing and building vocabulary**

Adriana Weisleder and Anne Fernald

New York University School of Medicine and Stanford University

How does language experience influence the development of language skills known to be critical for academic success? In this chapter, we build on a long history of research examining sources of variability in children's lexical development, and offer a new perspective that focuses on the development of efficiency in real-time language processing. Examining origins of individual differences in language proficiency, we review new research showing that the amount and quality of child-directed speech in infancy contributes to the development of language processing skills, which in turn facilitate vocabulary growth. These findings reveal that early language experience can have cascading effects for later learning and school success.

## **Social environments shape children's language experiences, strengthening language processing and building vocabulary**

The claim that talking to children is somehow important for language development seems hard to disagree with. Yet questions about whether, how, and why particular aspects of children's success in language learning might depend on particular aspects of their early experience with language are quite controversial. Over the course of a long and influential career, Eve Clark approached these questions with original insights that often met with strong headwinds. She began her work at a time when the dominant theoretical perspectives focused on the quest for universal patterns in early language growth, which were assumed to be propelled by innate grammatical knowledge and highly constrained strategies for lexical learning. Clark (1979, 1987, 1993) challenged some of these assumptions by showing how infants learn new words through social and conversational interactions with attentive caregivers who provide structure for learning, a view that emphasized an important role for experience. This theme also resonates in more recent work, where

Clark argues that identifying and elucidating sources of the striking differences observed among children in their trajectories of language learning – which vary substantially with differences in early experience – is as important as searching for similarities in patterns of development (Clark, 2003; Arnon & Clark, 2011).

Here we present research exploring how children's early experience with language from caregivers influences the development of language skills in infancy and later childhood. The first section reviews research on how SES differences in speech to children are linked to vocabulary learning. The next two sections describe the early development of language processing efficiency, why it matters, and how the development of this critical language skill varies in different SES groups. The fourth section examines the origins of differences in early language proficiency, with new data showing remarkable variability among low-SES families in the amount of verbal engagement with infants. We also provide evidence that richer language experience strengthens language processing skill, which in turn facilitates language growth. In the final section we discuss implications of the finding that early language experience can have cascading effects for later learning and school success. By showing how the amount and quality of child-directed speech in infancy contribute both to language processing skill and to lexical development over time, we affirm the value of the intuitive word-teaching strategies Clark has illuminated in her studies of higher-SES mothers (Clark, 2010; Clark & Estigarribia, 2011). However, our results also reveal that such supportive behaviors are used much less frequently by mothers from lower-SES populations, and that these differences have potentially serious consequences for their children's language learning.

## **1. Sources of variability in children's language learning**

Although all typically-developing children in normal environments learn to talk, there is substantial variability among them in rates of language learning and in the linguistic proficiency they eventually achieve. To some extent, these differences are part of the normal variation that exists in all human abilities. But differences in trajectories of language learning are revealing about where children come from and what their futures will look like, and thus are of considerable social relevance. Individual differences in language development mirror divisions in our societies, with children from disadvantaged families learning language more slowly, on average, than those from families higher in socioeconomic status (SES) (Farkas & Beron, 2004; Hart & Risley, 1995; Huttenlocher, Waterfall, Vasilyeva, Vevea, & Hedges, 2010; Rowe, Raudenbush, & Goldin-Meadow, 2012). Moreover, these differences have far-reaching consequences for children's academic and later life

(Durham, Farkas, Hammer, Tomblin, & Catts, 2007; Senechal, 2006; Tramontana, Hooper, & Selzer, 1988; Walker, Greenwood, Hart, & Carta, 1994).

What is it about socioeconomic background that leads to disparities among children in their language skills? Many experiential factors associated with living in poverty could contribute to variability in language learning. Aspects of the physical environment, such as sanitation, noise level, and exposure to toxins and dangerous conditions, differ dramatically in lower- and higher-SES families (Evans, 2004). There are also SES differences in access to adequate nutrition and health care, and in the availability of social and psychological support (Bradley & Corwyn, 2002; Engle & Black, 2008; Huston, Mcloyd, & Garcia Coll, 1994). These environmental factors are all implicated in children's physical, cognitive and social development (Bradley, Corwyn, Burchinal, McAdoo, & Coll, 2001; Brooks-Gunn & Duncan, 1997; Evans, 2006), and thus are likely to have consequences for children's language outcomes.

The quality of parent-child interaction is also related to SES. Parents under greater stress tend to be less nurturing and to respond less sensitively to their children (Conger, McCarty, Yang, Lahey, & Kropp, 1984; McLoyd, 1990; Mesman, van IJzendoorn, & Bakermans-Kranenburg, 2012). In addition, parents who are more affluent and better educated tend to invest more time and resources in their children's intellectual development (Phillips, 2011; Ramey & Ramey, 2009), buying more books and educational toys for their children, and engaging with them more often in cognitively enhancing activities (Bradley, Corwyn, McAdoo, & Coll, 2001; Raikes et al., 2006).

One of the most significant differences in the early experiences of lower- and higher-SES children is the amount and nature of the language they hear. Hart and Risley (1995) estimated that by 36 months, children from high-SES families had heard 30 million more words directed to them than those growing up in poverty, a stunning difference that predicted important long-term outcomes (Walker et al., 1994). In interactions with their children, higher SES mothers gesture more, use more diverse vocabulary and more complex syntax, ask more questions and use fewer directives, and respond more contingently to their infants' vocalizations than do lower SES mothers (Hart & Risley, 1995; Hoff-Ginsberg, 1991, 1998; Huttenlocher, Vasilyeva, Waterfall, Vevea, & Hedges, 2007). These differences in the quantity and quality of language input have been shown to account, at least in part, for disparities in children's lexical and grammatical development (Hoff, 2003; Huttenlocher, Waterfall, Vasilyeva, Vevea, & Hedges, 2010; Rowe & Goldin-Meadow, 2009).

How is it that child-directed speech promotes children's lexical development? One explanation proposed for the association between exposure to more child-directed speech and faster vocabulary growth has been that more language from caregivers provides children with more models to learn from as they begin to build

a lexicon. Children who hear more speech from their caregivers are exposed to more different words and to more instances of those words in a variety of contexts. Thus, they have more opportunities to learn new word forms and more information to figure out the meanings of these words, leading to growth in vocabulary (e.g., Hoff & Naigles, 2002). Another explanation for how child-directed speech promotes lexical development emphasizes socio-pragmatic aspects of conversation, such as joint attention and maternal responsiveness (e.g., Akhtar & Tomasello, 2000). According to this view, children learn language in the context of mutually understood social interactions (Clark & Clark, 1977), or joint attentional formats (Bruner, 1983), during which they can leverage inferences about speakers' communicative intentions to learn word meanings. Thus, children will exhibit faster vocabulary growth when they have more opportunities to interact with an attentive caregiver and engage in more episodes of joint attention.

In this chapter, we argue that language experience contributes to lexical development in another way as well. By hearing words used repeatedly in a variety of contexts, children have more opportunities to practice processing words they already know, and for interpreting the meanings of words in relation to the linguistic and extra-linguistic context. As a result, infants with more exposure to child-directed speech are faster and more accurate in interpreting familiar words in real time. Recent evidence from our lab supports the idea that practice with language strengthens children's processing skill, which then helps to drive vocabulary growth (Weisleder & Fernald, 2013)

## **2. What is processing efficiency? And why does it matter?**

Next we describe new ways to characterize developmental gains in verbal ability by assessing infants' fluency in interpreting spoken language, and show that variability in language-processing efficiency predicts both concurrent and later language and cognitive outcomes. Many studies of comprehension in young children have relied on *offline* measures, responses made after the offset of the speech stimulus. While offline measures can provide evidence that a child responds systematically in a way that indicates understanding, they do not tap into the dynamic nature of language understanding and thus reveal little about the child's developing efficiency in interpreting familiar words in fluent speech. In contrast, most studies of adult speech processing rely on *online* measures that monitor the time course of the listener's response to spoken words in relation to key points in the speech signal. Because comprehension happens quickly and automatically, it is revealing to study the listener's interpretation *during* processing of the speech signal rather than waiting until processing is complete.

Here we describe recent research using the “looking-while-listening” (LWL) paradigm (Fernald, Zangl, Portillo, & Marchman, 2008) to monitor the time course of comprehension by very young language learners. In the LWL procedure, infants look at pairs of pictures while listening to speech naming one of the pictures, and their gaze patterns are video-recorded as the sentence unfolds in time. Using this paradigm, we have shown that speed and efficiency in infants’ recognition of familiar words increase substantially over the second year for English- and Spanish-learning children (Fernald, Pinto, Swingley, Weinberg, & McRoberts, 1998; Hurtado, Marchman, & Fernald, 2007), that young children make use of linguistic information at multiple levels in real-time processing (Lew-Williams & Fernald, 2007; Swingley, Pinto, & Fernald, 1999), and that individual differences in early processing efficiency are related to lexical and grammatical development, both concurrently and at later ages (Fernald & Marchman, 2011; Fernald, Perfors, & Marchman, 2006; Hurtado, Marchman & Fernald, 2008).

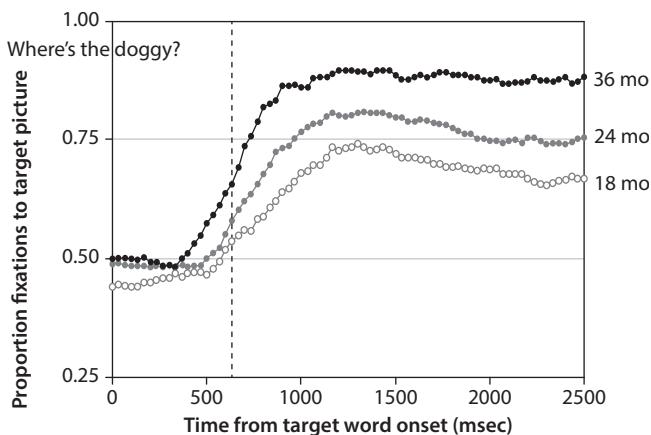
### *What is fluency in understanding?*

To follow a conversation, adults process 10-15 phonemes per second as they continuously integrate acoustic information with linguistic and conceptual knowledge. Fluent understanding of a speaker’s meaning requires the ability to listen *predictively*, anticipating what is coming next in the speech stream by integrating different sources of linguistic knowledge with nonlinguistic information from the context in which the words are spoken. Many studies show that adult listeners exploit linguistic and nonlinguistic knowledge on multiple levels in anticipating upcoming words, and that these predictions are made continuously and instantaneously as the speech signal unfolds (Allopenna, Magnuson, & Tanenhaus, 1998; Altmann & Kamide, 1999).

In cross-sectional and longitudinal studies using the LWL task, we have shown that infants make dramatic gains in receptive language skill over the second year of life, by increasing the speed and accuracy with which they identify familiar words in spoken language and match them with the appropriate referent (Fernald et al., 2006, 1998, 2008).

Figure 1 shows one representation of the data from a study of children at 3 different age points; 18, 24, and 36 months (Zangl & Fernald, 2007). The x-axis shows time in ms from the onset of the target noun; the y-axis shows the mean proportion of fixations to the target picture. As the figure shows, older children shifted to the target picture sooner – indicating faster processing of the familiar words – and they reached a higher asymptote – indicating higher accuracy.

Studies have also shown that young children, like adults, are already able to make use of potentially informative sources of contextual information, processing



**Figure 1.** Mean proportion of trials on which children in three age groups are looking at the target picture at each 33-msec interval as the stimulus sentence unfolds. The dashed vertical line represents target noun offset; error bars represent SEs over participants.

speech incrementally from moment to moment (Henderson, Weighall, Brown, & Gaskell, 2013; Swingley et al., 1999). One example illustrates incremental processing at the lexical level, when the listener identifies a word based on partial phonetic information without waiting to hear the whole word. A child who hears *Where's the dog?* in the presence of a dog and a doll is confronted with a temporary ambiguity, since *dog* and *doll* overlap phonetically and thus are indistinguishable for the first 300 ms or so. In this situation, 24-month-olds delayed their response by about 300 ms until disambiguating information became available (Swingley et al., 1999), parallel to findings from studies on adult speech processing (Allopenna et al., 1998).

A second example illustrates incremental processing at a morphosyntactic level. In languages such as Spanish, all nouns have grammatical gender, with obligatory gender-marking on preceding articles. To explore whether children could exploit grammatical gender cues in real-time sentence interpretation, Lew-Williams and Fernald (2007) tested Spanish-learning children in the LWL procedure. Children saw pairs of pictures with names of either the same (e.g., *la pelota*, 'ball'; *la galleta*, 'cookie') or different grammatical gender (e.g., *la pelota*, *el zapato*, 'shoe'), as they heard sentences referring to one of the pictures (e.g., *Encuentra la pelota*, 'Find the ball'). On same-gender trials, the article could not be used to identify the referent before the noun was spoken; on different-gender trials, the gender-marked article was potentially useful in predicting the subsequent noun. This study showed that children were reliably faster to identify the referent on different-gender trials, as were native Spanish-speaking adults tested in the same

procedure. Thus, young children learning Spanish as their first language already demonstrated a processing advantage that is typical of adult native speakers but not of second-language learners (Guillelmon & Grosjean, 2001; Grüter, Lew-Williams & Fernald, 2012). This ability to exploit morphosyntactic information in incremental processing reveals another dimension of children's early emerging fluency in understanding.

### *Stability and predictive validity of online processing measures*

Another central goal in our research is to characterize the causes and consequences of variation in language proficiency among young children. We know that at every age, there is substantial variation among children in offline measures of vocabulary and grammar (Bates, Dale, & Thal, 1995; Fenson et al., 1994). Research using online processing measures of language understanding can also address important questions about differences among children: Is speed of lexical processing a stable measure across age for individual children? And if so, how do individual differences in early speech processing efficiency relate to later language growth, as assessed by standardized measures of lexical and grammatical knowledge?

We first addressed these questions in a longitudinal study of 59 English-learning infants tested four times in the LWL procedure between 15 and 25 months (Fernald et al., 2006). Children's efficiency in identifying familiar words increased significantly over this period, and measures of early processing skill were moderately stable from one age to the next. That is, children who responded more quickly on average in identifying familiar words at earlier ages also responded relatively more quickly at later ages. Parental reports of vocabulary and grammar on the MacArthur-Bates Communicative Development Inventory (CDI; Fenson et al., 2006) were also gathered across the second year, enabling us to explore the relation of online measures of speech processing skill to more traditional measures of linguistic development. Speed and accuracy in speech processing at 25 months were robustly related to lexical and grammatical development across a range of measures from 12 to 25 months, and those children who had faster RTs at 25 months also showed more accelerated vocabulary growth across the second year. We have recently replicated this finding in a prospective longitudinal study, showing that early processing efficiency predicts vocabulary growth in a larger sample of typically-developing children and in children who showed delays in the onset of productive vocabulary (Fernald & Marchman, 2011).

Given the stability and short-term predictive validity of these online measures in the infancy period, the next question was to what extent individual differences in early processing efficiency predict *long-term* language and cognitive outcomes. In a follow-up study, 30 children from the original Fernald et al. (2006) longitudinal

sample were tested at 8 years of age on two standardized assessments of cognitive and language skills (Marchman & Fernald, 2008). Multiple regression analyses were used to evaluate the predictive validity of processing speed and expressive vocabulary in infancy, in relation to school-age outcomes. Vocabulary size at 25 months was correlated with later cognitive and language skills, but knowing mean RT in addition to vocabulary *doubled* the predictive power, accounting for 58% of the variance in working memory at 8 years. This longitudinal study was the first to reveal the long-term predictive validity of early measures of real-time processing efficiency, showing that individual differences in fluency of understanding at two years predict children's cognitive and language outcomes in later childhood.

### 3. SES-differences in language processing skill

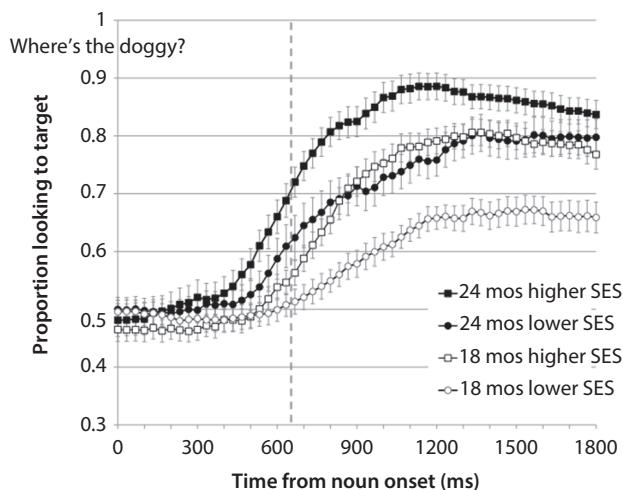
In the research described above, language-processing skill was studied in English-learning children from families high in SES (Fernald et al., 2006; Marchman & Fernald, 2008) as well as in Spanish-learning children from families low in SES (Hurtado et al., 2007; Hurtado, et al., 2008). These studies revealed age-related increases in processing efficiency in both populations, with comparable links between processing efficiency and vocabulary. However, this research did not allow us to determine whether variation in language-processing skill was related to socio-economic differences, given that SES was confounded with both language background and ethnicity in the families that participated in these studies.

To examine the influence of SES on the development of language-processing skill, we needed to recruit English-speaking families from a much broader demographic range than is typical of most psychological research (Arnett, 2008; Fernald, 2010). With the goal of extending our research beyond the convenience sample of high-SES families available as participants at our university-campus lab (Site 1), we outfitted a mobile lab to enable us to conduct research in areas where it is possible to recruit equivalent numbers of lower- and middle-SES English-speaking families. Site 2 is in an urban area comparable in population size to the area in which Site 1 is located. However, these two regions differ substantially in median family income, cost-of-living, and percentage of children living in poverty, allowing us to include a much more diverse sample of English-learning children in our research.

In the first study to examine SES differences in the development of processing efficiency and vocabulary, Fernald, Marchman, and Weisleder (2013) recruited 48 children from monolingual English-speaking families in which maternal education ranged from less than a high school degree to post-BA training. Based on the

Hollingshead Four Factor Index of Socioeconomic Status (HI; Hollingshead, 1975), families were divided into lower- ( $n = 23$ ) and higher-SES ( $n = 25$ ) groups. Of the children from families in the higher-SES group, 19 were recruited at Site 1 and six at Site 2; of those in the lower-SES group, one was recruited at Site 1 and 22 at Site 2.

In both groups we found common developmental patterns, with increases in vocabulary and in language-processing skill from 18 to 24 months. Older children had larger vocabularies, and were more likely than younger children to interpret incoming speech incrementally, fixating the target picture as soon as they had enough information to identify the referent. In both groups we also found reliable links between skill in early language processing and vocabulary development, replicating previous studies (Fernald et al., 2006; Fernald & Marchman, 2011). But the differences between higher- and lower-SES children in language-processing skill and vocabulary were striking. By the age of 18 months, children from higher-SES families were faster and more accurate in identifying referents of familiar words than were children from lower-SES families. As shown in Figure 2, mean accuracy for lower-SES children increased from .59 to .69 between the ages of 18 and 24 months; however, mean accuracy for the higher-SES children was already .69 at 18 months, increasing to .77 by 24 months. Similar patterns were observed for reaction time and vocabulary. Thus, these findings revealed the equivalent of a 6-month gap in processing efficiency between higher- and lower-SES children, which was already well established by the age of 24 months.



**Figure 2.** Mean proportion of looking to the target as a function of time in ms from noun onset for Lower- and Higher-SES children. Open squares/circles represent the time course of correct looking at 18 months; filled squares/circles represent the time course of looking in the same children at 24 months. Error bars represent SE of the mean over participants. Adapted from Fernald et al., 2012.

#### 4. Where do these differences come from?

What explains these disparities in children's language-processing skills? Many studies described earlier have shown that variability in lexical and grammatical development is linked to differences in children's exposure to language from caregivers. But is variation in early language experience also related to individual differences in children's real-time language processing? Two different studies from our lab shed light on this question.

In the first study, Hurtado, Marchman, and Fernald (2008) recorded 27 Latina mothers interacting with their 18-month-old infants during a short play session in the laboratory. Most of the mothers were native Spanish speakers with limited English proficiency who had recently immigrated from Mexico. At both 18 and 24 months, the children were assessed in the LWL procedure, and mothers completed a Spanish-language CDI. Hurtado et al. then examined links between features of maternal talk and children's vocabulary size and speech processing efficiency. Measures of mothers' speech included total number of utterances, word tokens, word types, and mean length of utterance.

Consistent with previous findings with English-learning children, the quantity and quality of maternal speech were associated with children's vocabulary outcomes. Infants who heard more, more varied, and more complex speech from their mother at 18 months had larger vocabularies at 24 months, even when controlling for early vocabulary differences. Vocabulary size was also related to speech-processing efficiency, as in previous studies with English-learning children (Fernald et al. 2006): 18-month-old children who were faster to identify familiar words in fluent speech made greater gains in vocabulary from 18 to 24 months. But the important new finding from this study was that those children whose mothers spoke more words and used more complex utterances during the play session at 18 months were significantly faster in online comprehension at 24 months, as compared to those who had heard less maternal talk – even after controlling for differences in mean RT at 18 months. This research provided the first evidence that variability in maternal speech predicts children's efficiency in language processing.

In a more recent study, we extended this research in three important ways (Weisleder & Fernald, 2013): First, rather than relying on short samples of mothers' speech in dyadic interactions with infants, we collected more extensive and representative recordings of infants' language environments during a 10-hour day at home, capturing children's interactions with different family members in diverse settings and activities. Second, we used unobtrusive technology to record families in their natural home environments, minimizing artifacts introduced by the presence of an observer or by parents' reactions to an unfamiliar laboratory

setting. And third, by focusing on a relatively homogeneous group of low-SES, Latino families, we examined variability in the daily language experiences of children from similar socioeconomic and cultural backgrounds.

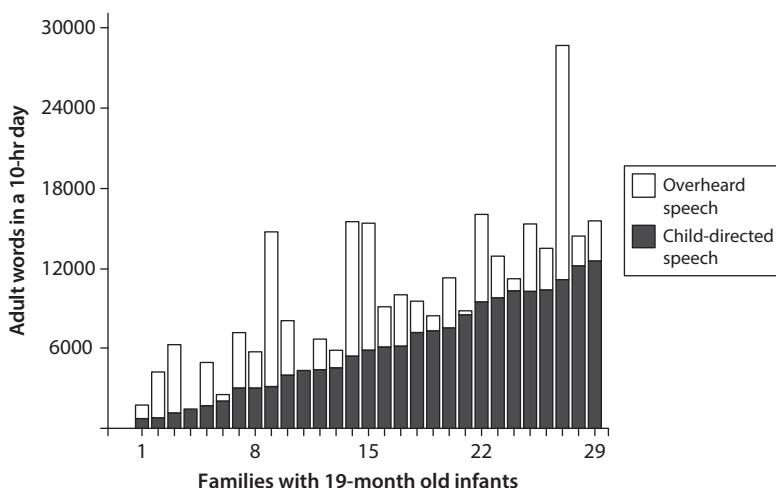
By examining how measures of spontaneous caregiver speech at home relate to measures of both language processing and expressive vocabulary, we could address three sets of questions: (1) How much do children's daily language experiences vary within this population of low-SES Latino families? (2) How is children's exposure to speech related to differences in their vocabulary development? Is vocabulary predicted only by adult speech that is directed to the child, or is speech overheard by the child also linked to vocabulary outcomes? (3) Does early language experience explain differences in children's language-processing skills? And if so, do differences in processing skill help explain the link between language experience and vocabulary development?

Parents in the 29 participating families were all native Spanish speakers. Most had not completed high school, and the majority reported a yearly family income below the federal poverty line. To measure the amount of adult speech their children were routinely exposed to, audio-recordings were made during a typical day at home when the child was 19 months old. These audio-recordings were made using a LENA™ digital language processor (DLP; <<http://www.lenafoundation.org>>), a digital recorder that is placed in the chest pocket of specialized clothing worn by the child and is designed to record the audio environment surrounding the child for up to 16 hours (Ford, Baer, Xu, Yapanel, & Gray, 2009). This allowed us to unobtrusively collect whole-day recordings of children's natural language environments. At 19 and 24 months, children's efficiency in online comprehension was assessed in the LWL procedure and mothers completed a Spanish-language CDI.

The recordings were first processed by LENA (Language ENvironment Analysis) software, which provides reliable estimates of different components of the infant's language environment, including the number of *adult word tokens* identified as potentially accessible to the child (Oller et al., 2010; Xu et al., 2008). Native Spanish-speaking coders then listened to each of the recordings in order to differentiate between speech directed to the target child and speech directed to other adults or children nearby. Each 5-min segment of speech was classified as predominantly "child-directed" or "overheard" based on the content of the speech. When speech was directed to a group of people that included the target child, the segment was classified as "child-directed". However, speech that was clearly directed to a child other than the target child was classified as "overheard". The number of adult word tokens in segments classified as child-directed, divided by the duration of the recording, served as our measure of "child-directed speech". The

number of adult word tokens in segments classified as overheard, divided by the duration of the recording, served as our measure of “overheard speech.”<sup>1</sup>

We found striking variability among families in the total amount of adult speech accessible to children, which ranged from almost 29,000 adult words to fewer than 2,000 words over a 10-hour day, as shown in Figure 3. These differences were even more extreme when we included only child-directed speech. In one family, caregivers spoke more than 12,000 words to the infant, while in another the infant heard only 670 words of child-directed speech over an entire day – an 18-fold difference. This shows that although differences in the quantity and quality of caregiver speech are consistently linked to factors related to SES, there is also considerable variability in children’s language experiences *within* a social class.



**Figure 3.** Variability across 29 families in amount of adult speech during a typical day at home. The height of each bar indicates the total number of adult words spoken near the target child in one family, calculated by averaging the word counts per waking hour and extrapolating to a 10-hr day. The proportion of words that was child-directed speech is indicated in gray, with overheard speech in white. Adapted from Weisleder & Fernald (2013).

1. It is important to keep in mind that language “input” does not necessarily imply language “uptake” by the child. In an audio recording it is fundamentally impossible to know whether a child is actually paying attention to the speaker, regardless of whether the speech is directed to the child or to someone else present. Thus, the category of “child-directed speech” designates all speech that is potentially accessible and directed to the target child, while the category of “overheard speech” designates all speech that is potentially accessible to the target child but not directed to them.

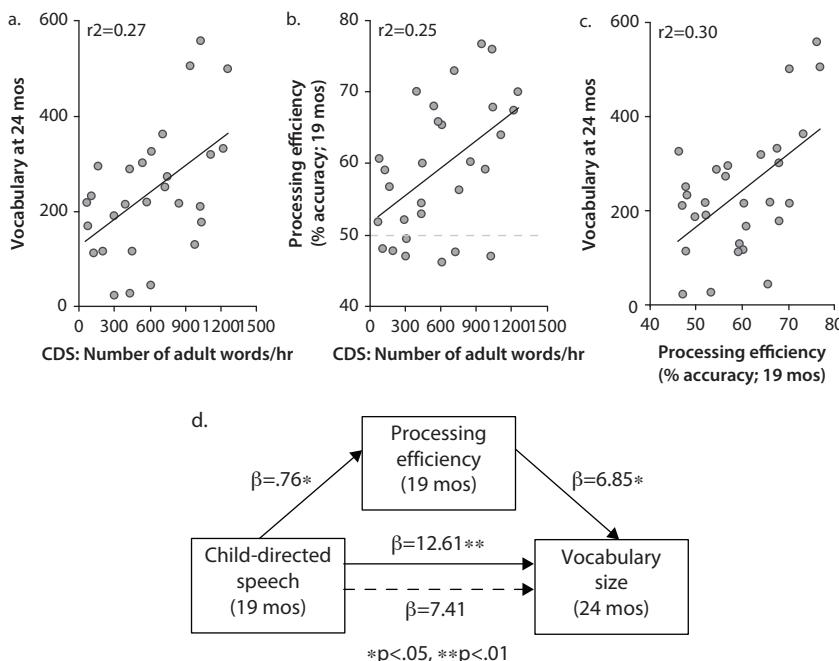
Our next question was whether differences in child-directed speech were related to differences in children's lexical development. Consistent with previous studies (Hoff, 2003; Hurtado et al., 2008), children who heard more child-directed speech at 19 months had larger vocabularies at 24 months, showing that even among families similar in SES and cultural background, differences in amount of caregiver speech are related to children's lexical development. However, differences in exposure to *overheard* speech were not related to vocabulary size, consistent with recent studies of children in middle-class English-speaking families in the U.S. (Arroyo, Levine, & Goldin-Meadow, 2013) and in Yucatec Mayan families (Shneidman & Goldin-Meadow, 2013). These results provide compelling evidence that language spoken directly to the child is more supportive of early lexical development than speech simply overheard by the child.

There are several reasons why child-directed speech might be more effective than overheard speech in fostering children's early lexical development. First, prosodic properties of infant-directed speech (IDS) may play a role in engaging infants' attention, which is critical for learning in natural settings. A number of experiments have shown that very young infants prefer to listen to infant-directed speech as compared to adult-directed speech (Cooper & Aslin, 1990; Fernald, 1985, 1992; Fernald & Kuhl, 1987), and that phonetic information is more clearly specified in IDS (Kuhl et al., 1997). Speech directed to children is also more likely to occur in contexts that are relevant to the child's attentional focus than overheard speech, thereby facilitating word learning. Experimental studies have shown that it is easier for children to learn new labels when the speaker follows in on the infant's focus of attention (Tomasello & Todd, 1983), something that middle-class American parents often do when introducing new words to their children (Clark & Estigarribia, 2011; Clark, 2010). Observational studies have also shown that children who spend more time in joint attention with their mothers have more advanced vocabulary development than children who spend less time in episodes of joint attention (Akhtar, Dunham, & Dunham, 1991; Tomasello & Farrar, 1986). Finally, speech directed to children typically uses simpler syntax, fewer unique word types, and contains more single-word utterances than speech directed to adults (Phillips, 1973; Soderstrom, 2007). This may be helpful because the meaning of specific referents may be clearer to children when presented in simple and familiar syntactic frames (Cameron-Faulkner, Lieven, & Tomasello, 2003). Indeed, in some studies, these kinds of simplifications have been shown to facilitate children's early lexical development (Brent & Siskind, 2001; Furrow, Nelson, & Benedict, 1979; Murray, Johnson, & Peters, 1990).

To address our final set of questions, we used mediation analyses to examine whether processing skill at 19 months helped to explain the link between child-directed speech and 24-month vocabulary (while controlling for maternal

education, recording length, and infant vocalizations at 19 months). The scatter plots in Figure 4 illustrate the first three steps of the mediation analysis: (1) Exposure to child-directed speech at 19 months predicted vocabulary at 24 months, (2) Exposure to child-directed speech also predicted processing efficiency at 19 months, and (3) 19-month processing efficiency predicted 24-month vocabulary, even when controlling for child-directed speech. Finally, as shown at the bottom of Figure 4, the path coefficient between the predictor variable (child-directed speech) and the outcome variable (vocabulary) was significantly reduced when the mediator variable (processing efficiency) was included in the model. This final step confirms that the mediation was significant, showing that 19-month processing efficiency mediated the link between child-directed speech and 24-month vocabulary.

These results suggest that a critical factor in the path from early language experience to later vocabulary knowledge is the influence of language exposure on



**Figure 4.** The scatter plots show zero-order correlations between (a) child-directed speech (CDS) at home and vocabulary size at 24 mos, (b) CDS and processing efficiency at 19 mos, and (c) processing efficiency at 19 mos and vocabulary at 24 mos. The bottom portion of the figure (d) represents the analysis showing that processing efficiency mediates the relation between CDS and 24-mo vocabulary. Adapted from Weisleder & Fernald (2013).

infants' speech-processing skill. Infants who hear more talk have more opportunities to interpret language, and to exercise skills such as segmenting speech and accessing lexical representations that are vital to word learning (Saffran, Newport, & Aslin, 1996; Gershkoff-Stowe, 2002). As a result, infants with more exposure to child-directed speech are faster and more accurate to orient to familiar words in real time, enabling them to learn new words more quickly and facilitating rapid vocabulary growth.

## 5. Conclusions

The studies reviewed here build on a long history of research examining how language experience supports language development. By using gradient experimental measures of efficiency in moment-to-moment processing that are robustly related to a range of long-term language and cognitive outcomes, we have moved a step closer to understanding the mechanisms underlying the association between early language experience and the development of verbal proficiency. One central insight emerging from this research is that early language experience influences language development through multiple and mutually influential pathways. Previous studies have identified features of parent-child interactions that provide important supports for language learning, including semantic, syntactic, and pragmatic cues to word meaning (Clark, 2010; Clark & Estigarribia, 2011; Frank, Tenenbaum, & Fernald, 2013; Goodman, McDonough, & Brown, 1998; Hoff & Naigles, 2002). What the current studies contribute is evidence that everyday verbal interactions with caregivers also provide opportunities for exercise in interpreting language, enabling children to practice skills such as segmenting speech, accessing lexical representations, and monitoring cues to word meaning. In this way, exposure to child-directed speech helps children become faster and more accurate in identifying familiar words in increasingly complex contexts, sharpening processing skills that enable faster language growth.

However, the story does not end there. New results affirming that verbal engagement with young children is beneficial for language learning in multiple ways are certainly consistent with demonstrations of how middle-class parents provide valuable scaffolding for early word learning. But because our work included a much broader sample of young children than is typically represented in laboratory studies of language development, the results force us to confront another implication of these classic studies – namely that a relative *lack* of parental support for word learning could potentially have adverse consequences. What we found was that the amount and nature of parental support for language learning varied significantly among families even within a low-SES sample, and that these differences

were linked to early emerging and consequential differences in both processing efficiency and lexical development.

Numerous studies have shown that variability in both processing speed and vocabulary could have long-term developmental consequences. Since vocabulary size predicts IQ in both adults and children (Matarazzo, 1972; Vance, West & Kutsick, 1989), an early advantage in lexical development could have cascading benefits for other aspects of language learning as well (Bates et al., 1988). Vocabulary knowledge also provides a foundation for later literacy (Lonigan, Burgess & Anthony, 2000), and preschool language skills are predictive of academic success (Alexander, Entwistle & Horsey, 1997). These findings make it clear that the early emerging differences we found in language proficiency between children from different SES backgrounds have potentially serious implications for their future developmental trajectories. However, these findings also suggest that offering young language learners more opportunities for verbal interaction has the potential to strengthen critical processing skills, which in turn would enable more efficient learning. In our ongoing research, an important goal is to show that interventions designed to increase parents' verbal engagement with their infants could change the course of early vocabulary growth and improve long-term outcomes for disadvantaged children.

## References

- Akhtar, N., Dunham, F., & Dunham, P.J. (1991). Directive interactions and early vocabulary development: The role of joint attentional focus. *Journal of child language*, 18(1), 41–49. DOI: 10.1017/S0305000900013283
- Akhtar, N., & Tomasello, M. (2000). The social nature of words and word learning. In R. M. Golinkoff & K. Hirsh-Pasek (Eds.), *Becoming a word learner: A debate on lexical acquisition*. Oxford: OUP.
- Alexander, K., Entwistle, D., & Horsey, C. (1997). From first grade forward: Early foundations of high school dropout. *Sociology of Education*, 70, 87–107. DOI: 10.2307/2673158
- Allopenna, P.D., Magnuson, J.S., & Tanenhaus, M.K. (1998). Tracking the time course of spoken word recognition using eye movements: Evidence for continuous mapping models. *Journal of Memory and Language*, 38, 419–439. DOI: 10.1006/jmla.1997.2558
- Altmann, G.T.M., & Kamide, Y. (1999). Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition*, 73, 247–264. DOI: 10.1016/S0010-0277(99)00059-1
- Arnett, J.J. (2008). The neglected 95%: Why American psychology needs to become less American. *The American Psychologist*, 63, 602–14. DOI: 10.1037/0003-066X.63.7.602
- Arnon, I., & Clark, E.V. (2011). Introduction. In I. Arnon & E.V. Clark (Eds.), *Experience, variation, and generalization: Learning a first language (TiLAR 7)* (pp. 1–11). Amsterdam: John Benjamins. DOI: 10.1075/tilar.7
- Bates, E., Bretherton, I., & Snyder, L. (1988). *From first words to grammar: Individual differences and dissociable mechanisms*. Cambridge: CUP.

- Bates, E., Dale, P.S., & Thal, D. (1995). Individual differences and their implications for theories of language development. In P. Fletcher & B. MacWhinney (Eds.), *Handbook of Child Language* (pp. 96–151). Oxford: Basil Blackwell.
- Bradley, R.H., & Corwyn, R.F. (2002). Socioeconomic status and child development. *Annual Review of Psychology*, 53, 371–399. DOI: 10.1146/annurev.psych.53.100901.135233
- Bradley, R.H., Corwyn, R.F., Burchinal, M.R., McAdoo, H.P., & Coll, C.G. (2001). The home environments of children in the United States Part II: Relations with behavioral development through age thirteen. *Child Development*, 72, 1868–86. DOI: 10.1111/1467-8624.t01-1-00383
- Bradley, R.H., Corwyn, R.F., McAdoo, H.P., & Coll, C.G. (2001). The home environments of children in the United States, Part I: Variations by age, ethnicity, and poverty status. *Child Development*, 72, 1844–1867. DOI: 10.1111/1467-8624.t01-1-00382
- Brent, M.R., & Siskind, J.M. (2001). The role of exposure to isolated words in early vocabulary development. *Cognition*, 81(2), B33–44. DOI: 10.1016/S0010-0277(01)00122-6
- Brooks-Gunn, J., & Duncan, G.J. (1997). The effects of poverty on children. *The Future of Children*, 7(2), 55–71. DOI: 10.2307/1602387
- Bruner, J. (1983). *Child's talk*. New York, NY: Norton.
- Cameron-Faulkner, T., Lieven, E.V.M., & Tomasello, M. (2003). A construction-based analysis of child directed speech. *Cognitive Science*, 27, 843–73. DOI: 10.1207/s15516709cog270\_6\_2
- Clark, H., & Clark, E.V. (1977). *Psychology and language*. New York, NY: Harcourt Brace Jovanovich.
- Clark, E.V. (1987). The principle of contrast: A constraint on language acquisition. In B. MacWhinney (Ed.), *Mechanisms of language acquisition* (Vol. 1, p. 33). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Clark, E.V. (2010). Adult offer, word-class, and child uptake in early lexical acquisition. *First Language*, 30, 250–269. DOI: 10.1177/0142723710370537
- Clark, E.V., & Estigarribia, B. (2011). Using speech and gesture to introduce new objects to young children. *Gesture*, 11, 1–23. DOI: 10.1075/gest.11.1.01cla
- Clark, E.V. (1979). *The Ontogenesis of Meaning*. Wiesbaden: Athenaion.
- Clark, E.V. (1993). *The Lexicon in Acquisition*. Cambridge: CUP. DOI: 10.1017/CBO9780511554377
- Clark, E.V. (2003). Acquisition of language: Meanings and forms. In W. Frawley (Ed.), *Oxford international encyclopedia of linguistics*, 2nd ed. (pp. 19–22). Oxford: OUP.
- Conger, R.D., McCarty, J.A., Yang, R.K., Lahey, B.B., & Kropp, J.P. (1984). Perception of child, child-rearing values, and emotional distress as mediating links between environmental stressors and observed maternal behavior. *Child Development*, 55, 2234–2247. DOI: 10.2307/1129795
- Cooper, R.P., & Aslin, R.N. (1990). Preference for infant-directed speech in the first month after birth. *Child Development*, 61, 1584–1595. DOI: 10.2307/1130766
- Durham, R.E., Farkas, G., Hammer, C.S., Tomblin, J.B., & Catts, H.W. (2007). Kindergarten oral language skill: A key variable in the intergenerational transmission of socioeconomic status. *Research in Social Stratification and Mobility*, 25, 294–305. DOI: 10.1016/j.rssm.2007.03.001
- Engle, P.L., & Black, M.M. (2008). The effect of poverty on child development and educational outcomes. *Annals of the New York Academy of Sciences*, 1136, 243–256. DOI: 10.1196/annals.1425.023

- Evans, G.W. (2004). The environment of childhood poverty. *American Psychologist*, 59, 77–92. DOI: 10.1037/0003-066X.59.2.77
- Evans, G.W. (2006). Child development and the physical environment. *Annual Review of Psychology*, 57, 423–451. DOI: 10.1146/annurev.psych.57.102904.190057
- Farkas, G., & Beron, K. (2004). The detailed age trajectory of oral vocabulary knowledge: Differences by class and race. *Social Science Research*, 33, 464–497. DOI: 10.1016/j.ssresearch.2003.08.001
- Fenson, L., Dale, P.S., Reznick, J.S., Bates, E., Thal, D.J., & Pethick, S.J. (1994). Variability in early communicative development. *Monographs of the Society for Research in Child Development*, 59. DOI: 10.2307/1166093
- Fenson, L., Marchman, V.A., Thal, D.J., Dale, P.S., Reznick, J.S., & Bates, E. (2006). *MacArthur-Bates communicative development inventories: User's guide and technical manual*, 2nd ed. Baltimore, MD: Paul Brookes.
- Fernald, A. (1985). Four-month-old infants prefer to listen to motherese. *Infant Behavior and Development*, 8, 181–195. DOI: 10.1016/S0163-6383(85)80005-9
- Fernald, A., & Kuhl, P.K. (1987). Acoustic determinants of infant preference for motherese speech. *Infant Behavior and Development*, 10, 279–293. DOI: 10.1016/0163-6383(87)90017-8
- Fernald, A. (1992). Human maternal vocalizations to infants as biologically relevant signals: An evolutionary perspective. In J.H. Barkow, L. Cosmides, & J. Tooby (Eds.), *The adapted mind: Evolutionary psychology and the generation of culture* (pp. 391–428). Oxford: OUP.
- Fernald, A. (2010). Getting beyond the “convenience sample” in research on early cognitive development. *The Behavioral and Brain Sciences*, 33, 91–92. DOI: 10.1017/S0140525X1000294
- Fernald, A., & Marchman, V.A. (2011). Individual differences in lexical processing at 18 Months predict vocabulary growth in typically developing and late-talking toddlers. *Child Development*, 83, 203–222. DOI: 10.1111/j.1467-8624.2011.01692.x
- Fernald, A., Marchman, V.A., & Weisleder, A. (2013). SES differences in language processing skill and vocabulary are evident at 18 months. *Developmental Science*, 16, 234–248. DOI: 10.1111/desc.12019
- Fernald, A., Perfors, A., & Marchman, V.A. (2006). Picking up speed in understanding: Speech processing efficiency and vocabulary growth across the 2nd year. *Developmental Psychology*, 42, 98–116. DOI: 10.1037/0012-1649.42.1.98
- Fernald, A., Pinto, J.P., Swingley, D., Weinberg, A., & McRoberts, G.W. (1998). Rapid gains in speed of verbal processing by infants in the 2nd year. *Psychological Science*, 9, 228. DOI: 10.1111/1467-9280.00044
- Fernald, A., Zangl, R., Portillo, A.L., & Marchman, V.A. (2008). Looking while listening: Using eye movements to monitor spoken language comprehension by infants and young children. In I.A. Sekerina, E.M. Fernandez, & H. Clahsen (Eds.), *Developmental psycholinguistics: Online methods in children's language processing* (pp. 97–135). Amsterdam: John Benjamins.
- Ford, M., Baer, C.T., Xu, D., Yapanel, U., & Gray, S. (2009). *The LENA Language Environment Analysis System: Audio Specifications of the DLP-0121* (pp. 1–8).
- Frank, M.C., Tenenbaum, J.B., & Fernald, A. (2013). Social and discourse contributions to the determination of reference in cross-situational word learning. *Language, Learning and Development*, 9, 1–24. DOI: 10.1080/15475441.2012.707101
- Furrow, D., Nelson, K., & Benedict, H. (1979). Mothers' speech to children and syntactic development: Some simple relationships. *Journal of Child Language*, 6, 423–442. DOI: 10.1017/S0305000900002464

- Gershkoff-Stowe, L. (2002). Object naming, vocabulary growth, and the development of word retrieval abilities. *Journal of Memory and Language*, 46, 665–687. DOI: 10.1006/jmla.2001.2830
- Goodman, J.C., McDonough, L., & Brown, N.B. (1998). The role of semantic context and memory in the acquisition of novel nouns. *Child Development*, 69, 1330–1344. DOI: 10.2307/1132269
- Grüter, T., Lew-Williams, C., & Fernald, A. (2012). Grammatical gender in L2: A production or a real-time processing problem? *Second Language Research*, 28, 191–215. DOI: 10.1177/0267658312437990
- Guillelmon, D., & Grosjean, F. (2001). The gender marking effect in spoken word recognition: The case of bilinguals. *Memory & cognition*, 29, 503–11. DOI: 10.3758/BF03196401
- Hart, B.M., & Risley, T.R. (1995). *Meaningful differences in the everyday experience of young American children*. Baltimore, MD: Paul Brookes.
- Henderson, L., Weighall, A., Brown, H., & Gaskell, G. (2013). Online lexical competition during spoken word recognition and word learning in children and adults. *Child Development*, 84(5), 1668–1685. DOI: 10.1111/cdev.12067
- Hoff, E. (2003). The specificity of environmental influence: Socioeconomic status affects early vocabulary development via maternal speech. *Child Development*, 74, 1368–1378. DOI: 10.1111/1467-8624.00612
- Hoff, E., & Naigles, L.R. (2002). How children use input to acquire a lexicon. *Child Development*, 73, 418–433. DOI: 10.1111/1467-8624.00415
- Hoff-Ginsberg, E. (1991). Mother-child conversation in different social classes and communicative settings. *Child Development*, 62, 782–796. DOI: 10.2307/1131177
- Hoff-Ginsberg, E. (1998). The relation of birth order and socioeconomic status to children's language experience and language development. *Applied Psycholinguistics*, 19, 603–629. DOI: 10.1017/S0142716400010389
- Hollingshead, A.B. (1975). *Four factor index of social status*. Unpublished manuscript, Yale University.
- Hurtado, N., Marchman, V.A., & Fernald, A. (2007). Spoken word recognition by Latino children learning Spanish as their first language. *Journal of Child Language*, 33, 227–249. DOI: 10.1017/S0305000906007896
- Hurtado, N., Marchman, V.A., & Fernald, A. (2008). Does input influence uptake? Links between maternal talk, processing speed and vocabulary size in Spanish-learning children. *Developmental Science*, 11, F31–39. DOI: 10.1111/j.1467-7687.2008.00768.x
- Huston, A.C., Mcloyd, V.C., & Garcia Coll, C. (1994). Children and poverty: Issues in contemporary research. *Child Development*, 65, 275–282. DOI: 10.1111/j.1467-8624.1994.tb00750.x
- Huttenlocher, J., Vasilyeva, M., Waterfall, H.R., Vevea, J.L., & Hedges, L.V. (2007). The varieties of speech to young children. *Developmental Psychology*, 43, 1062–83. DOI: 10.1037/0012-1649.43.5.1062
- Huttenlocher, J., Waterfall, H., Vasilyeva, M., Vevea, J., & Hedges, L.V. (2010). Sources of variability in children's language growth. *Cognitive Psychology*, 61, 343–65. DOI: 10.1016/j.cogpsych.2010.08.002
- Kuhl, P.K., Andruski, J.E., Chistovich, I.A., Chistovich, L.A., Kozhevnikova, E.V., Ryskina, V.L., Stolyarova, E.I., et al. (1997). Cross-language analysis of phonetic units in language addressed to infants. *Science*, 277, 684–686. DOI: 10.1126/science.277.5326.684

- Lew-Williams, C., & Fernald, A. (2007). Young children learning Spanish make rapid use of grammatical gender in spoken word recognition. *Psychological Science*, 18, 193–198. DOI: 10.1111/j.1467-9280.2007.01871.x
- Lonigan, C.J., Burgess, S.R., & Anthony, J.L. (2000). Development of emergent literacy and early reading skills in preschool children: Evidence from a latent variable longitudinal study. *Developmental Psychology*, 36, 596–613. DOI: 10.1037/0012-1649.36.5.596
- Marchman, V.A., & Fernald, A. (2008). Speed of word recognition and vocabulary knowledge in infancy predict cognitive and language outcomes in later childhood. *Developmental Science*, 11, F9–16. DOI: 10.1111/j.1467-7687.2008.00671.x
- Matarazzo, J.D. (1972). *Wechsler's Measurement and Appraisal of Adult Intelligence*, 5th ed. Baltimore, MD: Williams and Wilkins.
- McLoyd, V.C. (1990). The impact of economic hardship on black families and children: psychological distress, parenting, and socioemotional development. *Child Development*, 61, 311–346. DOI: 10.2307/1131096
- Mesman, J., Van IJzendoorn, M.H., & Bakermans-Kranenburg, M.J. (2012). Unequal in opportunity, equal in process: Parental sensitivity promotes positive child development in ethnic minority families. *Child Development Perspectives*, 6, 239–250. DOI: 10.1111/j.1750-8606.2011.00223.x
- Murray, A., Johnson, J., & Peters, J. (1990). Fine-tuning of utterance length to preverbal infants: Effects on later language development. *Journal of Child Language*, 17, 511–525. DOI: 10.1017/S0305000900010862
- Oller, D.K., Niyogi, P., Gray, S., Richards, J.A., Gilkerson, J., Xu, D., Yapanel, U., et al. (2010). Automated vocal analysis of naturalistic recordings from children with autism, language delay, and typical development. *Proceedings of the National Academy of Sciences of the United States of America*, 107, 13354–13359.
- Phillips, J.R. (1973). Syntax and vocabulary of mother's speech to young children: Age and sex comparisons. *Child Development*, 44(1), 182–185. DOI: 10.2307/1127699
- Phillips, M. (2011). Parenting, time use, and disparities in academic outcomes. In G.J. Duncan & R.J. Murnane (Eds.), *Whither opportunity? Rising inequality, schools, and children's life chances* (pp. 207–228). New York, NY: Russell Sage Foundation.
- Raikes, H., Pan, B.A., Luze, G., Tamis-LeMonda, C.S., Brooks-Gunn, J., Constantine, J., Tarullo, L.B., et al. (2006). Mother-child bookreading in low-income families: Correlates and outcomes during the first three years of life. *Child Development*, 77(4), 924–53. DOI: 10.1111/j.1467-8624.2006.00911.x
- Ramey, G., & Ramey, V.A. (2009). *The rug rat race*. Cambridge, MA. DOI: 10.3386/w15284
- Rowe, M.L., & Goldin-Meadow, S. (2009). Differences in early gesture explain SES disparities in child vocabulary size at school entry. *Science*, 323, 951–953. DOI: 10.1126/science.1167025
- Rowe, M.L., Raudenbush, S.W., & Goldin-Meadow, S. (2012). The pace of vocabulary growth helps predict later vocabulary skill. *Child Development*, 83, 508–25. DOI: 10.1111/j.1467-8624.2012.01805.x
- Saffran, J.R., Newport, E.L., & Aslin, R.N. (1996). Word segmentation: The role of distributional cues. *Journal of Memory and Language*, 35, 606–621. DOI: 10.1006/jmla.1996.0032
- Senechal, M. (2006). The misunderstood giant: On the predictive role of early vocabulary to future reading. In D.K. Dickinson & S.B. Neuman (Eds.), *Handbook of early literacy research*, Vol. 2 (pp. 173–184). New York, NY: Guilford Press.

- Shneidman, L.A., Arroyo, M.E., Levine, S., & Goldin-Meadow, S. (2013). What counts as effective input for word learning? *Journal of Child Language*, 40, 672–686. DOI: 10.1017/S0305000912000141
- Shneidman, L.A., & Goldin-Meadow, S. (2012). Language input and acquisition in a Mayan village: How important is directed speech? *Developmental Science*, 15, 659–673. DOI: 10.1111/j.1467-7687.2012.01168.x
- Soderstrom, M. (2007). Beyond babytalk: Re-evaluating the nature and content of speech input to preverbal infants. *Developmental Review*, 27, 501–532. DOI: 10.1016/j.dr.2007.06.002
- Swingley, D., Pinto, J.P., & Fernald, A. (1999). Continuous processing in word recognition at 24 months. *Cognition*, 71, 73–108. DOI: 10.1016/S0010-0277(99)00021-9
- Tomasello, M., & Farrar, M.J. (1986). Joint attention and early language. *Child Development*, 57, 1454–1463. DOI: 10.2307/1130423
- Tomasello, M., & Todd, J. (1983). Joint attention and lexical acquisition style. *First Language*, 4, 197–211. DOI: 10.1177/014272378300401202
- Tramontana, M.G., Hooper, S.R., & Selzer, S.C. (1988). Research on the preschool prediction of later academic achievement: A review. *Developmental Review*, 8, 89–146. DOI: 10.1016/0273-2297(88)90001-9
- Vance, B., West, K., & Kutsick, K. (1989). Prediction of Wechsler pre-school and primary scale of intelligence IQ scores for pre-school children using the PPVT Test-R and the expressive one word picture vocabulary test. *Journal of Clinical Psychology*, 45, 642–644. DOI: 10.1002/1097-4679(198907)45:4<642::AID-JCLP2270450421>3.0.CO;2-Q
- Walker, D., Greenwood, C., Hart, B.M., & Carta, J. (1994). Prediction of school outcomes based on early language production and socioeconomic factors. *Child Development*, 65, 606–621. DOI: 10.2307/1131404
- Weisleder, A., & Fernald, A. (2013). Talking to children matters: Early language experience strengthens processing and builds vocabulary. *Psychological Science*, 24, 2143–52.
- Xu, D., Yapanel, U., Gray, S., Gilkerson, J., Richards, J., & Hansen, J. (2008). Signal processing for young child speech language development. *Workshop on Child, Computer and Interaction*. Chania, Crete, Greece.
- Zangl, R., & Fernald, A. (2007). Increasing flexibility in children's online processing of grammatical and nonce determiners in fluent speech. *Language Learning and Development*, 3(3), 199–231. DOI: 10.1080/15475440701360564



# The interactional context of language learning in Tzeltal

Penelope Brown

Max Planck Institute for Psycholinguistics, Nijmegen

This paper addresses the theories of Eve Clark about how children learn word meanings in western middle-class interactional contexts by examining child language data from a Tzeltal Maya society in southern Mexico where interaction patterns are radically different. Through examples of caregiver interactions with children 12-30 months old, I ask what lessons we can learn from how the details of these interactions unfold in this non-child-centered cultural context, and specifically, what aspects of the Tzeltal linguistic and interactional context might help to focus children's attention on the meanings and the conventional forms of words being used around them.

## Introduction

As a newcomer to the field of language acquisition, entering mid-career as an anthropologist committed to exploring relationships between language, culture and thought via studying language and social interaction in specific cultural settings, I found myself bewildered by the enormous range of theory and data on child language that had accumulated over the previous 30 years. Eve Clark provided a beacon guiding this adult novice into a new world, through her warm friendship, her prolific writings, and through the Stanford Child Language Research Forum meetings which she was responsible for hosting over so many years. (The late lamented CLRF remains my idea of a model format for academic meetings). In this paper, I want to confront Eve's lucid exposition of the interactional and pragmatic contexts in which American and European children learn words with the following questions: What happens where these 'child-centered' interactional contexts do not exist? What other kinds of pragmatic and interactional ways of highlighting word/meaning connections can help children in culturally different contexts to grasp the meanings of words? I will try to give the flavor of interactions with small children in a Mayan society in southern Mexico, and pinpoint some interactional

features that can draw children's attention to relevant word/context connections even though parents and their interlocutors are not particularly trying to do so.

### Word learning in a pragmatic context

Eve Clark's word-learning theories have evolved over the years (1979, 1987, 1991, 1993, 1999, 2001, 2002, 2004, 2009a,b,c, 2012; Clark & Amaral, 2010; Clark & Estigarribia, 2011; Estigarribia & Clark, 2007), but her current position – indeed, since the mid-1980s – focussing on interactional pragmatics is clearly articulated. Rejecting a constraints approach to word learning (e.g., Markman, 1987), she argues that children learn new words from participating in the flow of conversation with adults, and they learn the meanings of these words from how the speakers use them in the context. They learn words in contexts of joint attention, physical and conversational co-presence, and they are guided by specific pragmatic assumptions. They implicitly adopt Grice's (1975) principle of cooperation, assuming that adults use words conventionally and choose their words intentionally; if the adult uses a different word they mean something different. The principles of **conventionality** and **contrast** guide children's understanding of what words mean when used in a context. They can fast-map a word-meaning correspondence in a specific context from just one exposure, and initially may over- or under-extend the meaning to other contexts, but over time they adapt their understanding of word meanings to align with the conventions of their speech community. This process can (e.g., in the case of some deictic terms) take years. Children actively seek new word names, for example, by prolifically asking "What's that?" Adults help children along the way by offering new words to the child in specific syntactic frames with the new word utterance-final and/or prosodically highlighted (e.g., 'This is an X'), they offer additional information about the concept labeled by the new word (taxonomic relations, synonyms or contrasting terms), and they repeat and rephrase the child's utterances to replace faulty or unclear pronunciation and syntax with conventionally correct models (Clark, 2001, 2002, 2009c, 2012).

This, like other theories of child language acquisition, is heavily dependent on assumptions about the nature of caregiver-child interaction. Yet it is well known – to all students of language acquisition since the 1970s – that the kinds of interaction described here, and illustrated by abundant examples from the data analyzed by Clark, do not characterize child-caregiver interactions in all cultural settings. Indeed, I would venture that here – like in the realms of psychology and language – in comparison with most other people in the world, Americans and Europeans are WEIRD (western, educated, industrialized, rich, democratic; Henrich, Heine, & Norenzayan, 2010). The work of anthropologists and cross-cultural psychologists

studying child language and development cross-culturally has made it abundantly clear that this kind of ‘child-centered’ interaction, with adults adapting to the child’s perspective and overtly drawing the child out, trying to ‘teach’ words and concepts, is not normally what children experience outside of Euro-American middle-class families (Gaskins, 2006; Lieven, 1994; Ochs & Schieffelin, 1984), or indeed, for most of the time humans have existed. Elinor Ochs and Bambi Schieffelin have been perhaps most adamant in this claim, arguing for an alternative perspective on child language learning which they call ‘language socialization’, focusing not on language acquisition *per se* but on how through interactional practices not only language but also cultural habits, beliefs, feelings, and orientations are learned (Ochs, 1988; Schieffelin, 1990; Schieffelin & Ochs, 1986a,b)<sup>1</sup>. They distinguish orientations to child-caring which are ‘child-centered’, with adults accommodating to the child – predominantly Anglo-American middle class child-care practices – from those where the child is expected to accommodate to adult-defined situations, as is characteristic of much of the rest of the world.

While conceding that there is class- and ethnicity-based variation in caregiver-child interaction even within Western societies (e.g., Clark, 2009, 2011), Clark, like most scholars of child language, does not investigate the interactional styles of these variant groups to see how well her theories apply in these kinds of caregiving contexts.

Rather than simply reiterating ‘It’s not like that everywhere!’, I would like to portray what it actually *is* like in a particular cultural setting where interaction with small children is very different from this child-centered pattern. Where are word meanings exposed, how is the child’s attention focused on relevant aspects of the environment so that form/meaning connections are made clear, how is interaction with a small child constructed and how might aspects of its construction help the child learn language? I will show that without labeling and teaching of words and concepts, the simple fact of intentional interaction between child and caregiver in contexts presupposing joint attention, with the child’s vocalizations reliably (if not frequently) responded to and ratified, is sufficient for the Tzeltal child to learn the language.

The setting is a rural Tzeltal Mayan community in Chiapas, Mexico, where I’ve worked for over forty years, and over the past twenty years worked my way down in age to small children, then to babies. The data are drawn from a corpus of roughly 700 hours of video-and/or audio-taped longitudinal, naturally-occurring interaction in the homes of five Tzeltal families recorded every four to six weeks over two and a half years or more for each focal child (see Table 1 for

1. For overviews of work in language socialization see; Garrett & Baquedano-López, 2002; Kulick & Schieffelin, 2004; Duranti, Ochs & Schieffelin, 2011; Brown & Gaskins, in press.

details); twelve older or younger siblings are also represented in this data. This corpus is supplemented by occasional video-recordings of child-caregiver interactions in thirty other families, as well as with recordings of more focused elicitations, parental interviews, and a variety of structured interactional ‘games’. In some ways this corpus is not ideal for studying word learning in relation to studies of language development in WEIRD societies: (i) the children are not intensively sampled, (ii) the focal children were often supplied by the researcher with toys that otherwise would not be available, and (iii) parents often vanished when I arrived, leaving me as the sole caregiver, along with older children in the family. This conduct is usual – parents often relegate the caregiver role to others – but it means that mother-child interactions are comparatively rare in this data set, as is the introduction of objects for which the child does not already know the name.

Elsewhere I have reported on aspects of children’s language and characteristics of Tzeltal grammar, semantics and interactional style in this community that appear to facilitate learning words (Brown, 1997, 1998a,b,c, 2001, 2002, 2007, 2008; Brown & Levinson 2000, 2009; Brown, Pfeiler, de León & Pye, 2013). As a lens on caregivers’ interactional style, in this chapter I focus on detailed examples of interaction over objects involving children aged 12–30 months. I then draw out the implications of this interactional style for children’s word learning in this cultural context.

**Table 1.** Tzeltal child-caregiver interaction recordings.

Focal child	Gender	Age range	Number of video sessions (taped by researcher)	Number of audio sessions (taped by researcher or parents)
Lus*	F	1;5 – 4;9	64	38
Xan*	F	1;2 – 3;7	47	32
Mik	M	0;7 – 4;7	56	50
Xaw	F	1;0 – 5;3	74	47
Pet	M	0;5 – 5;2	67	35
Total focal child samples			308	202
Occasional samples				
19 boys	M	0;9 – 1;3	1-2 each	0
10 girls	F	0;7 – 1;3	1-2 each	0
Total occasional samples			33	0

\*Lus and Xan are cousins living in the same extended household; 23 video sessions and 8 audio are joint, with both children present

### Tzeltal caregiver-child interactions

The community of study is a Tzeltal Mayan group of corn farmers who live in the municipio of Tenejapa in a rural area of mountainous southeastern Mexico. Settlement is dispersed; people live in extended family households amongst their cornfields and coffee patches. The language of the home is (with rare exceptions) exclusively Tzeltal. Increasingly, children are learning Spanish in school and, as young adults, entering the bilingual world of Mayan-Spanish interactions in the local cities and away even as far as the USA. But prior to attending school, small children inhabit a very small social world of their family and neighboring households, with minimal travel beyond. Childcare is generally carried out by every household member from roughly age three up.

#### *Prelinguistic babies (age 0;9 – 1;5): The many functions of ‘eh’ and ‘hm’*

For the first nine months, there is not much social interaction with Tzeltal babies, who spend most of their time in a sling on the mother’s (or other caregiver’s) back, and are slung around to the front for feeding (Brown, 1998b, 2002, 2011). More direct interaction begins when the baby shows its readiness to initiate social encounters at around nine to twelve months. At this time, joint attention and pointing episodes appear; they are rare, but they happen (Brown, 2011). Family members engage in brief interactional routines (e.g., object exchanges, body part naming) with the infant. Some examples will illustrate how these unfold.<sup>2</sup>

#### **Example 1: Xmik and mother over bird [2006 v26]**

Participants: Xmik (Xm, 12 months) and her mother (Mo).

Mother and infant are inside the house, Mother sitting on a chair, baby Xmik sitting on the floor facing away from Mo (see Figure 1). Xmik initiates joint attention to a pet bird, and a proto-conversation ensues:

[Xmik is looking around the room; her attention comes to focus on the bird hopping across the floor]

---

2. The Tzeltal speech in examples is transcribing using a practical orthography, in which x = [ʃ], ch = [tʃ], tz = [ts], j = [h], and ‘ marks a glottal stop or glottalization of the preceding consonant. Transcription conventions include the following: ° indicates very soft speech, ^ indicates sudden shift to high pitch, a square bracket [ indicates location of speech overlapping speech in the prior turn, and // indicates part of transcript omitted. TC followed by numbers identifies time code locations in the video data.



Figure 1. Xmik, mother and bird

1. Xm: *jm.* [pointing at bird]  
‘Hm’
2. Mo: *jm.*  
‘Hm.’ [gazing at bird]
3. Xm: [looks behind her, at Mo]  
*jm.* [gesturing at bird]
4. Mo: [*um jm. la' me uta.*  
‘Um hm. Come here, say to it.’]
5. Xm: *me.* [looking back at bird and gesturing toward it]  
[*me* is a particle, part of *la' me* ‘Come here’ modeled in prior turn]
6. Mo: *jm. ila'wil.*  
‘Hm. Look.’
7. Xm: *a' me.* [gesturing to bird]  
[part of *la' me* ‘Come here’]
8. Mo: *jm.*  
‘Hm’
9. Xm: [turns to look at Mo]  
A few moments later, Mother re-initiates attention to bird:
10. Mo: *ile' ch'i.*  
‘Look.’  
[points over her shoulder to bird]
11. Xm: [turns and gazes at bird]

12. Mo: [gazes at Xmik]  
*way.*  
'You see.'
13. Xm: *hee* [pointing to bird]
14. Mo: *in.*  
'This one.'
15. Xm: [looks away]
16. Mo: *in te lumine.* [points over her shoulder again, looks at bird, then back at Xmik]  
'This (one) over there.'
17. Xm: [gazes where Mo points, then shifts gaze and points to where bird has now gone]  
*hee.*
18. Mo: *in. li' bajt li'i.* [pointing to floor, where bird has now hopped to]  
'This (one). Here it went here.'
19. Xm: *jm.* [pointing]  
'Hm.'
20. Mo: *jm.*  
'Hm.'

In line 1 of this exchange Xmik calls her mother's attention to the bird and her mother responds non-committally until Xmik looks around at her, then Mother suggests what she (Xmik) should say to the bird ('Come here, say to it'). This practice, like the Kaluli *elema* routines made famous by Schieffelin (Ochs & Schieffelin, 1984; Schieffelin, 1990), is a common device to encourage a young child to say something appropriate in the context. Twice (ll. 5 and 7) Xmik partially repeats the instructed words ('me', 'a me'), and Mo's response is minimal ('Hm.' 'Look'). For the most part, the infant's gaze tracks the object, not the interlocutor's attention; she is presuming (or possibly indifferent to) the mother's visual attentional focus. The mother does not treat this moment as an opportunity to teach the child words. In the interaction over this bird, which recurs intermittently over 7 minutes or so, the mother never mentions the word 'bird'. Instead, she says things like 'look at it', 'there it is', 'go get it', 'where is it', 'fetch it down here', 'there it went', 'do you want to play with it'. This referent is presumed to be 'given' in the context (Prince, 1981), even on first mention. How to interact with people (and in this case, animals) is the focus of the interchanges, not what they are called.

A second example shows canonical joint attention over an object, plus provision of its label:

**Example 2: Lus [2005v5B 29:17- 29:24]**

Participants: Lus (13 months) and her father (Fa)

**Figure 2.** Lus, joint attention and pointing

The whole family is sitting around relaxing in the cooking house. Here Father's gaze is on the infant, Lus, who has his attention.

1. Lus: [looks up to clothesline, points up at her pants hanging on the line over the fire]  
*Eee*  
[looks at her father].
2. Fa: [gazes up to referent] *ba'ay?*  
'Where is it?'
3. Lus: *ee* [gazes at Fa, while point is held out toward referent]
4. Fa: *eh in nix yael a, a'panta[lon.*  
'Oh look there [gestures to referent], your pants.'
5. Lus: [drops toy in left hand, looks down at it]  
[The interaction episode ends as Lus's attention shifts to a toy on the ground even before the label 'pantalon' is finished]

Father's expression in line 2, *ba'ay* 'Where is it?' is a conventional response to a child's pointing at something. Rather than labeling the object, the interlocutor presupposes its identity and asks where it is, with the response often a continued pointing gesture. Here Father responds to the child's reiterated *ee* and point with a label (1.4), recognizing her action as drawing attention rather than requesting, but the child is distracted and doesn't acknowledge the label. Indeed, the link between pointing and interlocutor looking is fragile in this family: often the mother points

and says something, but the child doesn't look, or vice versa. There is no pressure on the child to interact.

Ordinary household objects are used as distractors, and many interactions over these objects transpire without speech. For example, earlier in this session (TC4:38-4:54:) Lus's Father gets up, gets a flashlight, hands it to Mother and turns it on, Mother silently hands it to Lus. The same scenario occurs at TC1:10, when Mother silently hands Lus an empty bottle to play with. Note that none of these objects used to distract or entertain her are new to her. Even when they are, they are ordinary household objects (e.g., cooking spoon, keys) being used to distract the child; their identity as objects is not in focus.

Various object-exchange routines occur with infants at this age. In Example 3, baby Lus offers the object in her hand and says *eh*. The recipient says *ixtal* (the formula for 'I receive it'), but the baby doesn't give it, thereby prolonging the exchange.

**Example 3: Lus [2005v5B, 1:29: – 1:44:]**

Participants: Lus (13 mo.), Mother, elder sister (ElSi, age 7)



Figure 3. Lus, bottle exchange

Family in kitchen house, Lus sitting on her mother's lap, facing outwards. She is holding an empty plastic bottle as a 'toy'. Father sitting at Mother's left, slightly behind her. Elder Sister sitting at Mother's right, with Elder Brother sitting behind her.

1. Lus: *eh*. [gazes at sister, holds out empty bottle towards her]
2. ElSi: *jm*. 'Hm.' [turns her gaze to Lus, takes bottle in her hand, but Lus doesn't release it. Instead, holds it out to her Mother]
3. Lus: *eh*. 'Eh'

4. Mo: *jm, ak'be Papa eke* [pointing at father] *ak'be Papa tey ek ine.*  
'Hm, give it to papa too [pointing], give it to papa over there.'  
[Mother takes bottle from Lus and points with her head, indicating what direction to give it to Father; puts bottle into Lus's hand]
  5. Mo: [whispers] <sup>°</sup>*ak'be Papa tey ek ine. ak'be teye.*  
'Give it to papa over there. Give it there.'
  6. Lus: [turns her gaze to Fa, behind her]
  7. Mo: *ila uta.*  
'Here, say to him.'
  8. Fa: *ixtal ek.* [holding out hand]  
'Taking it too' [as she holds out bottle, he takes it from her; *ixtal* is the conventional response on taking something from someone holding it out]
- [Father offers it back to her]:
9. Fa: *ya 'kan.* [holds up bottle, showing it to her]  
'You want it.'
  10. Lus: [reaches out for it]
  11. Fa: *yu' ya 'kan.*  
'Because you want it.'
  12. Lus: [withdraws her hand, still gazing at him]
  13. Lus: *eh eh eh.* [squirming, reaching]
  14. Mo: [laughs, resettles Lus on her lap, takes bottle from Father and gives it to her]

There is interactional scaffolding here but it is in support of the interaction or the object exchange routine (e.g. mother's prompt in l. 7), not for word learning.

Occasionally, however, the object is labeled, as for example a few moments later (TC32:54 – 58:)

Lus sitting on ground at her mother's feet. Lus shows/offers Mother the toy car, Mother labels it:

- Lus: *eh.* [holding up toy car]
- Mo: *ba'ay. ay a'karo. Eso.* [reaches to take it, it falls. Lus handles it on the ground]  
'Where is it. There's your car. Thus.'

In addition, the mother may repeat the child's early object labels, but does not necessarily worry about the accuracy or informativeness of the label, as in example 4 with another child, Xmal.:

**Example 4: [2008May\_v21 1:01:34: – 1:01:48 B]**

Participants: Xmal (12 months) sitting on ground outside, Mother behind her.



**Figure 4.** Xmal pointing out earphones

1. Xmal: [looks at researcher's earphones hanging on back of chair, reaches towards them, points]
2. Xmal: *titi*. [babble word for something pretty or desirable, or alternatively something dangerous that bites]
3. Mo: *titi. titi*.
4. Xmal: *titi*. [holds both hands out to earphones] *ile*.  
'Titi. Look'
5. Mo: [whispering] *tzaka*.  
'Take it'

Again, although this is a novel object for the child, the mother's response in line 5 tells her what to do with the object, not what it is called. No gaze is exchanged; the child does not comply with mother's instruction to 'take' the earphones.

Other examples show that the identity of the object a child is pointing to is not necessarily clear, nor important to the interlocutor:

**Example 5: [2008May\_v33, - 23:10: - 20:]**

Participants: Xaw (15 months), great-aunt (GrA)



**Figure 5.** Xaw and great-aunt

Xaw is tied in shawl onto his great-aunt's back. Great Aunt is standing in the kitchen house looking out the open door. She feels Xaw move, turns her head back towards him and smiles, he stands up in his shawl, leans forward looking out the doorway and says:

1. Xaw: *binti?* 'What?'
2. GrA; *jai?* 'What (did you say)?'
3. Xaw: *binti?* 'What?'
4. GrA: [gazes and points out the door, gazes back at him] *ee.* 'Eh.'
5. GrA: [looks back at him, shakes her finger 'no'] *juh uh.* 'No'  
[They both gaze out the window]

Whatever Xaw's intention in drawing Great Aunt's attention to something outdoors, she doesn't respond to its identity. Instead, she indicates her interpretation by saying 'No', i.e. we won't be going out there.

By 15 months of age Tzeltal infants are regularly engaging in functional object interchanges. In these interactions, sometimes the object is named, not necessarily by the mother:

**Example 6: APBTz2006Nov\_v1\_XunFel\_exp2.mpg**

Participants: Xun (15 mo.), her mother, her elder sister (age 6)



**Figure 6.** Xun and toothpaste tube

Xun and her mother sitting on bed in sleeping house, with two older sisters standing nearby. Baby Xun, holding toothbrush (a ‘toy’) in right hand, holds out her hand (without pointing) saying:

1. Xun: *eh.*
2. Mo: [turns her gaze to Xun, nods and smiles]
3. Xun: *eh. pa.*
4. Mo: [gazes where baby is gazing]
5. Xun: [gestures with toothbrush in her hand]
6. Mo: *ya 'kan pasta.* ‘You want ‘pasta’ (toothpaste tube).’
7. ElSi: *pasta.* ‘Pasta.’ [spoken from behind Xun, i.e., not in joint attention with her]
8. [Mother gets up and gets toothpaste from ‘toy’ bucket<sup>3</sup> on bedside table, gives it to Xun, saying *hii*]

Some time later, in the same situation, Xun holds out her left hand towards the bed table, and says *eh*, summoning her mother’s attention:

9. Xun: *eh.*

---

3. ‘Toys’ in this household are ordinary functional objects, in this case a few – keys, toothpaste, a cup – have been collected in a bucket sitting on a table next to the bed.

10. Mo: [turns gaze to Xun, nods, says something inaudible]
11. Xun: [points with right hand]
12. Mo: *ba'ay. ja' ini* [looking where Xun points to]  
‘Where is it. It's this.’ [getting keys on string, shakes them, hands them to Xun]
13. ElSi: *yawe.*  
‘Key.’ [label, but not for a newly introduced object; this is a routine ‘toy’ for this child]

Again, the labeller is not in joint attention with the infant. Later, Mother and Xun interact over the toothpaste tube again:

**Example 7: [APBTz2006Nov\_v1\_XunFel\_exp4.mpg**



**Figure 7.** Xun, toothpaste into bucket

14. Xun: [picks up toothpaste tube, turns to her mother, silently holds it out to her]
15. Mo: *ba kak?*  
‘Where do I put it?’
16. Xun: [points to ‘toy’ bucket on side table]
17. Mo: *lumine.*  
‘Over there.’  
[Mo takes toothpaste tube and puts it in bucket]

Here the mother treats Xun's action (l. 14) as communicative and interprets it as a request; she asks (in l. 15) for clarification), gets it (l. 16), and complies.

These small interactions over objects occur with many contrasting intonations of the proto-word *eh*, with which the baby indicates her level of urgency and her expectation of getting the object. Naming of the object is occasional, not prosodically highlighted as a 'here's the name of the object' instructional moment, and not necessarily uttered by someone in joint attention with the infant. The instructional formula 'Say X' is used not to teach words but to encourage the child to say what is a culturally appropriate thing to say in the context; as often as not the child does not comply. Yet crucially, these are scenes of 'structured social reciprocity' (Clark, 2013: 574): mother and child presume joint attention through these interchanges, interpret each other's communicative intentions and tie their utterances to each other's contributions, resulting in a coordinated interchanges that build up common ground, turn by turn.

#### *Beginning to speak (age 1;6 – 2;6)*

At the one-word stage the Tzeltal child starts off with a few socio-pragmatic expressions ('look', 'no'), some deictic expressions ('this', 'here'), a handful of nouns, and with bare verbs, in the absence of recurrent object-labeling practices and in light of the verb-friendly nature of Tzeltal (Brown, 1998a). In this cultural context there are no books, very few toys, and relatively few objects (e.g., small household implements, natural objects like stones, flowers, chickens, dogs, birds), most of which are very familiar. These supply the repertoire of objects with which infants are distracted and occasionally entertained.

The one exception is interaction over toys that I provided. The next example comes from a session at the researcher's house with toys supplied, which is the closest situation in my data to a laboratory in a post-industrial society, where the mother 'plays' with her child, using experimenter-provided toys. Yet the interactions look nothing like those in Euro-American developmental psychology laboratories. 'Play' often ensues with no labeling and no mutual gaze. Rather, the interactions are characterized by silent placement of toy objects, interspersed with brief engagements over them.

**Example 8: 1994Aug10, 8:30: -39:**

Participants: Pet (1 day short of 24 mo.), mother, elder sister (3;6)



**Figure 8.** Pet with farm animal set-up

Pet and his mother are seated at small table next to each other. Elder Sister is seated at Mother's right, playing separately setting up a corral for farm animals. Pet has been picking up individual pieces (a toy animal, a water trough), while Mother sits there, silent.

[Pet picks up toy corral piece and says:]

1. Pet: *te.* 'Tree/wood.' [his label for the toy corral piece]
2. Mo: *jm.* 'Hm.'
3. Pet: [looks up at Mother] *te.* 'Tree/wood.'
4. Mo: *aka xan ta xujk. aka ta xujk.*  
'Put it again at the side. Put it at the side.'
5. [Mother takes the corral piece from him and sets it up to fence in the animals]
6. Pet: *ile.* [pointing at her set-up corral piece]  
'Look.'
7. Mo: [gesturing toward array] *ijk'a bayel.*  
'Really lots.'

Here the mother responds to the child's offer of a label for the corral piece with 'Hm' in l. 2, not with a correct label ('corral'), and again in l. 4 she does not ratify his proffered term but tells him what to do with the object. Her response in l. 6 to his 'Look' – expressing appreciation – shows attention to the presupposed goal of

the activity (setting up lots of toy animals) rather than any concern with words. As a result the session consists of brief coordinated interchanges like this one, interspersed with long silences where she or Pet are setting up toys.

Later in the same interaction, in a more routine interchange [TC22:38 – 56:] Pet is standing next to the table, next to his seated mother. He turns suddenly to face her, grabs her shirt, tries to climb onto her lap but slides off. He grabs her shirt again as she lifts it to offer him her breast. He latches on and nurses standing up. No words are exchanged.

Nursing is done on demand, and may be interspersed with interaction over an object, as in Example 9:

### Example 9

Mik, age 2;0: [9/12/95 38:24: – 42:]

Participants: Mik, his elder brother Al (age 3;6), and Mother.



**Figure 9.** Mik and popbead

Mik is on his mother's lap, nursing. Al comes over, takes toy (popbead, supplied by researcher) out of Mik's hand, squats on ground to attach popbead to string of popbeads in his hand.

1. Mo: [to Mik] *°ya me 'wak'be i.*  
‘You (should) give it to him [i.e., to Alux].’
2. Al: *jo.*  
‘Hm.’
3. Mo: *°ya laj yak'bet i. ya laj yak'bet.*  
‘He'll give it (back) to you he says. He says he'll give it (back) to you.’

4. Mik: [starts to cry]
5. Mo: *ak'be laj.* [chuckles]  
‘Give it to him, he [Mik] says.’ [to Al]
6. Al: *jo, yu' ma ya jna'ix ya'tik i* [playful cajoling intonation]  
‘Hm, because don't I know (how to do it, i.e., attach popbeads to each other) now?’ (implies: ‘I do know’).  
(.4)
7. *ya to ka'bet tz'i [kala tat.]*  
‘I'll give it (back) to you my little “father”’
8. Mo: *[ya yak'bet li' ta ajk' ini.* [she turns Mik around to other breast, he's still crying]  
‘He'll give it (back) to you in a minute’
9. Mo: *ya laj spas, ya laj spas. jm, ya laj spas.*  
‘He's doing it [attaching popbeads] he [Al] says, he's doing it he says. Hm, he's doing it he says.’  
[Mik settles down to nursing again.]

This triadic interchange illustrates the characteristic catering to the youngest child's wants, with Al cajoling him (ll. 6, 7) into letting him (Al) keep the popbead temporarily. A high priority for caregivers is keeping the youngest contented, not crying, and children from roughly age three are already adept at producing this effect in their youngest siblings.

The next example shows how objects are presupposed in context even when they haven't been mentioned or alluded to all day:

**Example 10: Xan [Lus\_Xan1994Sep29pt1 19:26 – 45:]**

Participants: Xan (23 mo.), Mother, researcher PB



**Figure 10.** Xan's foot

[Mother notices Xan limping slightly, approaches and addresses her from about 8 feet away]

1. Mo: *a'bojox a'wakan antz.*  
‘You cut your foot “woman”’
2. Xan: *joo?*  
‘Huh?’
3. Mo: *a'bojox a'wakan.*  
‘You cut your foot.’
4. Xan: *jo'o.*  
‘Uh’uh.’ [i.e., no.]
5. PB [calling from about 8 feet away]: *yak yilel, melel ila'wil bit'il ya xben.*  
‘Yes it appears, really, see how she walks [with a limp]!’
6. Mo: *jm, ja' sbojox wan.*  
‘Hm, she cut (it) perhaps.’

[Mother walks over to Xan, looks down at her foot]

7. Mo: *banti a'bojox?*  
‘Where did you cut it?’
8. Xan: *jmm?*  
‘Huh?’
9. Mo: *banti a'bojox?* [bends, peering at Xan’s foot]  
‘Where did you cut it?’
10. Xan: *lum to ine.* [no gaze or point, stares down at toy she’s holding]  
‘Way over there.’
11. Mo. [bends over, touches X’s toes] *la'wile.*  
‘You saw it.’
12. [Xan looks at her own foot.]
13. Mo: *ma me a'tzak me asaronae*  
‘Don’t take/grab that hoe.’ [presumed source of cut foot]
14. Xan: *jo.*  
‘Hm’

[Mother walks away]

Here we see how, with minimal help from the child, mother infers a prior activity and takes the opportunity (in l. 13) to offer a general rule of behavior (don’t handle dangerous things like hoes) which is minimally accepted (with ‘Hm’) by the child. Note also how the child uses a repair strategy (Huh? Hm? in ll. 2, 8) to participate in the conversation even when her word-production abilities are still fairly minimal.

At a slightly older age, when children can participate more contentfully in social interchanges, more extended interactions over objects transpire. The final example is drawn from a rare occurrence when a mother actually sits down and plays with her child using toys supplied by the researcher. They engage in a conversation – or rather, a running commentary – over the toys as they handle them. She speaks very softly, almost as if to herself, even when repeating correctly or expanding her son's mispronounced or ungrammatical utterances. Notice that the words she repeats are mostly *verbs*.

**Example 11: Pet [16/1/95 10:20: – 12:56:]**

Participants: Pet (2;5), mother



Figure 11. Pet, mother, and toy car

Pet and his mother are sitting on a plastic sheet on the ground with toys all around. Mother makes a move to play with the toy truck, picks it up, and says very softly 'Let's do it', and they proceed to put toy people and animals into the back of the truck. The excerpt begins when the mother notices that there is no cord to pull the truck along:

1. Mo: *jn. (.8) pero ma'yuk lek yak'ul.*  
'Hm. But it doesn't have any cord' [to pull the car with]  
[Mother looks around for something to make a cord with]
2. Pet: *ay!* [to self, re toy he put into the back of the truck]  
'Oh!'
3. Mo: [looking back at Pet's activity with the truck] *butkej.*  
'It tipped over.'
4. Pet: *mm butkej. (.7) ut. ujj* [effort sounds] *yukube.* REPEAT (verb)  
'Hm it tipped over. (.7) ut ujj. yukube' [unintelligible word]

5. Mo: °*jm.*  
     ‘Hm.’
6. Pet: *ukbe.* (2.4) *jnnnnn.* [making sound of truck engine]  
     ‘Ukbe [unintelligible word, possible attempt at *butk’ej*] Hnnnnn.’  
     (4.1) [Pet is running truck along]
7. Pet: *uj* [frustration sound]
8. Mo: °*ya xbutk’ej* [toys in back of truck] REPEAT/EXPANSION (verb)  
     ‘It falls over.’
9. Pet: *uj.*  
     ‘Uh’ [noncommittal agreement]
10. Mo: °*yak.*  
     ‘Yes.’  
     (5.6) [Mother runs the truck along]
11. Pet: *wajcho.* [re: toy that fell out of truck]  
     ‘It fell over.’ [target: *balch’oj*]
12. Mo: *ya xbalch’o jilel.* EXPANSION (verb + directional)  
     ‘It falls down remaining behind.’
13. Pet: *jilel.* (1.7) *uuy.* REPEAT (directional adverb)  
     ‘Remaining behind. uuy.’  
     (1.6) [Pet stands toy back up]
14. Mo: °*tejk’ajix.*  
     ‘Now it’s standing up.’
15. Pet: *tek’ajix.* REPEAT (verb)  
     ‘Now it’s standing up.’
16. Mo: °*jm.*  
     ‘Hm.’  
     // [After a few seconds of Pet playing with toys in back of truck, Mother reaches behind chair for a dried leaf, shreds it to make a cord]  
     (3.0)
17. Pet: *ay!*  
     ‘Oh!’
18. Mo: *uy* [mild dismay sound]
19. Pet: *bajch’oj!*  
     ‘It fell over!’ [his toy in back of truck] [target: *balch’oj*]

20. Mo: *ya xbalch'oj.* EXPANSION (verb)  
‘It's falling over.’  
(1.3)
21. Pet: *tut mut.*  
‘Small chicken.’ [toy animal]  
(1.3) [Mo is making a cord]
22. Mo: *°ya xbalch'oj ala mut.* EXPANSION (verb + noun)  
‘The little chicken is falling over.’  
(1.1)
23. Pet: *sok ay kado.* [holds up small toy car]  
‘And there’s a car.’ [target: *karo*]
24. Mo: *°ay karo.* REPEAT (predicate + noun)  
‘There’s a car.’
25. Pet: *jm.*  
‘Hm.’
26. Mo: *mmm.*  
‘Mmm.’  
(1.4)
27. Pet: *kuchotik bel.*  
‘(It) carries us away.’
28. Mo: *ya xkuchotik bel* EXPANSION (verb + directional)  
‘It carries us away.’
29. Pet: *ja’ ni.*  
‘It’s this.’
30. *ay! kuchix bel. ii ii.* REPEAT (verb + directional)  
‘Oh! (It) carried it away. ii ii’
31. *ee. (.8) ee.*  
‘Eh. Eh.’  
(4.8)
32. Pet: *jich ini. jich ini ek.*  
‘Like this, like this too.’
33. Mo: *jn maili i, ya kala chuktik.*  
‘Hm, wait, we’ll just tie it.’  
[Mother ties her handmade cord onto the front of the truck]

34. Pet: *jich ini ek.*  
           'Like this too.'
35. Mo: *jm.*  
           'Hm'  
           (1.5)
36. Pet: *kara. (1.2) .nit. (1.0) mm.*  
           'Car. Pull. Hm.'  
           (3.5) [Mother is finishing tying the cord onto the truck]
37. Mo: *ba nita bel i.* EXPANSION (verb + directional)  
           'Go pull it awaywards.'  
           (1.4)
38. Pet: *mali to.*  
           'Wait.'
39. Mo: *ya xkuch jente.*  
           'It carries people.'
40. Pet: *jente.* REPEAT (noun)  
           'People.'  
           (.9)
41. Mo: *ya xkuch kirsanu.* EXPANSION + REPHRASE (synonym)  
           'It carries people.'  
           (1.4)
42. Pet: *jm, nita mama.*  
           'Hm, pull it mama.'
43. Mo: *jm.*  
           'Hm.'
44. Pet: *jm, ja' ini.* [takes up corral piece]  
           'Hm, it's this.'
45. Mo: *°jn, ya xkuch koral.*  
           'Hm, it's carrying (the) fence'
46. Pet: *kolal.* REPEAT (noun)  
           'Fence'
47. Mo: *°ya xkuch koral.* REPEAT/EXPANSION (verb + noun)  
           'It's carrying (the) fence'
48. Pet: *jm yan.*  
           'Hm. Another (one)?'

49. Mo: *jn.* (.8) *ja' ni.* [offers the truck with pull-line tied on it]  
'Hm. It's this.'
- (3.4) [Pet puts another small toy in back of truck]
50. Pet: *yan cheb.*  
'Another (one) two.'
51. Mo: *cheb.* REPEAT (number)  
'Two.'
52. Pet: *jn.*  
'Hm.'
53. Mo: *jn.*  
'Hm.'

In this exchange the game involves sharing the activity of pulling a toy car along with toy people and animals in it. The mother facilitates this play activity by making a cord to pull the car. In the joint attention over these toys, what can be done with them (e.g., put toys into truck, pull truck along) and what happens to them (fall over, stand up) are the focus. Hence, verbs and other predicates are foregrounded. This child has pronunciation difficulties at this age, which make him hard to understand. For the most part, his mother does not try to probe further what he (unintelligibly) said but carries on with the play activity. When she does understand him, she sometimes repeats his utterances or expands them with corrected pronunciation and morphemes (ll. 12, 20, 22, 24, 28, 47), providing feedback about the conventional form of the terms. Notably, she repeats in the manner conventional among adults for conversational (next turn) repeats – largely repeating the predicate – while reformulating it, inserting the aspect and person marking that is still mostly missing in Pet's utterances. Thus, with no communicative intention to teach words, she models the verb forms appropriately in this specific context and provides alternative terms for the same object (*jente*, *kirsanu*, both meaning 'people' – though note that in this case there is no evidence of contrasting meanings for the two terms).

Space precludes supplying further examples, but as the children's speaking competence increases, longer and more content-full interactions occur, especially with older siblings or caregivers other than the parents. While there continues to be no evidence of explicit attention to teaching words and concepts, children are increasingly drawn into interactions, oriented to practical things that they must learn (obedience, eating practices, care of younger siblings), and exposed to teasing routines. Many children today attend pre-school from the age of three or four, bringing them into daily contact with a much wider range of interlocutors and exposing them to Spanish in the form of songs and routines. (Preschool teachers

are native speakers of Tzeltal, and most of the class is conducted in Tzeltal.) At home, older children draw the small ones into their games of role-play (e.g., pretending to be cooking tortillas, travelling on a bus, or visiting the local clinic) and telling stories and dreams. By age five, children of both sexes are generally good caregivers of smaller children and competent errand-runners to other local households. They grow up quickly.

## Discussion

Tzeltal children, like those in WIERD cultures, learn new words from participating in the flow of conversation with others, and they learn the meanings of these words from how speakers use them in particular contexts. Caregivers' communicative goals shape the linguistic and social contexts in which children learn, and these goals vary across cultures. Tzeltal caregivers of all ages do not put effort into focusing attention on infants' learning words and their meanings. Their interactions with small children are about the here and now, about what needs to be done and what can be done or must not be done with the objects and events at hand. These encounters are most often initiated by the child, so that most of the time the caregiver presupposes the child's attention and presumes the referent is given in the context. Topics of talk are local and cover the small range of objects and activities available to children. In the absence of books, museums, or travel to unfamiliar places, small children have no access to exotic objects like zebras, crocodiles, or dinosaurs that need to be labeled for them. Routine activities (e.g., nursing or eating, going to sleep, bathing) often are carried out without any verbal interchange at all or with merely murmured reassurances. Children do not regularly ask 'What's that?' and adults do not regularly label 'That's an X,' much less provide information about semantic domains or related words. Metalinguistic talk about words, while it may well be helpful in promoting the rapid build-up of a child's vocabulary, is clearly not an essential part of the language socialization process.<sup>4</sup>

What is less clear is how these differences in the communicative goals of interaction impact the word learning process. Clark is surely right to emphasize the nature of word learning as relying upon pragmatic inference. Such reliance seems even more apparent in the highly indexical, presupposing contexts of Tzeltal

4. The target size of vocabulary – the number of words adults master – is doubtless quite different in nonliterate societies like the one considered here from the vocabulary size of middle-class people in literate cultures. And vocabulary size is not a preoccupation of Tzeltal speakers, though if a child doesn't begin to speak and interact by about 2;6 or 3, cultural remedies are brought into play.

speakers' interactions with small children. There is extensive direct corroboration of interlocutors understanding the child's communicative intent, as well as indirect correction of child errors, provided by older Tzeltal speakers' repeats and reformulations of child utterances as a natural part of talking together (Chouinard & Clark, 2003). Clark is also surely right to emphasize the central importance of joint attention episodes in early interactions as the kind of context where children initially pick up word meanings. Around the age of twelve months children across the world do appear to engage with others in episodes of joint attention and initiate pointing behavior with them (Carpenter, Nagell, & Tomasello, 1998; Liszkowski, Brown, Gallagher, Takada, & de Vos, 2012; de León, 2005, 2011; Takada, 2011; Tomasello, 2003), though the frequency of these interactions varies radically (Salomo & Liszkowski, 2013), as does the likelihood that they will be accompanied by a label for the object or activity in focus (Brown, 2011).

It is what happens in the year or two *after* this apparently universal orientation to jointly attending in interaction that might be expected to make a significant difference to the process or speed of language learning. For example, attention is organized differently in different cultural settings. Various groups of Mayans have been observed to have quite a different style with respect to attention management from that characteristic of e.g. American middle-class parents, resulting in a 'distributed' pattern of attention in Mayan children (Chavajay & Rogoff, 1999; de León, 1998, 2005, 2011; Gaskins, 2006; Gaskins & Paradise, 2010; Rogoff, Mistry, Gonci & Mosier, 1993; Rogoff, Moore, Najafi, Dexter, Correa-Chávez, & Solis, 2007; Rogoff, Paradise, Mejía-Arauz, Correa-Chávez, & Angelillo, 2003). The variant nature of input speech in Mayan families has also entered the discussion (Pye, 1986; Schneidman & Goldin-Meadow, 2012a,b), as has the nature of adult Tzeltal interaction with its intensive conversational repetition (Brown, 1998b) and its avoidance of mutual gaze (Rossano, Brown, & Levinson, 2009). Monitoring of joint attention may well be done without overt gaze shifts in this cultural setting.

The Tzeltal study raises the question of whether the role of *joint* attention in word learning has perhaps been overemphasized. Tzeltal children, socialized to attend keenly to what is going on all around them rather than focusing intently on one activity, may pick up word meanings from hearing others use words around them, not necessarily in joint attention with them. The role of such overheard speech is contentious in current theorizing: some research finds that children can indeed learn words from overhearing the speech of others (e.g., Akhtar, 2005; Akhtar & Gernsbacher, 2007; Akhtar, Jipson & Callanan, 2001; Crago, Allen, & Hough-Eyamie, 1997). Clark (2013:578) points out that children attend to overhearing their own name spoken by others, and their ability to enter a conversation they've been overhearing provides evidence that they can process speech not addressed to them. But a number of recent studies have found that overheard speech

does not predict a child's vocabulary size at a later age, whereas speech uttered in joint attention with a child does (see Wiesleder & Fernald, this volume). Most relevant here are the studies of Schneidman in a Yucatec Mayan village where children have an input and interactional environment comparable to the Tzeltal one (Schneidman, Arroyo, Levine, & Goldin-Meadow, 2013; Schneidman & Goldin-Meadow, 2012): in a quantitative study they too found a correlation between directed speech and child's later vocabulary but no correlation with overheard speech. Yet there are methodological doubts with these findings: in this and other studies that compare child-directed vs. overheard input speech, all speech in the situation not directed to the child is defined as 'overheard'. Yet clearly, the relevant comparison is with speech not addressed to the child but to which the child *is at least potentially attending*. Speech around a non-attending child (e.g. speech on the telephone, or radio) is not likely to have any effects. Even if overheard speech is less effective in predicting later vocabulary size than speech addressed directly to the child, its role in Mayan lexical acquisition seems worth pursuing further.

That the interactional skills and practices of caregivers in this Tzeltal community are notably different from those in middle-class families in WEIRD societies is not the end of the story. It behooves us to ask what aspects of this interactional style in its linguistic and cultural context enable Tzeltal and other Mayan children to map words to contexts at a rate that does not appear to be markedly different from that of western children, although, as mentioned (fn. 4), target vocabulary size may be much smaller. Positive feedback in the form of patterns of repetition, rewording, and expansion of child utterances (see e.g., Brown, 1998b; Clark & Bernicot, 2008; Küntay & Slobin, 1996) is no doubt an important part of the answer, even if their frequency is far less than it is in the child-centered style of middle-class caregivers in post-industrial societies. Further research is needed to pin down and quantify the role of language-and-culture-specific properties that might influence the ease of word learning. For Tzeltal, several have been proposed: the semantic specificity of Tzeltal verbs (Brown, 1998a, 2001, 2007, 2008), Tzeltal interactional practices like conversational repetition (Brown, 1998b), and the centrality of gesture and gaze in learning the Tzeltal absolute system of spatial reckoning (Brown & Levinson, 2000, 2009). In light of the fact that more than thirty years of anthropological scholarly reiteration that cultural contexts of language learning differ around the world appears to have had no impact on word learning theories, one has to wonder: will ethnographically-based language socialization factors ever enter psycholinguistic theorizing? At the very least we need to create a more refined typology of the kinds of interactional practices in which children are socialized in different societies and make specific predictions of how such cultural and contextual differences influence the processes of word and language learning. Then, ideally, we could collect the kind of intensively-sampled and

culturally-meaningful interactional data around the world that would enable testing of these predictions. This academic undertaking has hardly begun.

## References

- Akhtar, N. (2005). The robustness of learning through overhearing. *Developmental Science*, 8(2), 199–209. DOI: 10.1111/j.1467-7687.2005.00406.x
- Akhtar, N., & Gernsbacher, M.A. (2007). Joint attention and vocabulary development: A critical look. *Language and Linguistics Compass*, 1(3), 195–207. DOI: 10.1111/j.1749-818X.2007.00014.x
- Akhtar, N., Jipson, J., & Callanan, M.A. (2001). Learning words through overhearing. *Child Development*, 71(2), 416–430. DOI: 10.1111/1467-8624.00287
- Brown, P. (1997). Isolating the CVC root in Tzeltal Mayan: A study of children's first verbs. In E.V. Clark (Ed.), *Proceedings of the 28th Annual Child Language Research Forum* (pp. 41–52). Stanford, CA: CSLI.
- Brown, P. (1998a). Children's first verbs in Tzeltal: Evidence for an early verb category. *Linguistics*, 36(4), 713–53. DOI: 10.1515/ling.1998.36.4.713
- Brown, P. (1998b). Conversational structure and language acquisition: The role of repetition in Tzeltal adult and child speech. *Journal of Linguistic Anthropology*, 8(2), 197–221. DOI: 10.1525/jlin.1998.8.2.197
- Brown, P. (1998c). Early Tzeltal verbs: Argument structure and argument representation. In E.V. Clark (Ed.), *Proceedings of the 29th Annual Child Language Research Forum* (pp. 129–140). Stanford, CA: CSLI.
- Brown, P. (2001). Learning to talk about motion UP and DOWN in Tzeltal: Is there a language-specific bias for verb learning? In M. Bowerman & S.C. Levinson (Eds.), *Language acquisition and conceptual development* (pp. 512–543). Cambridge: CUP. DOI: 10.1017/CBO9780511620669.019
- Brown, P. (2002). Everyone has to lie in Tzeltal. In S. Blum-Kulka & C. Snow (Eds.), *Talking to adults: The contribution of multiparty discourse to language acquisition* (pp. 241–275). Mahwah, NJ: Lawrence Erlbaum Associates.
- Brown, P. (2007). Culture-specific influences on semantic development: Learning the Tzeltal 'benefactive' construction. In B. Pfeiler (Ed.), *Learning indigenous languages: Child language acquisition in Mesoamerica* (pp. 119–154). Berlin: Mouton de Gruyter. DOI: 10.1515/9783110923148.119
- Brown, P. (2008). *Verb specificity and argument realization in Tzeltal child language*. In M. Bowerman & P. Brown (Eds.), *Crosslinguistic perspectives on argument structure: Implications for language acquisition* (pp. 167–189). Mahwah, NJ: Lawrence Erlbaum Associates.
- Brown, P. (2011). The cultural organization of attention. In A. Duranti, E. Ochs, & B.B. Schieffelin (Eds.), *Handbook of language socialization* (pp. 29–55). Oxford: Blackwell. DOI: 10.1002/9781444342901.ch2
- Brown, P. & Levinson, S.C. (2000). Frames of spatial reference and their acquisition in Tenejapan Tzeltal. In L. Nucci, G. Saxe, & E. Turiel, (Eds.), *Culture, thought, and development* (pp. 167–197). Mahwah NJ: Lawrence Erlbaum Associates.

- Brown, P. & Levinson, S.C. (2009). Language as mind tools: Learning how to think through speaking. In J. Guo, E. Lieven, N. Budwig, S. Ervin-Tripp, K. Nakamura, & S. Ozcaliskan (Eds.), *Crosslinguistic approaches to the psychology of language: Research in the traditions of Dan Slobin* (pp. 451–464). Mahwah NJ: Lawrence Erlbaum Associates.
- Brown, P., Pfeiler, B., de Leon, L., & Pye, C. (2013). The acquisition of agreement in four Mayan languages. In E. Bavin & S. Stoll (Eds.), *The acquisition of ergativity* (pp. 271–306). Amsterdam: John Benjamins.
- Carpenter, M., Nagell, K., & Tomasello, M. (1998). Social cognition, joint attention, and communicative competence from 9 to 15 months of age. *Monographs of the Society for Research in Child Development*, 63(4/255), 1–174. DOI: 10.2307/1166214
- Chavajay, P., & Rogoff, B. (1999). Cultural variation in management of attention by children and their caregivers. *Developmental Psychology*, 35(4), 1079–1090. DOI: 10.1037/0012-1649.35.4.1079
- Chouinard, M.M., & Clark, E.V. (2003). Adult reformulations of child errors as negative evidence. *Journal of Child Language*, 30(3), 637–669. DOI: 10.1017/S0305000903005701
- Clark, E.V. (1979). *The Ontogenesis of Meaning*. Wiesbaden: Athenaion.
- 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. (1991). Acquisition principles in lexical development. In S.A. Gelman & J.B. Bynnes (Eds.), *Perspectives on language and thought* (pp. 31–71). Cambridge: CUP. DOI: 10.1017/CBO9780511983689.003
- Clark, E.V. (1993). *The lexicon in acquisition*. Cambridge: CUP. DOI: 10.1017/CBO9780511554377
- Clark, E.V. (1999). Acquisition in the course of conversation. In A. Goldberg & E.H. Antonsen (Eds.), *Studies in the Linguistic Sciences (Forum Lectures from the 1999 Linguistic Institute)*, 29(2), 1–18.
- Clark, E.V. (2001). Grounding and attention in language acquisition. In M. Andronis, C., Ball, H. Elston, & S. Neuvel (Eds.), *Papers from the 37th Meeting of the Chicago Linguistic Society*, Vol. 1 (pp. 95–116). Chicago, IL: Chicago Linguistic Society.
- Clark, E.V. (2002). Making use of pragmatic inferences in the acquisition of meaning. In D. Beaver, S. Kaufmann, B. Clark, & L. Casillas (Eds.), *The construction of meaning* (pp. 45–58). Stanford, CA: CSLI.
- Clark, E.V. (2009a). Lexical meaning. In E. Bavin (Ed.), *Cambridge handbook of child language* (pp. 283–299). Cambridge: CUP.
- Clark, E.V. (2009b). What shapes children's language? Child-directed speech, conventionality, and the process of acquisition. In V.C. Mueller Gathercole (Ed.), *Routes to language: Studies in honour of Melissa Bowerman* (pp. 233–254). Mahwah, NJ: Lawrence Erlbaum Associates.
- Clark, E.V. (2009c). *First language acquisition, 2nd ed.* Cambridge: CUP. DOI: 10.1017/CBO9780511806698
- Clark, E.V. (2012). Children, conversation, and acquisition. In M. Spivey, K. McRae & M. Joanisse (Eds.), *Cambridge handbook of psycholinguistics* (pp. 573–588). Cambridge: CUP. DOI: 10.1017/CBO9781139029377.039
- Clark, E.V., & Amaral, P.M. (2010). Children build on pragmatic information in language acquisition. *Language & Linguistics Compass*, 4(7), 445–457. DOI: 10.1111/j.1749-818X.2010.00214.x

- Clark, E.V., & Bernicot, J. (2008). Repetition as ratification: How parents and children place information in common ground. *Journal of Child Language*, 35, 349–371. DOI: 10.1017/S0305000907008537
- Clark, E.V., & Estigarribia, B. (2011). Using speech and gesture to inform young children about unfamiliar word meanings. *Gesture*, 11(1), 1–23. DOI: 10.1075/gest.11.1.01cla
- Crago, M.B., Allen, S.E.M., & Hough-Eyamie, W.P. (1997). Exploring innateness through cultural and linguistic variation. In M. Gopnik (Ed.), *The biological basis of language* (pp. 70–90). Oxford: OUP.
- Duranti, A., Ochs, E., & Schieffelin, B. (Eds.). (2011). *The handbook of language socialization*. Oxford: Wiley-Blackwell. DOI: 10.1002/9781444342901
- Estigarribia, B., & Clark, E.V. (2007). Getting and maintaining attention in talk to young children. *Journal of Child Language*, 34, 799–814. DOI: 10.1017/S0305000907008161
- Garrett, P.B., & Baquedano-López, P. (2002). Language socialization: Reproduction and continuity, transformation and change. *Annual Review of Anthropology*, 31, 339–361. DOI: 10.1146/annurev.anthro.31.040402.085352
- Gaskins, S. (2006). Cultural perspectives on infant-caregiver interaction. In N.J. Enfield & S.C. Levinson (Eds.), *Roots of human sociality: Culture, cognition and interaction* (pp. 279–298). Oxford: Berg.
- Gaskins, S., & Paradise, R. (2010). Learning through observation. In D.F. Lancy, J. Bock, & S. Gaskins (Eds.), *The anthropology of learning in childhood* (pp. 85–117). Lanham, MD: Alta Mira Press.
- Grice, P. (1975). Logic and conversation. In P. Cole & J.L. Morgan (Eds.), *Syntax and semantics, Vol. 3: Speech acts* (pp. 41–58). New York, NY: Academic Press.
- Henrich, J., Heine, S.J. & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences*, 33(2–3), (pp. 1–75). DOI: 10.1017/S0140525X09991105
- Kulick, D., & Schieffelin, B. (2004). Language socialization. In A. Duranti (Ed.), *A companion to linguistic anthropology* (pp. 349–368). Malden, MA: Blackwell.
- Küntay, A., & Slobin, D.I. (1996). Listening to a Turkish mother: Some puzzles for acquisition. In D.I. Slobin, J. Gerhardt, A. Kyrtzis, & J. Guo (Eds.), *Social interaction, social context, and language: Essays in honor of Susan Ervin-Tripp* (pp. 265–286). Hillsdale, NJ: Lawrence Erlbaum Associates.
- de León, L. (1998). The emergent participant. *Journal of Linguistic Anthropology* 8(2), 131–161. DOI: 10.1525/jlin.1998.8.2.131
- de León, L. (2005). *La llegada del alma: Lenguaje, infancia y socialización entre los Mayas de Zinacantán*. Mexico: CIESAS-INAH-CONACULTA.
- Lieven, E. (1994). Crosslinguistic and crosscultural aspects of language addressed to children. In C. Gallaway & B.J. Richards (Eds.), *Input and interaction in language acquisition* (pp. 56–72). Cambridge: CUP. DOI: 10.1017/CBO9780511620690.005
- Liszkowski, U., Brown, P., Gallagher, T., Takada, A., & de Vos, C. (2012). A prelinguistic universal of human communication. *Cognitive Science*, 36, 698–713. DOI: 10.1111/j.1551-6709.2011.01228.x
- Markman, E.M. (1987). How children constrain the possible meanings of words. In U. Neisser (Ed.), *Concepts and conceptual development: Ecological and intellectual factors in categorization* (pp. 255–287). Cambridge: CUP.
- Ochs, E. (1988). *Culture and language development*. Cambridge: CUP.

- Ochs, E., & Schieffelin, B. (1984). Language acquisition and socialization: Three developmental stories. In R. Shweder & R. LeVine (Eds.), *Culture theory: Mind, self, and emotion* (pp. 276–320). Cambridge: CUP.
- Prince, E. (1981). Toward a taxonomy of given-new information. In P. Cole (Ed.), *Radical pragmatics* (pp. 223–255). New York, NY: Academic Press.
- Pye, C. (1986). Quiche' Mayan speech to children. *Journal of Child Language*, 13, 85–100.
- Rogoff, B., Mistry, J., Gonci, A., & Mosier, C. (1993). Guided participation in cultural activity by toddlers and caregivers. *Monographs of the Society for Research in Child Development*, 58(8/236), 1–179. DOI: 10.2307/1166109
- Rogoff, B., Moore, L., Najafi, B., Dexter, A., Correa-Chávez, M., & Solis, J. (2007). Children's development of cultural repertoires through participation in everyday routines and practices. In J.E. Grusec & P.D. Hastings (Eds.), *Handbook of socialization: Theory and research* (pp. 490–515). New York, NY: The Guilford Press.
- Rogoff, B., Paradise, R., Mejía-Arauz, R., Correa-Chávez, M., & Angelillo, C. (2003). Firsthand learning through intent participation. *Annual Review of Psychology*, 54(1), 175–203. DOI: 10.1146/annurev.psych.54.101601.145118
- Rossano, F., Brown, P., & Levinson, S.C. (2009). Gaze, questioning, and culture. In J. Sidnell (Ed.), *Comparative studies in conversation analysis* (pp. 187–249). Cambridge: CUP. DOI: 10.1017/CBO9780511635670.008
- Salomo, D., & Liszkowski, U. (2013). Socio-cultural settings influence the emergence of prelinguistic deictic gestures. *Child Development*, 84(4), 1296–1307.
- Schieffelin, B.B. (1990). *The give and take of everyday life: Language socialization of Kaluli children*. Cambridge: CUP.
- Schieffelin, B.B., & Ochs, E. (1986a). Language socialization. *Annual Review of Anthropology*, 15, 163–191. DOI: 10.1146/annurev.an.15.100186.001115
- Schieffelin, B.B., & Ochs, E. (Eds.) (1986b). *Language socialization across cultures*. Cambridge: CUP.
- Schneidman, L.A., Arroyo, M.E., Levine, S.C., & Goldin-Meadow S. (2013). What counts as effective input for word learning? *Journal of Child Language*, 40, 672–686.
- Schneidman, L.A., & Goldin-Meadow, S. (2012a). Mayan and US caregivers simplify speech to children. In A.K. Biller, E.Y. Chung, & A.E. Kimball (Eds.), *Proceedings of the 36th annual Boston University Conference on Language Development* (pp. 536–544). Sommerville, MA: Cascadilla Press.
- Schneidman, L.A., & Goldin-Meadow, S. (2012b). Language input and acquisition in a Mayan village. *Developmental Science*, 15(5), 659–673. DOI: 10.1111/j.1467-7687.2012.01168.x
- Takada, A. (2011). Pre-verbal infant-caregiver interaction. In A. Duranti, E. Ochs, & B.B. Schieffelin (Eds.), *Handbook of language socialization* (pp. 56–80). Oxford: Blackwell. DOI: 10.1002/9781444342901.ch3
- Tomasello, M. (2003). *Constructing a language: A usage-based theory of language acquisition*. Cambridge, MA: Harvard University Press.



# Conversation and language acquisition

## Unique properties and effects

Edy Veneziano

Université Paris Descartes – MoDyCo, CNRS

Conversational exchanges have unique properties that have a great potential for language acquisition such as the motivation to communicate, information about language meanings and structures, and the construction of new ways of expression. Results from the literature and from new data analyses show that conversational exchanges have both immediate and long-term effects. Two properties are considered here to have particularly great learning potential. One is the facilitated access to unities of meaning underlying the successive transformations in utterance shapes, cognitively analogous to the discovery of physical and logico-mathematical invariances under the transformations of shape and configuration. The other is the co-construction of sequences that foreshadow behaviors still beyond the individual child's present capacities.

*Conversation provides the primary setting for language acquisition, and it is in conversational settings that children display their emerging knowledge together with their skill in using a language*  
Eve V. Clark (2009, p. 41)

- (1) Camille (1;10.12) and her mother

Mother<sub>1</sub>: *c'est quoi ces oiseaux là?* [pointing at image of swans on book]  
'what are these birds there?'

Child<sub>1</sub>: *les oiseaux ils volent*  
'the birds, they fly'

Mother<sub>2</sub>: *comment ils s'appellent ces oiseaux là?*  
'what are they called, these birds here?'

Child<sub>2</sub>: *regarde les oiseaux ils volent*  
'look the birds they fly'

- Mother<sub>3</sub>: *oui les oiseaux ils volent mais ceux-là ils volent pas ils nagent là*  
 [pointing at image of swans on book]  
 ‘yes the birds they fly, but these do not fly, they swim there’
- Child<sub>3</sub>: *hein?*
- Mother<sub>4</sub>: *ils nagent ceux-là, ils volent pas # hein? ## ils sont dans l'eau*  
 [CHI looks at image and MOT  
 alternatively]## *c'est des cygnes?*  
 ‘those swim, they don't fly # hein? ##they are in the water, they  
 are swans?’
- Child<sub>4</sub>: *nh* [confirmation] [%act: looking at image on the book]
- Mother<sub>5</sub>: *t'en a déjà vu toi des cygnes #1.6 où c'est que t'as vu des cygnes  
 toi?*  
 ‘you have seen already swans you #1.6 where have you seen  
 swans you?’
- Child<sub>5</sub>: *dans l'eau*  
 ‘in the water’
- Mother<sub>6</sub>: *oui ils nagent dans l'eau les cygnes ## on en a vu au bord du lac  
 nous? ## on avait donné du pain*  
 ‘yes they swim in the water ## we have seen them on the  
 lake?## we gave them bread?’
- Child<sub>6</sub>: *oui*  
 %tra: yes
- Mother<sub>7</sub>: *nh? tu te rappelles?*  
 ‘nh? do you remember?’
- Child<sub>7</sub>: *edonné pain cygnes* [turning towards OBS]  
 ‘have given bread swans’
- Mother<sub>8</sub>: *oui on a donné du pain aux cygnes nous hier*  
 ‘yes, we gave bread to the swans yesterday’

Learning a language is a multitask activity that requires the mastery of many different strands, concomitantly and in a very quick manner. Children need to grasp the conventional sounds of the language as well as the meanings attributed to words and to larger constructions. At the same time, they need to keep track of context, shared knowledge, prosodic hints, and past discourse in order to infer their partner's intentions and goals, and to produce utterances that effectively allow the partner to retrieve the child's own intentions and aims.

Conversations with knowledgeable and cooperative partners are very privileged settings where these acquisitions can take place. In (1) above, mother and child start having different aims and interests but after a few skillful turns by the mother, and the cooperative participation of the child, the partners' interests and focus of attention begin to converge, eventually reaching out into the shared recall of past experience.

Eve Clark has been a leading researcher in showing that the construction of common ground is essential to successful communication and can itself be built through conversation. In contrast to nativist accounts that, at best, see language as simply "triggered" by the input they hear, Eve Clark assumes that conversational exchanges offer children adequate information to learn the language. Through detailed analyses of naturalistic observations and ingenious experimental settings that keep the flavor of everyday experience, Eve Clark and her collaborators have shown how caregivers allow children to focus on the relevant information and get hold of the intricacies of the language, both as a system of meanings and rules, and as a means for communication (e.g., Chouinard & Clark, 2009; Clark, 1999; 2010; Clark & deMarneffe, 2012).

This does not mean that children cannot simply learn from speech directed to them, or even from general input, just by overhearing language spoken around them (e.g., Akhtar, Jipson & Callanan, 2001; Gampe, Liebal & Tomasello, 2012). However, conversational exchanges are special experiences as they present several unique and facilitating properties concurrently.

In this chapter we will review the different properties that are specific to conversational exchanges: the motivation to communicate, the reciprocal information that partners can retrieve about each other and that can be updated in real time, the information about language and its use in real settings, as well as the co-construction inherent to their realization. We argue that the convergence of these properties confers to conversational exchanges a privileged status compared to other language experiences that children are exposed to. During the early stages of acquisition, conversations have a great potential to get the process of acquisition going and to facilitate children's first steps into the language system. Empirical evidence, based on results from the literature and from new data analyses, will be provided in support of this position.

## The unique properties of conversational exchanges

### *Motivation to communicate*

Conversation is the natural site of communication. Hence, children are highly motivated to make themselves understood and to understand their partners. Since

they have intentions and meanings to exchange with their partners, in conversation children are particularly likely to attend to their interlocutors' utterances, as well as to the effect that their own interventions have on them. This property fits well with the requirements of functional use approaches that consider language to be best learned in settings where children are *motivated* to use it and where adults talk about objects and events that are of interest to the children (Tomasello, 2003).

### *Information about partners*

Conversation is also the setting where caregivers and children get information about each other in real time, and where the two intertwined components of communication – production and comprehension – are alternatively solicited about the same topics. Caregivers can monitor children's understanding and production abilities, adjust their language to the way children talk, and respond to them. Children have the opportunity to see the effects that their utterances have on their caregivers, and to modify their way of talking as a function of this effect (e.g., Clark, 2009; Gallaway & Richards 1994; Snow & Ferguson 1977; Wales Shugar, 1981).

### *Information about language meaning and structure*

Within this optimal motivational and communicational context, children receive well-timed interpretations of their utterances in the form of expansions, reformulations and corrections, that provide information on vocabulary (phonological shape and semantics) and on morphosyntax, while creating opportunities to produce new, better-adjusted, utterances and to consolidate emerging expressions (e.g., Cazden, 1972; Chouinard & Clark, 2003; Clark & de Marneffe, 2012; Saxton, 2000; Veneziano, 2005, 2009). In the course of conversation, adults also correct children's grammatical errors. Corrections can be explicit, but more commonly they are implicit, as in the case of reformulations (e.g., Chouinard & Clark, 2003; Farrar 1992; Moerk 1991; Saxton, 1997, 2000; Strapp 1999; Veneziano, 2005). Reformulations can be considered *error detectors* (e.g., Chouinard & Clark, 2003). Indeed, contingent corrections are likely to reveal the difference between children's and adults' productions and to offer children ideal opportunities for changing errors into more conventional forms (e.g., Saxton, 2000).

All conversationally contingent responses share the property that changes in form occur while meaning remains invariant. As Brown (1968) and Brown, Cazden and Bellugi (1969) put it, the transformations that utterances undergo "as they shuttle between persons in conversation" may ultimately be one of the most crucial features of conversations, providing the "richest data available to the child" for acquiring language (1969:72).

### *Information about language use*

In conversation, children do not only learn about the language system, its lexicon and morphosyntactic rules, but also how to use linguistic resources to accomplish communicative acts in socially and culturally appropriate ways (e.g., Clark, 2009; Veneziano, 2010). They can learn that requests are made on the assumption that the speaker cannot carry out the object of the demand himself, that refusals should be justified (e.g., Dunn, 1991; Dunn & Munn, 1987; Goetz, 2010; Veneziano, 2001; Veneziano & Sinclair, 1995), or that the speaker's utterances should take into account the degree of shared knowledge and "common ground" among the interlocutors (e.g., Clark, 2004; Clark & Bernicot, 2008; Gundel, Hedberg, & Zacharski, 1993; Matthews, Lieven, Theakston, & Tomasello, 2006).

### *The emergence of new joint events*

Compared to both CDS or input in general, conversations have another unique property that has great potential for the acquisition of new language knowledge. As H.H. Clark puts it, "though conversations are created from individual utterances, they are more than the sum of their parts" (Clark, 1996: 318). They are 'participatory acts' whose individual meaning derives from them being part of a joint action (*ibid.*). As a consequence, events that were not foreseen by the participants at the time they entered the conversation emerge from their joint interactional functioning. This property fits well with a psychogenetic constructivist account of language acquisition based on Piaget's epistemology. One of its central tenets is that the acquisition of new knowledge results from a progressive construction growing out of the possibilities opened up by the functioning of earlier adapted behaviors (Piaget, 1981/1987). Accordingly, behaviors of level  $n + 1$  are considered to build upon behaviors typical of level  $n$ , the current level of functioning (e.g., Piaget & Garcia, 1983/1989), in such a way that the apparent discontinuities of structures result from underlying continuities in the processes of development (Piaget, 1964/1968).

In early language acquisition, the realization of successive single-word utterances, a transitional phenomenon between single-word and two-word utterances, illustrates this property particularly well (Bloom, 1973; Greenfield, Reilly, Leaper & Baker, 1985; Scollon, 1979; Veneziano, Sinclair & Berthoud 1990; Veneziano, 1999, in press):

- (2) Gael (1;10.17) holds a plastic chip that he places in front of a toy cow.  
 Child,: /su/  
 'coin' (money)

Mother<sub>1</sub>: *tu le donnes à qui le sou?*  
 ‘to whom do you give the coin?’

Child<sub>2</sub>: /vaS/  
 ‘cow’

(Veneziano, 2013)

In this particular kind of sequence, the child produces two single word utterances ('coin' and 'cow') that refer to different aspects of one event within a short period of time. The conversational framework within which these two utterances occur allows children to talk about two different aspects of the situation without planning their production beforehand, since the second single-word utterance in the sequence ('cow') is uttered answering a question asked by the mother. In this way, by simply following the turns of the conversation, children enrich their initial focus of attention and pave the way towards the production of multiword utterances (Veneziano, 1999, 2013).

### The effects of conversational exchanges

#### *Immediate effects: Continuing the topic of the conversation*

Does the conversational response of one partner produce further responding on the same topic by the other? This effect was found in some earlier studies (e.g., Folger & Chapman, 1978; Tomasello & Todd, 1983). In new analyses of prior longitudinal recordings, I found additional support for this immediate effect of conversations. The longitudinal recordings came from six Hebrew-acquiring children, followed from the age of 11 to 20 months, recorded once a month in everyday interaction with their mothers (Veneziano, 1984, 1988). Results show that children are three times more likely to stay on topic after their mother's responses to their previous interventions than after the mother's utterances that initiated a topic (see Table 1).

**Table 1.** Mean percentage of on topic response by children at three age periods, according to the type of mother's intervention (number of responded interventions in parentheses).

Age periods	Mean % of Child's response to Mother's interventions	
	Mother's topic initiation	Mother's response to child
11–13 months	10.6% (137)	30.8% (59)
14–15 months	11.5% (182)	44.9% (171)
16–17 months	17.4% (158)	48.1% (175)

At each developmental period, 2x2 contingency tables were constructed for children's on-topic and not-on-topic responses to the two types of turns (mother's initiations and mother's responses to the child), for the six children together. At each of the three age periods, children's on-topic responses were significantly related to mothers' responses to the child (at 11–13 months:  $\chi^2(1, N = 1478) = 57.51, p < .001$ ; at 14–15 months:  $\chi^2(1, N = 1970) = 222.48, p < .001$ ; at 16–17 months:  $\chi^2(1, N = 1285) = 114.38, p < .001$ ). The higher response of children to mothers' turns that are themselves a response to the child, may also be related to the fact that these responses place the word (or word approximation) uttered by the child in a focalized position in the turn. The mother's response may contain only that word, or emphasize it by separating it from the rest of the utterance. In (3) below, the mother's turn contains only the repeated interpretation of the child's word approximation:

- (3) Yuval (1;3) is trying to pull apart a toy drum.

Child<sub>1</sub>: /'a'ma/

Mother<sub>1</sub>: 'ima 'ima  
'mommy, mommy'

Child<sub>2</sub>: /'ima/  
'mommy'

This phenomenon is found also when other languages are involved. Example (4) below concerns a French-speaking mother and her French-acquiring child. The mother utters the word '*pique*' ('sting(s)'), produced previously by the child, in initial position, just before inserting it in a longer utterance that extends the child's meaning:

- (4) Camille (1;7.18) touches the rough spot on the image of a face in a picture book.

Child<sub>1</sub>: /pik/  
'sting(s)'

Mother<sub>1</sub>: *pique! ça pique oui, comme la barbe de papa*  
'sting(s)! it stings yes like father's beard'

Child<sub>2</sub>: /sapik/ 'ça pique'  
'it sting(s)' (Veneziano et al., 1990)

In both examples, the child's utterance in the third turn not only remains on topic but is also more adequate than the initial production.

In the Hebrew study mentioned above, *focalization* was present in 39% of all mothers' turns, but in as many as 76% of the mother's responses that related to the child's previous production (Veneziano, 1988: 136). Children were also more likely

to relate to focalized than to non-focalized words: 26.5% of focalized words and only 14.7% of non-focalized ones were picked up by the children, a statistically significant difference ( $p < .05$ ; Veneziano, 1988: 138–139). Repetition of a lexical item didn't have the same effect, as children were just as likely to pick up words that were repeated in the mother's turn than words that were not repeated (19.7 vs. 19.1%, respectively) (Veneziano, 1988: 139). These results suggest that *focalization* is a particularly effective property of mothers' responses to children's utterances, and is quite likely involved in the topic-continuation effect of the conversations discussed above.

Are mothers affected by the responses of their children? The Hebrew data of the longitudinal study mentioned above were reanalyzed to address this question more precisely. Results show that mothers were also more likely to stay on topic after children's responses to their previous interventions than after children's utterances that initiated a topic (Table 2).

Also here the relation was tested at each developmental period with 2x2 contingency tables presenting mothers' on-topic and not on-topic responding to the two types of children's turns (children's initiations and children's responses to the mother), for the six children together. At each of the three age periods, mothers' on-topic responses were significantly related to children's responses to the mother (at 11–13 months:  $\chi^2(1, N = 1194) = 154.38, p < .001$ ; at 14–15 months:  $\chi^2(1, N = 1671) = 250.61, p < .001$ ; at 16–17 months:  $\chi^2(1, N = 1392) = 134.42, p < .001$ ).

In other words, partners influence each other mutually. When either the mother or the child takes the first step in topic-related responding, joint conversational activity is enhanced, resulting in sequences of topic-related turns. Such extended on-topic conversations provide children with the opportunity to compare their own productions to their mother's, and to then modify their utterances towards a better match to targets (e.g., Chouinard & Clark 2003; Otomo, 2001).

**Table 2.** Mean percentage of on topic response by mothers at three age periods of their children according to the type of child's intervention (number of responded interventions in parentheses).

Age periods	Mean % of Mothers response to Child's interventions	
	Child's topic initiation	Child's response to Mother
11–13 months	10.8% (108)	44.7% (93)
14–15 months	14.7% (194)	53.8% (194)
16–17 months	19.4% (205)	48.8% (173)

### *Immediate effects: Increased matching*

During extended conversational exchanges children modify their verbal productions (e.g., Chouinard & Clark, 2003; Bertin, 2012). Modifications can occur in the phonological shape (as in (3) above) and/or in the semantics of target words, as well as in the grammatical structure of the utterance, as in (4) above and (5) below:

- (5) Sophie (2;4;6)
- |                      |  |
|----------------------|--|
| Adult <sub>1</sub> : | <i>Qu'est-ce qu'elle veut faire Nina'?</i><br>'what does Nina want to do?' |
| Child <sub>1</sub> : | <i>Nina [øjue]</i><br>'Nina [ø]play'                                       |
| Adult <sub>2</sub> : | <i>elle veut aller jouer dehors</i><br>'she wants to go play outside'      |
| Child <sub>2</sub> : | <i>[e]veut jouer dehors</i><br>[e]want play outside                        |
- (Bertin, 2012)

In the longitudinal data of six Hebrew-acquiring children mentioned above, Veneziano (2005) analyzed children's consecutive productions to see whether utterances occurring later in the exchange were more on-target than initial utterances. Excluding sequences where the child's initial production was already on target, it was found that before 14 months children very rarely modified their productions. In contrast, between 14 and 17 months, they improved their utterances on average in 34% of the conversational exchanges that contained two or more child turns (individual children's values ranged between 21% and 50%).

Modifications also occur when the mother does not understand the child's utterance and demands to clarify it with a repair-initiating sequence (Schegloff, 1992). By requesting clarification, caretakers let children know that they need to revise their initial utterance (e.g., Saxton, 2000) and children, not yet 2;0, respond to these requests by changing it or offering the required additional information (Anselmi, Tomasello & Acunzo, 1986; Golinkoff, 1993; Marcos, 1991; Saxton, 2000; Chouinard & Clark, 2003). Modifications occur also when caregivers misinterpret children's intentions (Chouinard & Clark, 2003; Golinkoff, 1986; Marcos, 1991). In the longitudinal study of one child between 1;6 and 2;3, Saxton (2000) noted that, after reformulations and clarification requests, the child corrected the morphological form of verbs more often once the knowledge of that form was emerging, producing it correctly 50% of the time.

Conversational exchanges and, in particular, reformulations and clarification requests, thus have the effect of stimulating more conventional productions from

the children. The extent of the modifications is limited, however, by the children's developmental level and by their language knowledge (Saxton, 2000).

### **Long-term effects of conversational exchanges**

Conversational exchanges also have long term effects. Several studies reported positive correlations between reformulations and recasts by caregivers, and MLU, noun phrase elaboration, grammatical morphemes and number of verb phrases and auxiliaries per utterance (Barnes et al., 1983; Farrar, 1992; Gleitman, Newport & Geltman, 1984; Hoff-Ginsberg, 1985). In training studies, reformulations and recasts have been found to have a positive effect on children's acquisition of the grammatical structures that were recasts, particularly on verbal constructions (Nelson, Carskaddon & Bonvillian, 1973, Nelson, 1977) and on passive structures (Baker & Nelson, 1984). Saxton (1997) found that 20 exposures to corrective reformulations were enough for 5-year-olds to learn the past irregular tense of unfamiliar verbs. However not all studies were able to confirm the positive effect of reformulations on later language acquisition (Cazden, 1972; Ellis & Wells, 1980).

As discussed earlier, conversations do not only involve the caregiver's responses to the child – they also involve the active participation of the children. Their responses to caregivers are part and parcel of language acquisition while providing, at the same time, information about and evidence of children's attention to caregivers' turns. In the early stages of acquisition, children's imitative uptakes of words produced by mothers are positively correlated with their lexical acquisition (Nelson, Baker, Denninger, Bonvillian & Kaplan, 1985; Rodgon & Kurdek, 1977; Snow, 1987). Bloom, Hood & Lightbown (1974) and Ramer (1976) showed that children differ in the extent to which they imitate; however, for all children, the majority of imitations involved words they didn't use spontaneously. In a study of the longitudinal data of the six Hebrew-acquiring children mentioned above, Veneziano (1997, 2005) showed that words imitated by the children, repeated or reformulated by their mother, or both, were produced as conventional word forms at later sessions. Moreover, at later sessions, these productions were recognized as meaningful words by the mother, and solicited comments that advanced the theme two to three times more often than words appearing in that session for the first time.

Slobin (1968) stated that sequences in which children imitate mothers' expansions and reformulations have great relevance for early language acquisition, and Chouinard and Clark (2003) have shown that children's responses to mothers' reformulations are an extremely valuable source of information on the attention that children bring to bear on mothers' reformulations. However, little work has been done to assess the specific impact of extended discourse of this kind compared to

**Table 3.** Correlations between mothers' reformulations at 14 months and children's gains in vocabulary occurring between 14 and 17 months, for three types of mothers' reformulations: (1) total reformulations, (2) reformulations followed by children's response, and (3) reformulations not followed by children's response.

Overall mothers' reformulations	0.794	n.s.
Mothers' reformulations followed by children's response	0.884	<i>p</i> < .05, df = 4
Mothers' reformulations not followed by children's response	0.072	n.s.

discourse where mothers' expansions and reformulations are not followed by children's response to them. In another study of the Hebrew data mentioned above (Veneziano 1997), sequences in which mothers' reformulations (including repeats and expansions) were followed by the children's repetition of the word focused upon by mothers were distinguished from mothers' reformulations that were not followed by children's repeats. For each mother-child dyad, these sequences were identified and measured at 14 months, when the six children had a similar number of words in their repertoire. Three months later, the number of words in the children's repertoire varied, with gains ranging from 11 to 48 words. Results showed that sequences in which children responded imitatively to their mother's reformulations at 14 months correlated positively and significantly with gains in vocabulary measured at 17 months; in contrast, mothers' reformulations that were not followed by children's responses didn't correlate with those gains (Table 3).

### Child-directed speech and conversational exchanges

Caregivers' talk in conversational exchanges is a particular kind of CDS that is closely related to children's focus of attention, to what children say and to how they say it. Both CDS and caregivers' talk in conversation are part of the input. Often input, CDS and caregivers' talk in conversational exchanges are confounded in studies of the effects of the environment on language acquisition. In a recent study of children's early verbs, the specific contributions of CDS and conversational exchanges were assessed (Veneziano & Parisse, 2010). The study was meant to investigate whether caregivers' talk could explain the verb forms children use during the period in which they produce verbs in one form only, a phenomenon found for languages with limited inflectional morphology like English, but also for languages with rich inflectional morphology, such as Hebrew, Turkish and Spanish, as well as moderate inflectional morphology like French (Bassano, 2000; Armon-Lotem & Berman, 2003; Aksu-Koç, 1998; Gathercole, Sebastian & Soto, 1999; Laaha et al., 2007; Tomasello, 1992; Veneziano et al, 1990; Veneziano & Sinclair, 2000). The two

children of the study were acquiring Swiss French and were in the single-form period until 1;11 for the girl and 2;3 for the boy. For example, they used /ka'se/ (to break/broken) for the verb *casser* 'to break' and /tuRn/ (turn/turns) for the verb *tourner* 'to turn'. Authors first analyzed CDS and showed that verb forms that mothers used most frequently explained 64% and 80% of the verb forms used by the two children, respectively. Then, they analyzed the degree of correspondence between the forms used by mothers and children when they were talking about jointly attended events. For about half of the children's verb forms, the information from CDS and conversational exchanges converged: the verb forms used by the children corresponded to the dominant forms used in CDS and were also used by both mothers and children in joint conversational exchanges. Among the remaining children's verb forms, three quarters corresponded exclusively to those used jointly by mothers and children in conversations. Fewer than 15% corresponded exclusively to the dominant forms in CDS. The residue was neither explained by CDS nor by conversation.

This study shows that information present in conversational exchanges can be more powerful than information provided by CDS alone. Moreover, the residue of unexplained verb forms reminds us that children have their own ideas and individual preferences (Shirai, 1998) that can override information contained in both CDS and conversational exchanges.

### Concluding remarks

Conversational exchanges are a particular kind of child-directed speech and a special case of input where various unique properties that have a great potential for language acquisition converge. Children are active participants, motivated to understand and to make themselves understood, and very likely to attend to what is being said to them. Moreover, Western middle class caregivers tune their conversation to the interests of their children, and often relate to what children say and are attending to. At the early stages of language acquisition, conversational exchanges offer children many facilitating features for learning language while they keep pursuing desired activities and goals. Results discussed above show that conversational exchanges have both immediate and long-term effects. They have the effect of stimulating further on-topic conversation by mothers and by children, to solicit modifications of the children's production, and to widen the number of aspects of a situation or intention that can be talked about. They also highlight the importance of joint co-constructed activity: what matters is not so much what each partner does individually but what mothers and children do *together*, the activity of one boosting that of the other, and vice versa. This conclusion is

supported by findings concerning the specific effect of children's response to mothers' reformulations for word learning and for the transition to articulated speech.

Two properties of conversational exchanges stand out as having a particularly great potential for learning. One is the fact that, in conversational exchanges, caregivers interpret, reformulate or expand their children's productions, and children are likely to continue relating to these interpretations and reformulations. In such conversations, meanings are the invariants of transformations in utterance forms, like physical matter, weight or number are the invariants of the transformations in shape. As transformations of shape help children discover physical invariances and the concept of conservation (Piaget & Inhelder, 1941/1974), transformations that utterances undergo 'as they shuttle between persons in conversation' can help children acquire grammar (Brown et al., 1969: 72).

The other important property of conversational exchanges is that they are co-constructed through the turns contributed by each participant. Although participants are in control of their contributions, each turn opens up new possibilities to the speakers, depending on what was said earlier and on the expectations of what the partner may do next. What is accomplished in conversations is thus "more than the sum of the parts" (Clark, 1996: 318). It is a new event that was neither planned beforehand nor foreseen by the participants at the time they entered the conversation. This property allows, for example, the production of successive single-word utterances, where utterances containing only one word are meaningfully related to each other through the unfolding of the conversation (Veneziano, 2013). Successive single-word utterances foreshadow the advent of a new behavior – multitword speech – by practicing well-mastered behaviors typical to the present level of knowledge – single-word speech. This process corresponds to a constructivist principle contained in Piaget's epistemological theory of developmental change (see, e.g., Piaget & Garcia, 1983), according to which discontinuities in the content of knowledge result from underlying continuities in the processes of acquisition: a progressive construction in which earlier acquisitions provide essential releasing steps (see also Gréco, 1985).

Another central tenet of Piaget's psychogenetic constructivism is that the acquisition of knowledge depends, on the one hand, on cognitive mechanisms of acquisition at the subject's disposal, and on the other hand, on the way the object is presented to the subject. Society cannot operate on the former but can modify the latter, and in particular the relation between the subject and the object to be known (e.g., Piaget & Garcia, 1983: 295). Child-directed speech and language not addressed to children both contain relevant information for language learning and children can learn from them. However, conversational exchanges, where many different facilitating properties converge, may be the optimal way in which language can be presented to the learning child, as well as the most powerful resource

in the early stages of language acquisition and at all those critical points where new knowledge is being acquired.

## References

- Akhtar, N., Jipson, J., & Callanan, M. (2001). Learning words through overhearing. *Child Development*, 72, 416–430. DOI: 10.1111/1467-8624.00287
- Aksu-Koç, A. (1998). The role of input vs. universal predispositions in the emergence of tense-aspect morphology: Evidence from Turkish. *First Language*, 18, 255–280. DOI: 10.1177/014272379801805402
- Anselmi, D., Tomasello, M., & Acunzo, M. (1986). Young children's responses to neutral and specific contingent queries. *Journal of Child Language*, 13, 135–144. DOI: 10.1017/S030500000349
- Armon-Lotem, S., & Berman, R.A. (2003). The emergence of grammar: Early verbs and beyond. *Journal of Child Language*, 30, 845–877. DOI: 10.1017/S0305000903005750
- Baker, N.D., & Nelson, K.E. (1984). Recasting and related conversational techniques for triggering syntactic advances in young children. *First Language*, 5, 3–22. DOI: 10.1177/014272378400501301
- Barnes, S., Gutfreund, M., Satterly, D., & Wells, G. (1983). Characteristics of adult speech which predict children's language development. *Journal of Child Language*, 10, 65–84. DOI: 10.1017/S0305000900005146
- Bassano, D. (2000). Early development of nouns and verbs in French: Exploring the interface between lexicon and grammar. *Journal of Child Language*, 27, 521–559 DOI: 10.1017/S0305000900004396
- Bertin, T. (2012). Rôle des reprises dans l'acquisition de la morpho-syntaxe chez des enfants francophones âgés de 2–3 ans. 12e Colloque de Logopédie “La parole reprise : formes, processus et fonctions”, Neuchâtel, Suisse.
- Bloom, L. (1973). *One Word at a Time*. The Hague: Mouton.
- Bloom, L., Hood, L., & Lightbown, P. (1974). Imitation in language development: If, when and why. *Cognitive Psychology*, 6, 380–420. DOI: 10.1016/0010-0285(74)90018-8
- Brown, R. (1968). The development of WH questions in child speech. *Journal of Verbal Learning & Verbal Behavior*, 7, 279–290. DOI: 10.1016/S0022-5371(68)80002-7
- Brown, R., Cazden, C.B., & Bellugi, U. (1969). The child's grammar from I to II. In J.P. Hill (Ed.), *Minnesota symposia in child psychology, Vol. II* (pp. 28–73). Minneapolis, MN: University of Minnesota Press.
- Cazden, C.B. (1972). *Child language and education*. New York, NY: Holt, Rinehart & Winston.
- Chouinard, M.M., & Clark, E.V. (2003). Adult reformulations of child errors as negative evidence. *Journal of Child Language*, 30, 637–669. DOI: 10.1017/S0305000903005701
- Clark, E.V. (1999). Acquisition in the course of conversation. *Studies in the Linguistic Sciences (Forum Lectures from the 1999 Linguistic Institute)* 29(2), 1–18.
- Clark, E.V. (2004). Pragmatics and language acquisition. In L.R. Horn & G. Ward (Eds.) *Handbook of pragmatics* (pp. 562–577). Oxford: Blackwell.
- Clark, E.V. (2009). *First language acquisition, 2nd ed.* Cambridge: CUP. DOI: 10.1017/CBO9780511806698

- Clark, E.V., & Bernicot, J. (2008). Repetition as ratification: How parents and children place information in common ground. *Journal of Child Language*, 35, 349–371. DOI: 10.1017/S0305000907008537
- Clark, E.V., & de Marneffe, M.-C. (2012). Constructing verb paradigms in French: Adult construals and emerging grammatical contrasts. *Morphology*, 22, 89–120. DOI: 10.1007/s11525-011-9193-6
- Clark, H.H. (1996). *Using Language*. Cambridge: CUP. DOI: 10.1017/CBO9780511620539
- Dunn, J. (1991). Understanding others: Evidence from naturalistic studies of children. In A. Whiten (Ed.), *Natural Theories of Mind* (pp. 51–61). Oxford: Basil Blackwell.
- Dunn, J., & Munn, P. (1987). Development of justification in disputes with mother and sibling. *Developmental Psychology*, 23, 791–798. DOI: 10.1037/0012-1649.23.6.791
- Ellis, R., & Wells, G. (1980). Enabling factors in child-adult discourse. *First Language*, 1, 46–62. DOI: 10.1177/014272378000100104
- Farrar, M.J. (1992). Negative evidence and grammatical morpheme acquisition. *Developmental Psychology*, 28, 90–98. DOI: 10.1037/0012-1649.28.1.90
- Folger, J.P., & Chapman, R.S. (1978). A pragmatic analysis of spontaneous imitations. *Journal of Child Language*, 5, 25–38. DOI: 10.1017/S0305000900001914
- Gallaway, C., & Richards, B.J. (Eds.) (1994). *Input and interaction in language acquisition*. Cambridge: CUP. DOI: 10.1017/CBO9780511620690
- Gampe, A., Liebal, K., & Tomasello, M. (2012). Eighteen-month-olds learn novel words through overhearing. *First Language*, 32, 385–397. DOI: 10.1177/0142723711433584
- Gathercole, V.C.M., Sebastian, E. & Soto, P. (1999). The early acquisition of Spanish verbal morphology: Across-the-board or piecemeal knowledge? *International Journal of Bilingualism*, 3, 133–182. DOI: 10.1177/13670069990030020401
- Gleitman, L.R., Newport, E.L., & Gleitman, H. (1984). The current status of the motherese hypothesis. *Journal of Child Language*, 11, 43–79. DOI: 10.1017/S0305000900005584
- Goetz, P.J. (2010). The development of verbal justifications in the conversations of preschool children and adults. *First Language*, 30, 403–420. DOI: 10.1177/0142723710370522
- Golinkoff, R. (1986). ‘I beg your pardon?’: The preverbal negotiation of failed messages. *Journal of Child Language*, 13, 455–476. DOI: 10.1017/S0305000900006826
- Gréco, P. (1985). Réduction et construction. *Archives de Psychologie*, 53, 21–35.
- Greenfield, P., Reilly, J., Leaper, C., & Baker, N. (1985). The structural and functional status of single-word utterances and their relationship to early multi-word speech. In M.D. Barrett (Ed.), *Children’s Single-Word Speech* (pp. 233–267). Chichester: Wiley & Sons.
- Gundel, J.K., Hedberg, N., & Zacharski, R. (1993). Cognitive status and the form of referring expressions in discourse. *Language*, 69, 274–307. DOI: 10.2307/416535
- Hoff-Ginsberg, E. (1985). Some contributions of mothers’ speech to their children’s syntactic growth. *Journal of Child Language*, 12, 367–385. DOI: 10.1017/S0305000900006486
- Marcos, H. (1991). Reformulating requests at 18 months: Gestures, vocalisations and words. *First Language*, 11, 361–375. DOI: 10.1177/014272379101103304
- Matthews, D., Lieven, E., Theakston, A., & Tomasello, M. (2006). The effect of perceptual availability and prior discourse on young children’s use of referring expressions. *Applied Psycholinguistics*, 27, 403–422. DOI: 10.1017/S0142716406060334
- Moerk, E.L. (1991). Positive evidence for negative evidence. *First Language*, 11, 219–251. DOI: 10.1177/014272379101103202

- Laaha, S., Gillis, S., Kilani-Schoch, M., Korecky-Kröll, K., Xanthos, A., & Dressler, W.U. (2007). Weakly inflecting languages: French, Dutch, and German. In S. Laaha & S. Gillis (Eds.), *Typological perspectives on the acquisition of noun and verb morphology* (pp. 21–33). Antwerp: University of Antwerp.
- Nelson, K.E. (1977). Facilitating children's syntax acquisition. *Developmental Psychology*, 13, 101–107. DOI: 10.1037/0012-1649.13.2.101
- Nelson, K.E., Baker, N.D., Denninger, M., Bonvillian, J.D. & Kaplan, B.J. (1985). Cookie versus do-it-again: Imitative-referential and personal-social syntactic-initiating language styles in young children. *Linguistics*, 23, 433–454. DOI: 10.1515/ling.1985.277.1.433
- Nelson, K.E., Carskaddon, G., & Bonvillian, J. (1973). Syntax acquisition: Impact of experimental variation in adult verbal interaction with the child. *Child Development*, 44, 497–504. DOI: 10.2307/1128005
- Otomo, K. (2001). Maternal responses to word approximations in Japanese children's transition to language. *Journal of Child Language*, 28, 29–57. DOI: 10.1017/S0305000900004578
- Piaget, J. (1964/1968). Genèse et structure en psychologie de l'intelligence. In D. Elkind (Ed.), *Six études de psychologie* (pp. 164–181). Genève: Gonthier (English Edition: Six Psychological Studies. New York, NY: Vintage Books, 1968).
- Piaget, J. (1981/1987). *Le possible et le nécessaire*, Vol. 1: *L'évolution des possibles chez l'enfant*. Paris, France: P.U.F. (English Edition: *Possibility and necessity*, Vol.1: *The role of possibility in cognitive development*. Minneapolis, MN: University of Minnesota Press, 1987).
- Piaget, J., & Garcia, R. (1983/1989). *Psychogenèse et histoire des sciences*. Paris: Flammarion. (English *Psychogenesis and the history of science*. New York, NY: Columbia University Press, 1989).
- Ramer, A. (1976). The function of imitation in child language. *Journal of Speech and Hearing Research*, 19, 700–717.
- Rodgon, M., & Kurdek, L. (1977). Vocal and gestural imitation in children under 2 years old. *Journal of Genetic Psychology*, 131, 115–123. DOI: 10.1080/00221325.1977.10533280
- Saxton, M. (1997). The contrast theory of negative input. *Journal of Child Language*, 24, 139–161. DOI: 10.1017/S030500099600298X
- Saxton, M. (2000). Negative evidence and negative feedback: Immediate effects on the grammaticality of child speech. *First Language*, 20, 221–252. DOI: 10.1177/014272370002006001
- Schegloff, E.A. (1992). Repair after next turn: The last structurally provided defence of intersubjectivity in conversation. *American Journal of Sociology*, 97, 1295–1345. DOI: 10.1086/229903
- Scollon, R. (1979). A real early stage: An unzipped condensation of a dissertation on child language. In E. Ochs & B.B. Schieffelin (Eds.), *Developmental pragmatics* (pp. 215–277). New York, NY: Academic Press.
- Slobin, D.I. (1968). Imitation and grammatical development in children. In N. Endler, L. Boulter, & H. Osler (Eds.) *Contemporary issues in developmental psychology* (pp. 437–443). New York, NY: Holt, Rinehart & Winston.
- Snow, C.E. (1987). Imitativeness: A trait or a skill? In G. Speidel & K.E. Nelson (Eds.), *The many faces of imitation in language learning* (pp. 73–90). New York NY: Springer.
- Snow, C.E., & Ferguson, C.A. (Eds.) (1977). *Talking to children: Language input and interaction*. Cambridge, UK: Cambridge University Press.

- Strapp, C.M. (1999). Mothers', fathers', and siblings' responses to children's language errors: Comparing sources of negative evidence. *Journal of Child Language*, 26, 373–391. DOI: 10.1017/S0305000999003827
- Tomasello, M. (2003). *Constructing a language: A usage-based theory of language acquisition*. Cambridge, MA: Harvard University Press.
- Tomasello, M., & Todd, J. (1983). Joint attention and lexical acquisition style. *First Language*, 4, 197–212. DOI: 10.1177/014272378300401202
- Veneziano, E. (1984). *Structural and formal precursors of language: An interactional approach*. Unpublished doctoral dissertation. The Hebrew University, Jerusalem.
- Veneziano, E. (1988). Vocal-verbal interaction and the construction of early lexical knowledge. In M.D. Smith & J.L. Locke (Eds.) *The emergent lexicon: The child's development of a linguistic vocabulary* (pp. 109–147). New York, NY: Academic Press.
- Veneziano, E. (1997). Echanges conversationnels et acquisition première du langage. In J. Bernicot, A. Trognon & J. Caron-Pargue (Eds.), *Conversation, interaction et fonctionnement cognitif* (pp. 91–123). Nancy: Presses universitaires de Nancy.
- Veneziano, E. (1999). Early lexical, morphological and syntactic development in French: Some complex relations. *International Journal of Bilingualism*, 3, 183–217. DOI: 10.1177/13670069990030020501
- Veneziano, E. (2001). Interactional processes in the origins of the explaining capacity. In K. Nelson, A. Aksu-Koç, & C. Johnson (Eds.), *Children's language*, Vol. 10 (pp. 113–141). Mahwah NJ: Lawrence Erlbaum Associates.
- Veneziano, E. (2005). Effects of conversational functioning on early language acquisition: When both caregivers and children matter. In B. Bokus (Ed.), *Studies in the Psychology of Child Language* (pp. 47–69). Warsaw, Poland: Matrix.
- Veneziano, E. (2009). Processus inter et intra dans l'acquisition et le fonctionnement du langage. *Cahiers du SRED*, 15, 407–414.
- Veneziano, E. (2010). Conversation in language development and use: An introduction. *First Language*, 30(3–4), 241–249. DOI: 10.1177/0142723710380531
- Veneziano, E. (2013). A cognitive-pragmatic model for the change from single-word to multi-word speech: A constructivist approach. *Journal of Pragmatics*, 56, 133–150. DOI: 10.1016/j.pragma.2013.03.013
- Veneziano, E., & Parisse, C. (2010). The acquisition of early verbs in French: Assessing the role of conversation and of child-directed speech. *First Language*, 30, 287–311. DOI: 10.1177/0142723710379785
- Veneziano, E., & Sinclair, H. (1995). Functional changes in early child language: The appearance of references to the past and of explanations. *Journal of Child Language*, 22, 557–581. DOI: 10.1017/S0305000900009958
- Veneziano, E., & Sinclair, H. (2000). The changing status of "filler syllables" on the way to grammatical morphemes. *Journal of Child Language*, 27, 1–40. DOI: 10.1017/S03050009000427X
- Veneziano, E., Sinclair, H., & Berthoud, I. (1990). From one word to two words: Repetition patterns on the way to structured speech. *Journal of Child Language*, 17, 633–650. DOI: 10.1017/S0305000900010928
- Wales Shugar, G. (1981). Early child discourse analyzed in the dyadic interaction unit. *International Journal of Psycholinguistics* 8(2), 55–78.



# Taking the floor on time

## Delay and deferral in children's turn taking

Marisa Casillas

Max Planck Institute for Psycholinguistics

A key part of learning to speak with others is figuring out when to start talking and how to hold the floor in conversation. For young children, the challenge of planning a linguistic response can slow down their response latencies, making misunderstanding, repair, and loss of the floor more likely. Like adults, children can mitigate their delays by using fillers (e.g., *uh* and *um*) at the start of their turns. In this chapter I analyze the onset and development of fillers in five children's spontaneous speech from ages 1;6–3;6. My findings suggest that children start using fillers by 2;0, and use them to effectively mitigate delay in making a response.

Children's linguistic input is a fundamental component of the language learning process – what children hear is essential to what they can learn. All theories of language acquisition must contend with the question of how children extract relevant information from the linguistic data around them. In the countless discussions that have taken place about children's input, researchers have focused their attention on domain-specific mechanisms, domain-general mechanisms, and a little bit of everything in between to explain how children manage to glean relevant details from the sum of available information.

In much of her work, Eve Clark has diverged from theory-driven perspectives on the internal mechanisms of language learning, encouraging us to instead reflect on the depth of children's linguistic environments. She has helped us understand that a child's linguistic upbringing is more than the sum of its words and syntactic structures. By taking a data-driven and usage-based approach to studying language, she has focused our attention on a vital truth: children learn language in interaction with others, and they learn language for the purpose of interacting with others. Consequently, children's input incorporates not only what they hear, but also what they say in response. I take up her insight in this chapter by focusing on turn taking, one key skill children must develop for interaction.

## Introduction

From the start, children have a major impact on the speech they hear from their caregivers.<sup>1</sup> Caregivers work hard to get and maintain their children's attention, and one of the most effective techniques they have for doing so is focusing their conversation on things that are relevant to the child in the here and now (Bloom, 1975; Bloom, 1984; Snow, 1977). Children as young as 0;3 respond contingently to caregivers' prompts with smiles and vocalizations which caregivers, in turn, treat as relevant to the ongoing interaction (Balog & Roberts, 2004; Bloom, 1988; Masataka, 1993).

Once they begin to speak, children exhibit further influence on the content of conversations by introducing new information and commenting on topics that had been brought to bear previously. Children begin to initiate conversational exchanges around their second birthdays and by age three they initiate up to two-thirds of exchanges (Clark, 2009; Wellman & Lempers, 1977). By 3;6, children fluently respond to others by both ratifying the prior utterance and adding new information to move the conversation forward (Clark, 2007; Clark & Bernicot, 2008; see also Saylor, Baird, & Gallerani, 2006). Children also adapt their conversational contributions, depending on who they are talking to. Because caregivers and other interlocutors variably accommodate to younger speakers, children adjust their conversational strategies from one conversation to another in order to proceed smoothly (Dunn & Shatz, 1989; Mannle & Tomasello, 1987). So not long after children begin to talk, they also begin to actively influence the language they hear and use.

In light of this, we can view the child's linguistic input as an ongoing function of the interaction, rather than an end result of it. Caregivers' contributions are designed for children's current interests and abilities (e.g., Roy, Frank, & Roy, 2009; Tomasello & Farrar, 1986; Uther, Knoll, & Burnham, 2007; also see Weisleder & Fernald, this volume) and are embedded in the immediate context shared between caregiver and child. Children can actively engage in conversation with others by taking turns at speaking, thereby influencing the input they receive. By this token, children's turn taking skills have far-reaching ramifications for the input they get

---

1. I focus here on available research concerning children's conversational skills. Unfortunately, this makes for a highly western-centric analysis. Cross-cultural data on children's language practice is critical to evaluating theories of acquisition, especially those rooted in specific formats for input. Turn taking, my topic here, shows striking cross-cultural similarity in adult timing patterns (Stivers et al., 2009), and so might be expected to show similar developmental trajectories cross-culturally. However, more diverse data on children's turn taking will be essential in assessing how general these effects are.

because a chance to take the floor is a chance to learn something new, test a hypothesis, and get feedback.

### Turn timing

When adults take conversational turns, there is usually minimal inter-speaker gap and minimal vocal overlap at the points of turn exchange. This seems to be a nearly universal property of spontaneous discourse (Stivers et al., 2009). While children's basic turn taking skills emerge in infancy (Masataka, 1993; Snow, 1977), their turn timing continues developing toward adult norms well after they turn four (Ervin-Tripp, 1979; Garvey & Berninger, 1981). On average, a four-year-old English-speaker's response latency is at least twice the norm for adults (Garvey & Berninger, 1981). Traditionally, turn taking is viewed as a competitive endeavor, during which speakers are held accountable for their lapses in timing.<sup>2</sup> The lack of an immediate response can result in a lost opportunity to take the floor (Sacks, Schegloff, & Jefferson, 1974), so children's slower timing can affect their ability to get a word in, depending on who else is participating in the conversation.

The number of speakers and their adjustment to the child's needs affect how easy it is to take the floor on time. Dyadic conversation (with exactly two speakers) simplifies the allocation of turns because there are fewer options for turn-exchange: selection of the other speaker or self-selection. When asking a question in dyadic conversation, there is only one speaker to whom talk could be addressed, so it may be easier for the responder to figure out when to begin to speak. Also, in making a response during dyadic conversation, a child does not have to compete with any other speakers while taking the floor.

In multi-party conversation (with three or more speakers), children need to work harder to take their turns. Young children at school and children with siblings commonly find themselves in multi-party conversations. Children must work harder in these conversations because, when turn-exchange is initiated, they have to figure out which of the two or more non-current speakers are supposed to speak next. When these children find a chance to speak, they often compete for the floor with speakers who have quicker turn-timing skills. And then when they manage to get a word in, they are likely to overlap with another person's speech, which could ultimately result in a misunderstanding or the appearance of irrelevance (Dunn & Shatz, 1989; Lieven, 1978). For these reasons, multi-party conversations can be a challenge for young children, but those who practice conversation

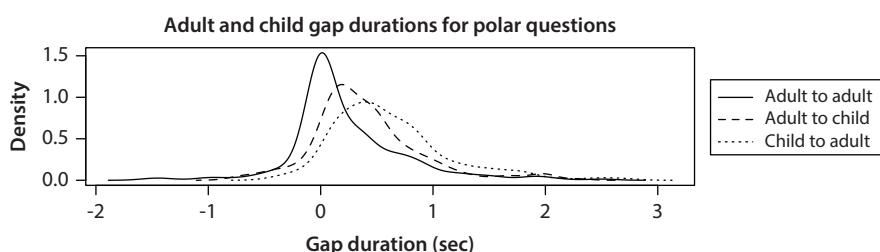
---

2. This holds for children too, but with ample slack (Balog & Roberts, 2004).

under challenging conditions (e.g., younger siblings) can hone their skills early on (Dunn & Shatz, 1989).

Children's success at turn taking also depends on how accommodating their interlocutors are. At age three, children's response latencies in conversation with their peers (Garvey & Berninger, 1981) are substantially longer than while speaking to their caregivers (Casillas, Bobb, & Clark, under review). This could be due to the fact that caregivers often design their utterances to elicit specific kinds of information; for example, caregivers frequently ask questions to which they already know the answers (Fitneva, 2012). In speaking to their children, caregivers appear to slow down their response latencies, responding to questions more slowly than they would if they were speaking to another adult (Figure 1; Casillas et al., under review). When passing the floor to their children, caregivers wait much longer than they would with adults before repeating or reformulating their questions. In a collection of nearly 930 questions directed to children, 10% of children's answers took more than 1.5 seconds to begin (Casillas et al., under review). The children were not interrupted or re-questioned by the caregiver, even though the child's response latency exceeded *six times* the adult average for question-answer sequences (Casillas et al., under review; Stivers et al., 2009; see also Balog & Roberts, 2004).

Many early interactions take place between a young child and a single caregiver. These one-on-one conversations allow even infants to appear conversational (Bloom, 1975; Bloom, 1984; Snow, 1977). But as children get older, their ability to take and hold the floor is tested. In competitive situations, children lose the floor if they do not take it on time, and they may be required to reserve their place before they have managed to prepare a response. So what do children do when it is their turn to speak and they are not quite ready to start?



**Figure 1.** Distribution of response times to yes-no questions for adult responders to adult questioners (—), adult responders to child questioners (—), and for child responders to adult questioners (· · ·).

### Marking delays in children's conversation

When adults expect a delay in their ability to take the floor, they use verbal and non-verbal signals such as *uh*, *um*, and facial expressions to indicate an upcoming pause. Clark and FoxTree (2002) found that speakers use *uh* and *um*<sup>3</sup> to indicate the initiation of short and long delays, respectively. They argue that these delay markers are contrasting lexical items used when speakers anticipate noteworthy delays in their speech.<sup>4</sup> Using evidence from prosody and syllabification, they argue that *uh* and *um* are planned parts of the speech signal. Word prolongation (e.g., *theee, ummm*) also initiated longer pauses when combined with *uh* and *um*, but the effect was independent, and the choice of *uh* or *um* had a much stronger role in predicting delay prolongation.

Children who are not able to respond in a timely manner could similarly use delay markers to their advantage when trying to get a word in during conversation, indicating the length of the upcoming delay while allowing them to keep the floor. To compensate for an inability to jump in at just the right time, young children may initially use physical or verbal intrusions to get a word in, such as tugging at a sleeve or tapping a shoulder repeatedly. Prior work has shown that, by four years, English-speaking children productively use *uh* and *um* when retelling stories and describing their toys, but don't use *uh* and *um* to indicate upcoming delay length until age five (Hudson Kam & Edwards, 2008). Even then, they only systematically us the contrast in describing toys, and not when telling a story (Van Der Wege & Ragatz, 2004).

The late acquisition of these delay markers is likely due to their complexity; to anticipate noteworthy delays and to determine if they will be long or short, children must develop sensitive self-monitoring skills and an awareness of conversational timing. But even before children work out the contrastive use of *uh* and *um*, these delay markers can serve children's turn taking. When used utterance-initially they help children get the floor, and when used utterance-internally they help children keep the floor.

Prior work on children's acquisition of *uh* and *um* has primarily relied on elicited speech in a restricted setting (Hudson Kam & Edwards, 2008; Van Der Wege & Ragatz, 2004). To get at children's spontaneous use of delay markers during everyday speech, we need to look at interaction in the home. Here I present data from a small sample of at-home interactions. These data do not represent a comprehensive study, but give some idea of trends in children's everyday delay marking. I find that children begin to use delay markers by age two, that children use a wide

3. Sometimes written *er* and *erm* in British English.

4. Contrasting delay marker pairs are also noted in German, Dutch, Swedish, Norwegian, Spanish, French, Hebrew, and Japanese (Clark & Fox Tree, 2002).

variety of conventional (e.g., *uh*, *um*, and prolongation) and unconventional (e.g., fronting) strategies, and that children who use conventional delay markers most often prefer to use *um* over *uh*. I also find initial patterns suggesting that, before 4;0, children may begin to differentiate their use of *uh* and *um* in the same way adults do, using *um* before longer delays and more disfluent speech.

### *Delay marking at home*

I extracted all of the transcribed instances<sup>5</sup> of *uh* and *um* from the Providence corpus of the CHILDES database (Demuth, Culbertson, & Alter, 2006; MacWhinney 2012). Providence includes a collection of at-home video and audio recordings of six children (*Female* = 3) between one and four years old. The recordings were made approximately every two weeks for about one hour each, though each child had a slightly different age range and recording frequency. This corpus is one of the densest longitudinal corpora in the CHILDES database and provides good enough audio for phonetic measurement of any pauses in speech.

There are 1,945 transcribed *uh* and *um* tokens in the corpus (1,354 *ums* and 591 *uhs*).<sup>6</sup> Overall, children showed an *um* bias, with 2.29 *um* tokens for every 1 *uh* token (but also see footnote 5). In general, *uh* and *um* are infrequent before children turn two years old. *Um* emerges as the overall more frequent delay marker from early on (Figure 2), but the frequency of use for each delay marker varied widely by child (Table 1). Half of the children actually show an *uh* bias – Ethan uses *uh* twice as often as *um* – but these children tend to have fewer *uhs* and *ums* overall (Table 1).

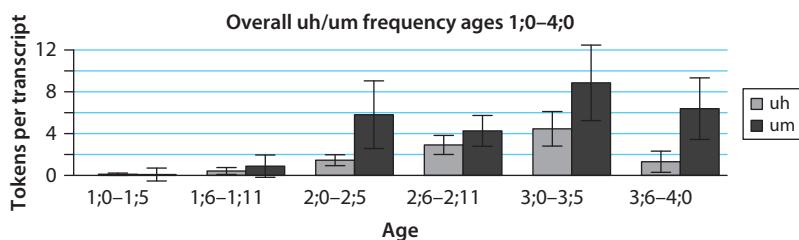


Figure 2. Use of *uh* and *um* from 1;0–4;0 collapsed over all six children. Vertical bars indicate the standard error of the mean. Each recording was approximately one hour long.

5. By carefully reviewing two randomly selected recordings for each of the six children, I determined that the transcriptions missed ~10% of the total uses of *uh* and *um* on average. Nearly two thirds of these misses were *uh* (about 1.25 missed tokens per recording) meaning that in the analyses I present below, *uh* tokens are slightly underrepresented.

6. I extracted tokens from 208 transcripts: Ethan (0;11–2;11, 19 recordings), Naima (1;0–3;10, 25 recordings), Alex (1;5–3;5, 32 recordings), Violet (1;2–4;0, 31 recordings), William (1;4–3;4, 30 recordings), and Lily (1;1–4;0, 71 recordings).

**Table 1.** Average delay marker use per transcript for each child in the Providence corpus, and the dominant option used by each child. Percent use of the dominant delay marker is given in parentheses.

Child (gender)	<i>Uhs + ums</i> per transcript	Preferred DM (and percent use)
Ethan (M)	1.08	<i>uh</i> (67%)
Naima (F)	1.08	<i>uh</i> (67%)
Violet (F)	3.94	<i>um</i> (55%)
Will (M)	4.38	<i>uh</i> (73%)
Alex (M)	5.23	<i>um</i> (77%)
Lily (F)	19.15	<i>um</i> (83%)

What is causing the wide dispersion in children's use of *uh* and *um*? In a separate study using two children's data from the same corpus (Providence: Demuth et al., 2006; see Casillas et al., under review), I analyzed children's turn-initial disfluency markers in a set of 1,280 question-answer sequences from ages 1;6-3;6. Of these, 142 had turn-initial delay markers; but only 49 of those were *uh* or *um*. The other 93 instances began with alternative delay markers, including repetition, recycling, fronting, and prolongation (see example 1). The two children used nearly an equal amount of delay markers overall, but expressed a consistent preference for one type over another throughout the corpus: Alex primarily used *uh* and *um* to indicate upcoming delays (40 *uh/um* and 25 other DMs), while Naima preferred other markers (9 *uh/um* and 68 other DMs). Thus, children who appear to use *uh* and *um* infrequently (as in Naima; Table 1) may rely more heavily on other delay-marking processes. Such as those in example 1.

- (1) Other types of delay markers in child data
  - a. *Repetition*: but but but but we could sing some other songs (nai71; 3;1)
  - b. *Restarts*: liz- lizards nurse (nai36; 1;10)
  - c. *Fronting*: wool cut her wool (nai37; 1;11)
  - d. *Prolongation*: uh: just brown (ale44; 3;1)
  - e. *Other*: name, jelly beans (nai26; 1;8)

When children use delay markers turn-initially, it should enable them to start speaking without delay, even when there is a hitch in planning their response. In the same study of question-answer sequences (Casillas et al., under review), I found that utterances with response-initial markers were significantly longer and more complex than utterances without. But the complexity of the marker-initial response did *not* result in slower turn onset. In fact there was no significant timing difference between utterances with and without markers ( $p = 0.14$ ), so children's

use of delay markers keeps gap lengths similar across more and less difficult responses. Because children use delay markers effectively we can infer that they (1) anticipated a noteworthy delay and (2) took action to mitigate the upcoming latency. Adults do this, but additionally differentiate short and long delays with *uh* and *um*, respectively. The prior work from elicited speech did not find such differentiation in children's use of delay markers before age five (Van Der Wege & Ragatz, 2004; Hudson Kam & Edwards, 2008), but conversational contexts may provide greater communicative need to indicate the timing of upcoming delays.

To determine whether or not children also differentiate *uh* and *um*, I randomly selected 40 *uh* and 40 *um* tokens from the last year of the corpus (3;0–4;0). I then extracted the audio clip for each *uh/um* utterance, and measured the pause between the end of the marker and the beginning of the next segment of speech, which was not necessarily the beginning of the rest of the proposition (example 2 below). I bounded my measurement by the next segment of speech, assuming that a subsequent instance of *uh* or *um* is an additionally planned production that initiates a new delay (Clark & FoxTree, 2002).

- (2) The following pause: Measured portions are marked with an “\_”
  - a. *Fluent speech*: and uh\_what about these? (wil41; 3;2)
  - b. *Another marker*: and uh\_um a worm (ale51; 3;4)
  - c. *Disfluency*: can you um\_can you put me can you put me right there? (lil79; 3;10)

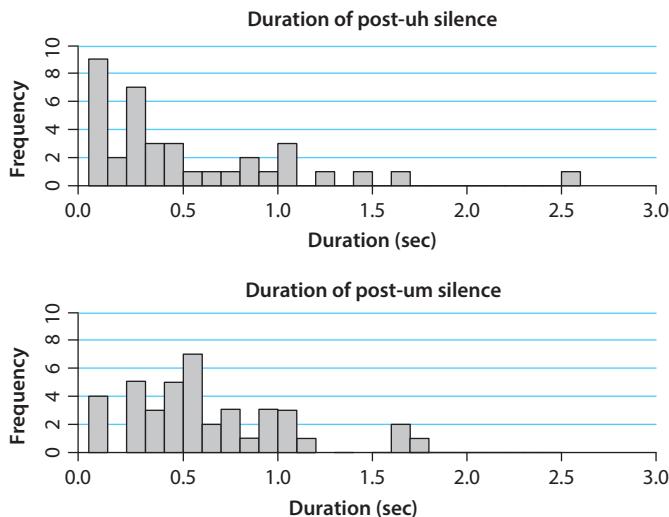
The results indicate that pauses following *uh* and *um* mirrored the speech of adults and older children: *um* initiated longer pauses than *uh* (median durations of 0.54 sec and 0.38 sec; Figure 3). The difference is not significant ( $p = 0.3$ ),<sup>7</sup> following prior work (Hudson Kam & Edwards, 2008; Van Der Wege & Ragatz, 2004). However, the trend may suggest that children are just beginning to use *um* to initiate longer delays and *uh* to initiate shorter ones.

These data alone do not provide enough evidence that children use *uh* and *um* to contrastively mark delay before age four. But by comparing these effects to other indirect measures of planning difficulty (e.g., word prolongation and speech fluency), we can get a more comprehensive picture of delay-based contrasts in children's speech. If we see consistency across different measures, it would indicate a stronger pattern in how children mark short vs. long delays.

Using the same 80 measurements as before, I coded whether or not the delay marker was prolonged (e.g., *uh* vs. *uhhh*) and what type of speech followed the

---

7. I entered the data into a linear mixed-effects model, with child as random effect and filler type and prolongation as fixed effects.



**Figure 3.** Histogram of silence duration following *uh* and *um* tokens. Phonetic measurements were made from the offset of the DM to the onset of the next bit of speech (See example 2).

**Table 2.** Median values (in seconds) for delay durations following regular and prolonged *uh* and *um*.

Delay Marker	Regular	Prolonged
<i>uh</i>	0.29	0.40
<i>um</i>	0.45	0.77

pause (fluent, disfluent, or other). Overall, *uh* was more likely to be prolonged than *um* (57.5% vs. 22.5%). For both words, prolongation indicated longer upcoming delays (Table 2). The effect of prolongation is not significant ( $p = 0.3$ ), but the pattern showed that, for both regular and prolonged tokens, delays following *um* are longer than delays following *uh*.

In most cases (63.8%), speech following the use of *uh* and *um* was a fluent continuation of the child's turn. But almost one-third of the time (31.3%), speech following *uh* and *um* started with another delay marker, recycled speech, or continued disfluency. The fluency of post-marker speech varied with marker type. *Uh* was more likely than *um* to be followed by fluent speech (72.5% vs. 55%), and slightly less likely than *um* to be followed by an interruption from the caregiver (2.5% vs. 7.5%; Table 3). This secondary source of evidence indirectly suggests that children use *um* more often than *uh* when their upcoming speech presents significant difficulty.

**Table 3.** Types of speech following children's use of *uh* and *um*.

Following Speech	<i>Uh</i>	<i>Um</i>
Fluent	72.5% (29)	55% (22)
Continue	67.5% (27)	52.5% (21)
Move on	5% (2)	2.5% (1)
Disfluent	25% (10)	37.5% (15)
New Disfluent	7.5% (3)	17.5% (7)
Delay marker	15% (6)	12.5% (5)
Recycle	2.5% (1)	7.5% (3)
Other-interruption	2.5% (1)	7.5% (3)

Taken together these data suggest a few things about children's development of delay markers. First, they emerge early in children's speech. *Uh* and *um* start regularly appearing by the time children turn two with use of both the regular and prolonged versions. *Um* emerges early on as the most frequent option, especially for the children who are most productive with *uh* and *um*. Before children turn four, their use of *uh* and *um* begins to pattern like adult data because their prolongation and choice of marker may indicate the durations of upcoming delays, though we would need more data to say if these effects are significant. Speech following *uh* is more fluent than speech following *um*, which suggests that children are facing less significant planning delays when they use *uh* compared to when they use *um*.

Strong evidence for discrimination between *uh* and *um* is minimized in a small sample like this one, but the secondary measurements of pause duration in prolonged tokens and post-pause speech fluency both suggest that before age 4;0 children have developed some differentiation in their use of *uh* and *um*. Related work suggests that children's use of *uh* and *um* helps them come in quickly when they're planning a complex contribution (Casillas et al., under review), and that when they're not using *uh* and *um*, children are likely using other ways of marking delay. These alternate delay markers may even emerge earlier in development because of their simpler execution (e.g., prolongation and repetition).

Once children begin to have an active role in conversation, their ability to get and keep the floor becomes critical for making controlled contributions. Even in contexts where turn taking isn't overtly competitive, it is to children's advantage to claim the floor as an indication of their intention to respond. If they don't, they run the risk of breaking the ongoing interaction, and may cause the conversation to veer in a new direction (e.g., toward an attempt at repair). At that point, the child's delayed response might no longer be relevant, or the child may need to come up with a new response. This may be cause for some bumpy interaction, e.g.:

## (3) Re-prompting and repair (nai64; 2;7.13)

- MOT: How many? (points to pencil cars; no response from child)
- MOT: What're these? (repeats point; no response from child)
- MOT: These're pencil cars.
- MOT: How many pencil cars?
- CHI: One, two, three, four, five, six, seven, eight.
- MOT: Yes.

To keep interaction smooth, it is to children's advantage to keep gaps short. By using delay markers, children uphold something akin to the minimal-gap-minimal-overlap timing pattern we see in adult conversation, only at a slower pace. Importantly, their use of delay markers implies that children know there is an appropriate time to respond, even if they can't manage to keep up with the pace of the conversation. This is curious at first, since we know children's response delays before 6;0 are lengthy compared to that of adults (Garvey & Berninger, 1981). But their eye movements reveal that by 24 months, children expect immediate turn-switches while watching the conversations of others (Casillas & Frank, 2013; Von Hofsten, Uhlig, Adell, & Kochukhova, 2009). Children's assumption that responses will come immediately displays an early inclination toward temporal contingency in conversation, one which may develop initially in pre-verbal interactions with their caregivers (Bloom, 1975; Bloom 1984; Masataka, 1993).

In light of this, it is less surprising that children – even at 2;0 – might have a sense of the “right” time to respond and mark delay accordingly.<sup>8</sup> We saw here that in marking delay, children pick up on all sorts of turn-initial cues from ambient adult speech, from *uh* to repetition to prolongation, and use them in conventional and creative ways to get and hold the floor on time. By age 2;0, children are active participants in conversation, and are thus partially responsible for keeping the interaction smooth and intact. So even in children's early conversations, delay markers can effectively aid turn taking by making it clear who is claiming the floor next and how long their delay will be in producing the content of a response.

## General discussion

The way children learn language is bound to their interactional environment. Simple as it seems, it is easy to take this fact for granted. In theories of language

<sup>8</sup>. As adults, their sense of the “right” timing may help them infer unspoken meanings when a response comes after “too short” or “too long” of a delay (Brennan & Williams, 1995; Burgoon, Buller, & Guererro, 1995; Roberts, Francis & Morgan, 2006; Smith & Clark, 1993).

acquisition, the conception of linguistic “input” is a primary determinant of proposed native and non-native skills and knowledge, and has been the basis for many arguments about the learnability of language (Gold, 1967; Wexler & Culicover, 1983). The term “input” itself suggests automatic processes as if, for the first five years of their lives, children have no role in their language learning. But everyday interactions between young children and their caregivers tell a different story. Caregivers tailor their utterances to their young children, and young children take control of conversation early in their development. As we have seen here, they even begin managing the timing of their interactions from age two.

In this chapter I have briefly investigated children’s acquisition of turn timing to begin addressing how they gain active influence over their input. Specifically, I discussed how they use *uh*, *um*, and other delay markers to get and hold the floor in conversation. We saw extensive individual variation in how children accomplish timely turn taking. This variation may be key to understanding the different ways that children engage in conversation to learn about the world around them.

By keeping turn taking smooth during interaction, children enable themselves to make more meaningful contributions, handling their responses within a time-frame *they* control. Smooth interaction allows children to make the most use of across-turn learning contexts (e.g., feedback and hypothesis testing) with less interruption from vocal overlap, conflicting prompts, and misunderstandings. In the results I have presented here, children begin to develop delay-marking and general turn taking skills early on. The realization of these skills, as they develop, helps children to engage successfully in conversation, even when their linguistic abilities lag behind those of their interlocutors.

Turn taking is one of many conversational skills that might fundamentally affect how children learn language. Many details about how to proceed in interaction are wrapped within culture-specific conventions, but basic principles of verbal exchange, like turn taking, are shared cross-culturally (Levinson, 2006; Stivers et al., 2009). While acquiring a language, children around the world learn to take turns effectively because it allows them to gain access to the linguistic and social world around them through coordinated interaction. By exploring the early stages of these skills, we can get insight into how the motive for learning language – to communicate with others – affects the trajectory of children’s language development and the ultimate outcome of their learning. Eve Clark has challenged us to fully embrace this perspective in the study of language acquisition.

## References

- Balog, H.L., & Roberts, F.D. (2004). Perception of utterance relatedness during the first-word period. *Journal of Child Language*, 31, 837–854.
- Bloom, K. (1975). Social elicitation of infant vocal behaviour. *Journal of Experimental Child Psychology*, 20, 51–58. DOI: 10.1016/0022-0965(75)90025-9
- Bloom, K. (1984). Distinguishing between social reinforcement and social elicitation. *Journal of Experimental Child Psychology*, 19, 209–222. DOI: 10.1016/0022-0965(75)90085-5
- Bloom, K. (1988). Quality of adult vocalizations affects the quality of infant vocalizations. *Journal of Child Language*, 15, 469–480. DOI: 10.1017/S0305000900012502
- Brennan, S.E., & Williams, M. (1995). The feeling of another's knowing: Prosody and filled pauses as cues to listeners about the metacognitive states of speakers. *Journal of Memory and Language*, 34, 383–398. DOI: 10.1006/jmla.1995.1017
- Burgoon, J.K., Buller, D.B., & Guerrero, L.K. (1995). Interpersonal deception: IX. Effects of social skill and nonverbal communication on deception success and detection accuracy. *Journal of Language and Social Psychology*, 14, 289–311. DOI: 10.1177/0261927X95143003
- Casillas, M., Bobb, S.C., & Clark, E.V. (Under review). Turn taking, timing, and planning in early language acquisition.
- Casillas, M., & Frank, M.C. (2013). The development of predictive processes in children's discourse understanding. In M. Knauff, M. Pauen, N. Sebanz, & I. Wachsmuth (Eds.), *Proceedings of the 35th Annual Meeting of the Cognitive Science Society* (pp. 299–304). Berlin: Cognitive Science Society.
- Clark, E.V. (2007). How young children describe events in time. In G.B. Flores d'Arcais & W.J.M. Levelt (Eds.), *Advances in psycholinguistics* (pp. 275–293). Amsterdam: North-Holland.
- Clark, E.V. (2009). *First language acquisition, 2nd Ed.* Cambridge: CUP. DOI: 10.1017/CBO9780511806698
- Clark, E.V., & Bernicot, J. (2008). Repetition as ratification: How parents and children place information in common ground. *Journal of Child Language*, 35, 349–372. DOI: 10.1017/S0305000907008537
- Clark, H.H., & FoxTree, J.E. (2002). Using *uh* and *um* in spontaneous speaking. *Cognition*, 84, 73–111. DOI: 10.1016/S0010-0277(02)00017-3
- Demuth, K., Culbertson, J., & Alter, J. (2006). Word-minimality, epenthesis, and coda licensing in the acquisition of English. *Language and Speech*, 49, 137–173. DOI: 10.1177/0023830960490020201
- Dunn, J., & Shatz, M. (1989). Becoming a conversationalist despite or because of having an older sibling. *Child Development*, 60, 399–410. DOI: 10.2307/1130985
- Ervin-Tripp, S. (1979). Children's verbal turn taking. In E. Ochs & B. Schieffelin (Eds.), *Developmental Pragmatics* (pp. 391–414). New York, NY: Academic Press.
- Fitneva, S.A. (2012). Questions and children's social information gathering. In J.P. de Ruiter (Ed.), *Questions: Formal, functional, and interactional perspectives* (pp. 165–178). Cambridge: CUP. DOI: 10.1017/CBO9781139045414.012
- Garvey, C., & Berninger, G. (1981). Timing and turn taking in children's conversations. *Discourse Processes*, 4, 27–57. DOI: 10.1080/01638538109544505
- Gold, E.M. (1967). Language Identification in the Limit. *Information and Control*, 10, 447–474. DOI: 10.1016/S0019-9958(67)91165-5
- Hudson Kam, C.L., & Edwards, N.A. (2008). The use of *uh* and *um* by 3- and 4-year-old native English-speaking children: Not quite right, but not completely wrong. *First Language*, 28, 313–327. DOI: 10.1177/0142723708091149

- Levinson, S.C. (2006). On the human “interaction engine”. In N.J. Enfield & S.C. Levinson (Eds.), *Roots of human sociality: Culture, cognition and interaction* (pp. 39–69). Oxford: Berg.
- Lieven, E.V.M. (1978). Conversations between mothers and young children: Individual differences and their possible implications for the study of language learning. In N. Waterson & C. Snow (Eds.), *The development of communication: Social and pragmatic factors in language acquisition* (pp. 173–187). New York, NY: Wiley.
- MacWhinney, B. (2012). *The CHILDES project*. Retrieved from: <[www.childepsy.cmu.edu](http://www.childepsy.cmu.edu)>.
- Mannle, S., & Tomasello, M. (1987). Fathers, siblings, and the Bridge Hypothesis. *Children's Language*, 6, 23–42.
- Masataka, N. (1993). Effects of contingent and noncontingent maternal stimulation on the vocal behaviour of three- to four-month-old Japanese infants. *Journal of Child Language*, 20, 303–312. DOI: 10.1017/S0305000900008291
- Roberts, F., Francis, A.L., & Morgan, M. (2006). The interaction of inter-turn silence with prosodic cues in listener perceptions of “trouble” in conversation. *Speech Communication*, 48, 1079–1093. DOI: 10.1016/j.specom.2006.02.001
- Roy, B., Frank, M.C., & Roy, D. (2009). Exploring word learning in a high-density longitudinal corpus. In N. Taatgen & H. van Rijn (Eds.), *Proceedings of the 31st Annual Meeting of the Cognitive Science Society* (pp. 2106–2111). Amsterdam: Cognitive Science Society.
- Sacks, H., Schegloff, E.A., & Jefferson, G. (1974). A simplest systematics for the organization of turn taking for conversation. *Language*, 50, 696–735. DOI: 10.2307/412243
- Saylor, M.M., Baird, J.A., & Gallerani, C. (2006). Telling others what's new: Preschoolers' adherence to the given-new contract. *Journal of Cognition & Development*, 7, 341–379. DOI: 10.1207/s15327647jcd0703\_7
- Smith, V.L. & Clark, H.H. (1993). On the course of answering questions. *Journal of Memory and Language*, 32, 25–38. DOI: 10.1006/jmla.1993.1002
- Snow, C.E. (1977). The development of conversation between mothers and babies. *Journal of Child Language*, 4, 1–22. DOI: 10.1017/S0305000900000453
- Stivers, T., Enfield, N.J., Brown, P., Englert, C., Hayashi, M., Heinemann, T., Hoymann, G., Rossano, F., de Ruiter, J.P., Yoon, K.-E., & Levinson, S.C. (2009). Universals and cultural variation in turn taking in conversation. *Proceedings of the National Academy of Science*, 106, 10587–10592.
- Tomasello, M., & Farrar, M.J. (1986). Joint attention and early language. *Child Development*, 6, 1454–1463. DOI: 10.2307/1130423
- Uther, M., Knoll, M.A., & Burnham, D. (2007). Do you speak E-NG-L-I-SH? A comparison of foreigner- and infant-directed speech. *Speech Communication*, 49, 2–7. DOI: 10.1016/j.specom.2006.10.003
- Van Der Wege, M.M., & Ragatz, E.C. (2004). Learning to be fluently disfluent. In K. Forbus, D. Gentner, & T. Regier (Eds.), *Proceedings of the 26th Annual Meeting of the Cognitive Science Society* (p. 1647). Chicago, IL: Cognitive Science Society.
- Von Hofsten, C., Uhlig, H., Adell, M., & Kochukhova, O. (2009). How children with autism look at events. *Research in Autism Spectrum Disorders*, 3, 556–569. DOI: 10.1016/j.rasd.2008.12.003
- Wellman, H.M., & Lemper, J.D. (1977). The naturalistic communicative abilities of two-year-olds. *Child Development*, 48, 1052–1057. DOI: 10.2307/1128359
- Wexler, K., & Culicover, P.W. (1983). *Formal principles of language acquisition*. Cambridge, MA: The MIT Press.

PART 2

**The role of paralinguistic information  
in language learning (three papers)**



# Temporal synchrony in early multi-modal communication\*

Barbara F. Kelly

University of Melbourne

Children's early gesture and speech combinations are typically asynchronous. As children develop the ability to employ speech and gesture together more regularly, the two modalities become synchronized and the word indicates one element, while the target of the gesture indicates another element. This chapter presents a longitudinal qualitative analysis of the path by which children develop the skill of temporal coordination in their gesture and spoken word communications. It investigates how this temporally coordinated information is taken up by the caregiver and examines synchrony of gesture and speech indexing the same element, in order to determine how this fits within the child's transition to reaching the stage of multi-element communications.

## A personal note

Eve's dynamic and ever-challenging acquisition work drew me into the field of language acquisition as an undergraduate, with her funny insights into quirky and laugh-aloud children's talk (Safeway – "is this where you get safe?"). Her influence and gentle guidance has continued to grow in importance for me at each stage of my academic life, as an encouraging and generous dissertation examiner, teaching mentor, colleague and friend. Beyond all this, her recommendations of terrific reads and fine films are cracking.

## Introduction

Multi-modality is crucial to both the acquisition of the lexicon and the development of grammar. It is only by looking across modalities and across interactants, that we

---

\* This paper was written while I was a visiting scholar at Stanford University and University of California, Santa Barbara. I thank Eve Clark and Patricia Clancy for wonderful discussions and ideas; naturally any errors are my own.

can go some way toward understanding how a child of twelve months who points and says “mama” becomes a child of twenty-eight months who says “Eve, this is a snow-car” (Clark, 1993:40). Eve Clark has long championed the role of such social interactions in language learning and in doing so has both echoed and extended De Laguna’s (1927:91) assertion that “in order to understand what the baby is saying you must see what the baby is doing”. In the current chapter, we look through the lens of this assertion toward an examination of children’s early speech and gesture combinations.

When children first begin combining gestures<sup>1</sup> with words and proto-words, they do so as a means of indexing targets and gaining caregiver attention. The two communication modalities occur asynchronously in these early uses and for English speakers the word generally indicates the element that the child is gesturing toward. As they use speech and gesture together more regularly, the two modalities become synchronized and the word indicates one element, while the target of the gesture another element.

In order to be able to temporally coordinate their utterances children must learn how to combine different pieces of information. For example, when combining a gesture + word with respect to a single event, temporal synchrony is crucial if the two communications are to be construed by an addressee as being related to the same event. Thus, the focus of the current chapter is a longitudinal qualitative analysis a longitudinal qualitative analysis of the path by which children develop the ability to employ the mentally challenging skill of temporal coordination and how this temporally coordinated information is construed by the caregiver is the focus of the current chapter.

### Stages of gesture-speech development

English-learning children follow similar paths of development, from using gestures that reflect those modeled by parents (Clark & Estigarribia, 2011, Estigarribia & Clark, 2007) to using multi-word utterances (Bates, 1976, Capirci, Caselli, Iverson, Pizzuto, & Volterra, 2002). Goldin-Meadow and colleagues have worked toward charting and understanding this development by investigating gesture-speech combinations and children’s production and comprehension of these at the one-word stage and beyond (Goldin-Meadow, 1999; Morford & Goldin-Meadow, 1992; Özçalışkan & Goldin-Meadow, 2005). This research suggests that when children begin combining speech and gesture, the two modalities

---

1. *Gesture* is defined as visible bodily actions by which meaning is represented (Kendon, 1982). For the analysis of temporal synchrony each of the gestures was coded according to the hand shape.

communicate about the same element, such as when holding out a shoe and saying “shoe”. Speech and gesture later come together as an integrated system much like in adult usage. For example, children use gestures together with speech to communicate more than one thing at once, such as when holding out a shoe and saying “my” (Butcher & Goldin-Meadow, 2000; Iverson & Goldin-Meadow, 2005; Goldin-Meadow & Butcher, 2003). When there is a mismatch between the target of a gesture and the element indicated by a word (such as pointing to a cup and saying *my*), this difference signals a readiness for cognitive growth and is a harbinger of two-word speech.

While there is good evidence that gesture-word synchrony plays a role in multi-word development, it is not always clear what the gestures, and in some cases the words, are actually conveying. A closer look at the types of examples in Goldin-Meadow (1999), Butcher and Goldin-Meadow (2000), and Goldin-Meadow and Butcher (2003) has raised questions regarding what information the gesture and the word carry, or are considered to carry by the caregiver. Pizzuto, Capobianco, and Devescovi (2005) make a case for an analysis in which early gestures contain a range of semantic possibilities. They argue that many of the examples in Goldin-Meadow and Butcher (2003) could be assessed as having different meanings if analyzed with no assumptions regarding grammatical categories, as expounded in Capirci, Pizzuto, and Volterra's (1996) method of analysis. For example, they suggest a POINT + “dog” communication may indicate ‘that dog’ (p. 185) and thus have a different semantic value from that of a simple noun label ‘dog’. Regardless of the semantic value being carried by the communication, Kelly (2011) has shown caregivers respond to multi-modal communications differently than to mono-modal communications.

In line with Capirci et al's (1996) model of not imbuing intentionality of semantic content or grammatical category on a child, the current study seeks to examine speech and gesture synchrony across child-caregiver interactions, focusing on the caregiver construal of child information and the unfolding process of developing temporal synchrony and multi-element communications. While it is not possible to determine the child's intention regarding the semantic content being communicated, the focus here is on the caregiver response to the communication and then the child response to the preceding action-response pairing. In light of the finding that temporal synchrony encoding different elements is important for the transition to multi-word utterances, the current study extends the research of Goldin-Meadow and colleagues by investigating an earlier step in development and examines synchrony of gesture and speech indexing the same element, in order to determine how this fits within the child's transition to reaching the stage of multi-element communications.

## Data

### *Participants*

Data for this study is a collection of longitudinal spontaneous interactions of caregivers and children aged about 12 months to 30 months in a daycare center. While the majority of longitudinal language development studies take place in the home with parents, the children in this study routinely spent around 40 hours a week at daycare, interacting with the same caregivers and children, in a different but not exceptional language-learning environment. Video recordings of five monolingual children were carried out at roughly weekly intervals over a period of 18 months, for 10-60 minute recordings (averaging around 45 minutes a session) and the data forms part of the Very Young Children corpus, as presented in Kelly (2006, 2011). The five children focused on in this research are known by pseudonyms and are one boy, Brailey and four girls: Chera, Lette, Caitlin, and Fiona.

### *Data analysis*

#### *Focal gestures + words*

Three gestures – visible bodily actions by which meaning is represented – occurred across all the children in their earliest recordings and these are what I term: *gimme* (holding out an open palm as if to receive an object while gaze is on the caregiver or the object); *point* (a movement of the index finger or outstretched hand toward an object while focusing a simultaneous gaze on the object or a caregiver) and *attention focus* (holding an object out while looking at the caregiver or holding it out and then throwing it down while looking at the caregiver). Children's vocalizations were coded as 'words' if they were construed as words by the caregiver, regardless of whether an analyst would classify them as words or proto-words.

#### *Synchronization*

Utilizing McNeill's (1992) analysis of gesture stages, this study of synchrony focuses on the stroke phase of the gesture – the peak of effort in gesture – because this is the phase at which the gesture is clearly taken by the caregiver to be indicating an object (Kelly, 2006). When a child's gesture is at its peak and the child utters a vocalization (word or proto-word), this is counted as being synchronous. The data were examined using ELAN, with synchrony being measured within the space of ten frames, on average. Throughout the paper times are

presented in increments of 0.5 of a second<sup>2</sup>. For example, 29:54.5 is 29 minutes 54.5 seconds<sup>3</sup>.

## The findings

The stages of gesture synchrony development for each of the children are given below. They are based on the order in which different communicative strategies are used, rather than being based on the children's MLU or solely on age. Previous research detailing developmental ordering of word and gesture communication has focused on age and MLU, and in doing so, has tended to collapse two stages of development. These are children's use of gestures with a single word to indicate the same element, at first not temporally synchronized, and later with temporal synchrony.

Table 1 below indicates the developmental ordering of strategies (numbered 1-6), for the five children in this study. All children except Fiona were using strategy 1 and 2 at the onset of recording. Fiona was absent for several early recordings and the lack of word-only uses is therefore likely to be a product of the recording frequency.

**Table 1.** *Developmental ordering of communicative strategy stages (age given in months;days).*

	STRATEGY STAGES	AGE OF INITIAL USE				
		Brailey	Caitlin	Chera	Fiona	Lette
1	gesture used as the sole strategy	12;2	10;2	13;3	13;0	13;2
2	words or proto-words used alone	12;2	10;2	13;3	17;1	13;3
3	gesture + single word indicating same element, not synchronized	12;4	12;1	14;1	13;3	13;3
4	<b>gesture + single word indicating same element, synchronized</b>	16;4	14;2	16;1	17;1	17;4
5	gesture + single word indicating different elements	17;4	15;3	17;0	17;3	18;2
6	onset of two-word speech	17;4	16;0	17;3	17;3	21;1

2. Interestingly, a comparison of a broader array of data coded for temporal synchronization on the basis of analogue recordings (Kelly, 2003) yields the same results – i.e. all crude measures of synchronization to the second and half-second are supported in this narrower coding.

3. An even more micro-level analysis of multi-modal temporal correspondence can be found in the work of Abry, Vilain, and Schwartz (2005) who investigated gestural correspondences, specifically pointing, and metrical phonological units in speech.

As I show in Table 1, Brailey, Caitlin, and Chera used gestures alone and words alone in their first recordings, so it is likely that they had acquired these abilities prior to the onset of recording. Fiona and Lette did not begin to use words until several weeks into the recordings. However, they may have also used them prior to this first documentation of use. This table indicates that the children's communicative gestures occur at Stage 1, before children begin to use proto-words and words (Stage 2)<sup>4</sup>, and these single-modality communications precede the use of a combination of modalities. In turn, combined use of different modalities to refer to different elements precedes the onset of two-word speech. These strategy stages follow a developmental ordering that researchers across a range of studies have noted and discussed (Iverson & Goldin-Meadow, 2005; Morford & Goldin-Meadow, 1992; Özçalışkan & Goldin-Meadow, 2005).

However, there is a missing link in the literature focusing on language acquisition and gesture, which I have labeled 'Stage 4' in Table 1. I have shown that at Stage 3, the children in this study first integrate a word and gesture using the two communication strategies to indicate the same element. Children's pointing gestures have been shown to provoke different responses within a conversation, according to whether they occur with vocalizations (Kelly, 2006; Filipi, 2009) but at this stage the communications are not temporally synchronized. For example, in one instance, Brailey points at water coming out of a hose and then says "[hoa]" (water) while retracting his outstretched hand.

At Stage 4, temporal synchronization between the gesture and the word becomes regular. At this stage the children can now coordinate their spoken word and gesture so that when a gesture and a word are used to indicate an element they are temporally synchronized. In one such example, Fiona used a GIMME gesture toward a ball that was slightly out of reach and said "ball" at the peak of the gesture.

At Stage 5 there is a change in the children's use of the two modalities: they begin using the two modalities to refer to different elements. For example, in one instance, Chera pointed to a hat belonging to Brailey and looked toward the caregiver while saying his name. She indicated the hat through her gesture and used her speech to convey additional information within the communication.

In Table 1 I present Stages 3 and 4 as two distinct developmental stages that need to be differentiated in studying gesture and spoken word communications. Treating them as a single developmental stage will result in an underestimation of a child's communicative ability. Also, separating out synchronous and asynchronous

---

4. Data collection for the current study began while the children were at stages 1 and 2, but the presentation order of these stages is reinforced by others' research (Bates, 1976; Butcher & Goldin-Meadow, 2000; Goldin-Meadow & Butcher, 2003; Werner & Kaplan, 1963).

gestures indicating the same element yields a vital clue that can be used to determine whether a child is ready to move on to using gesture and speech to indicate different elements. In the following sections I show caregiver responses to each child's multi-modal communications across each of the proposed stages. Each Table indicates the child's utterance and gesture, both of which are presented alongside the function of the communication, as evidence in the caregiver responses and the subsequent child response. This method of data analysis allows us to see when the caregiver's assumption about the child's intended meaning is problematic for the child. It is based on an assumption that no negative response from the child suggests the caregiver was correct in their meaning judgment (or at least not incorrect since this would surely be highlighted by the child).

### *Brailey*

When Brailey first began using a spoken word and gesture together, there were fairly large temporal gaps between the gesture stroke and the utterance. In an early recording, Brailey completed four out of seven utterances prior to beginning the preparation and pre-stroke of the gesture, and in two out of seven instances the gesture onset was after the beginning of the utterance. For example:

- (1) *Speech and gesture asynchrony (same element)* [12 months]

Function	Time	Strategy	Communication	Element
			gaze toward caregiver	
indicate toy	15:19	utterance	[dæ:]	toy
show toy	15:21.5	ATTENTION FOCUS	holds out toy to caregiver	toy
	15:21.5	cg response	looks at toy	
	15:22	child response	retracts arm	

In (1) Brailey utters the word [dæ:] and then, once he has established that the caregiver's attention is focused on him, he holds the toy up toward her, she looks and he then retracts his arm with the toy and begins playing with it again. He has not yet begun to synchronize his multi-modal communications. In (2), when he is aged 15 months, the timing of Brailey's speech and gesture was coordinated – the stroke of Brailey's gesture coincided temporally with his productions of [də]. He was able to use speech and gesture to communicate about a single element which the caregiver looked at and then took, and played with.

(2) *Speech and gesture synchrony (same element)*

[16 months]

Function	Time	Strategy	Communication	Element
			gaze toward caregiver	
indicate tissue	36:05	utterance	[də]	tissue
give tissue	36:05	ATTENTION-FOCUS	holds up tissue to caregiver	tissue
	36:05	cg response	looks at tissue	
	37.05	child response	lowers arm while moving away	

Example (3) is the first recorded use of Brailey combining speech and gesture to indicate more than one element. In this example Brailey is holding a book which another child (Chera) was playing with earlier and now wants again. She is crying and looking up at the caregiver. Brailey looks first at the caregiver and then holds the book out to Chera and says [bu] while still holding the book towards her. While the caregiver does not respond to Brailey to confirm the different semantic content, Brailey's gesture and speech communications are clearly conveying different information. Chera is the target of the gesture – it is toward her that Brailey holds out the object, and he appears to be holding it out as an 'offer' rather than a 'show' – while his utterance is an approximation of the category of the object he holds, a book. He holds on to the book through the caregiver's utterance and then passes it to Chera who immediately drops it onto the floor. Brailey then reaches for another book and falls forward onto the floor. Around 3.5 weeks after Brailey first temporally coordinates speech and gesture, he begins using the two modalities to communicate about more than one element in an event, as in (3).

(3) *Speech and gesture (different element)*

[17 months]

Function	Time	Strategy	Communication	Element
			gaze toward caregiver	
			gaze moves to Chera	
indicate book	15:20.5	utterance	[bu]	book
give book	15:22	ATTENTION-FOCUS	holds out book to Chera, who is crying	Chera

	15:24		caregiver looks at Brailey and says to Chera: "You can have his book. You can say 'thank you Brailey'"	
	15:26.5	<b>Chera response</b>	takes book and drops it to floor	
	15:27	<b>Brailey response</b>	reaches for another book on floor	

In the above example Brailey has coordinated speech and gesture and used them to communicate about two different elements in the scene, the book and Chera.

### *Caitlin*

Caitlin integrated speech and gesture at an earlier age than the other children. She is over three months younger than the oldest child, Chera, and conceivably, the interaction with the older children has helped Caitlin develop these skills at an earlier stage. In (4), Caitlin is aged 13 months. She uses speech and gesture to indicate a banana the caregiver is holding.

(4) *Speech and gesture asynchrony (same element)*

[13 months]

Function	Time	Strategy	Communication	Element
			gaze on caregiver	
index banana	03:44	<b>POINT</b>	to banana in caregiver's hand [6 secs]	banana
indicate banana	03:50:5	<b>utterance</b>	[a:na]	banana
	03:51:5	<b>cg response</b>	you want some banana?	
	03:52	<b>child response</b>	stops moving, turns to look at banana and waits for a piece to be cut off	

In (4) Caitlin points to a banana that the caregiver has been feeding to the children and now holds in her hand. Interestingly, this is not a POINT with an extended index finger, but with a closed fist – it is possible that Caitlin is holding something in her hand and does not want to open it fully and so has to keep her fist closed. The caregiver is focused on another child and does not move her attention to Caitlin. Caitlin maintains the POINT for 6 seconds in a post-stroke hold before retracting the gesture and saying, [a:na]. The caregiver responds by moving her attention to Caitlin and saying “you want some banana?”. In the following example, three months later, Caitlin has moved to a stage at which she can coordinate her gesture and speech. In this stage she uses the two modalities to communicate about the same element. For example:

(5) *Speech and gesture synchrony (same element)*

[14 months]

Function	Time	Strategy	Communication	Element
			gaze toward caregiver	
indicate broom	32:06	utterance	“beep”	broom
index broom	32:06	ATTENTION FOCUS	holds broom up toward caregiver	broom
	32:07	cg response	caregiver looks at broom, nods head and says “aha”	
	32:08	child response	retracts broom and pulls toward her body	

In (5) Caitlin uses a non-conventional word to indicate a broom. She says “beep” while holding up the broom. Caitlin may be requesting the caregiver to use the broom to sweep. However, given the caregiver’s reaction, which is simply to nod her head and say “aha”, treating the word as though it is the name of the object, it appears that the caregiver is reacting toward this word as she does toward labels, rather than treating it as a request for action. Caitlin does not maintain her ATTENTION FOCUS gesture<sup>5</sup> so that the caregiver might take the broom, nor does she reiterate the gesture to indicate that she desires a different response from the caregiver. She simply lets the broom fall slightly and then pulls it back toward her body, ready to play with it again.

5. This gesture tends to have fewer accompanying utterances than POINT and GIMME (Masur, 1983) but is clearly part of the child’s communicative system.

(6) *Synchronous integration with unclear elements* [15 months]

Function	Time	Strategy	Communication	Element
			gaze toward oranges on table	
indicate self	42:01	utterance	[mə]	Caitlin (self)
index oranges	42:01			
	POINT	toward oranges		orange
	42:03		gaze toward caregiver; hand still extended but fingers retracted	

The example in (6) above is an interesting speech and gesture integration with a different element. In part this is because it is not definitive as to whether the speech and gesture do involve different elements. The [mə] here is different from a similar utterance that we shall examine in Chera's uses. It is clearly not an approximation of the name of the element. However, it is not clear whether Caitlin is using speech to convey *more*, in which case the element is the same, but the caregiver appears to assume so since she responds by giving Caitlin more orange. In the following example Caitlin uses speech and gesture to communicate about two different elements.

(7) *Speech and gesture (different elements)* [14 months]

Function	Time	Strategy	Communication	Element
		gaze	toward caregiver	
direct attention to Chera	13:04.5	POINT	at Chera who has just put down a book	Chera
indicate book	13:06.5	utterance	[bu]	book
	13:08	cg response	you can get another book	
	13:08.5	child response	looks at caregiver, and then leans toward toys and books	

In the above example Caitlin directs her POINT toward Chera while uttering her word for “book”. In this instance it appears that Caitlin is attempting to communicate something about Chera’s possession of the book. It is not evident what this may be since immediately after she communicated this information she turned her attention to something else but the caregiver responded by saying “you can get another book” indicating that she took the communication to be about Chera’s possession of the book, not a label for the book. At this stage Caitlin now has the ability to talk about more than one element in an event and this developed only after she could temporally synchronize her speech and gesture.

### *Chera*

Although Chera’s earliest speech and gesture combinations were not temporally synchronized, there was often an overlap in the use of speech and gesture. Compared to the other children in the study, Chera held her gestures for substantial amounts of time. For example, at 13 months she held a pointing gesture for 13 seconds without uttering a sound. In example (8), Chera (aged 14 months) is having a snack. She points to a bottle of juice on far side of the table and 2 seconds later says [dʒu], which the caregiver responds to by looking at and saying “that’s your juice”. Chera then retracts her hand and arm but maintains her gaze on the juice. While Chera can clearly integrate gesture and speech, these are not synchronized at this stage.

(8) *Speech and gesture asynchrony (same element)*

[14 months]

Function	Time	Strategy	Communication	Element
			gaze on bottle of juice	
index juice	13:04	POINT	at bottle of juice	juice
indicate juice	13:06	utterance	[dʒu]	juice
	13:08	cg response	“that’s your juice”	
	13:08	child response	retracts arm and hand but maintains gaze on juice	

In (9), Chera (aged 16 months) is sitting eating lunch. She begins an interaction by pointing to a container of pasta and saying “baba”. The gesture and the utterance here are synchronized. The caregiver responds by saying, “Do you know what this is? This is pasta. Do you want to try a bite?” In the example below, Chera then

holds her gaze on the pasta and there is a 6 second pause between the caregiver's suggestion to try some and when Chera's says [ba], which she says slightly before she extends her arm and points to the pasta.

(9) *Speech-gesture synchrony & asynchrony (same element)* [16 months]

Function	Time	Strategy	Communication	Element
			gaze moves to pasta container	
indicate pasta container	07:06	utterance	[babab] (pasta?)	pasta container
index pasta container	07:06	POINT	to container of pasta	pasta container
	07:07.5	cg response	"Do you know what this is? This is pasta. Do you want to try a bite?"	
			gaze on pasta container in front of her	
indicate pasta	07:13	utterance	[ba]	pasta container
index container	07:14.5	POINT	to container of pasta	container

Chera's speech-gesture combinations in the above example are synchronous in her first use, at 07:06, and asynchronous in her second, at 07:13 and 07:14.5. This suggests that she is on the cusp of being ready to move on to using speech and gesture to communicate about different elements.

In (10), when Chera is 17 months, three weeks older than in the previous example, we see her first use of a synchronized gesture-speech combination: she holds out her shoe to the caregiver and says [gu] (her word for shoe). The hold is maintained for 1.5 seconds and then the shoe is dropped when the caregiver responds with a head nod and the utterance "shoe" but does not take the shoe, which is then dropped by Chera.

(10) *Speech and gesture synchrony (same element)*

[17 months]

Function	Time	Strategy	Communication	Element
			gaze on caregiver	
indicate shoe	26:16	utterance	[gu] (shoe)	shoe
call attention to shoe	26:16	ATTENTION FOCUS	holds shoe out towards caregiver	shoe
	26:16.5	cg response	nods head “shoe”, then looks away	
	26:18	child response	drops shoe	

In (10) the communications are clearly synchronous and the caregiver's response of “shoe” suggests the communications are considered to be indexing the same element. Chera can now coordinate speech and gesture temporally, and it is likely that she will soon be able to use the two modalities to communicate about different elements.

In example (11), speech and gesture are synchronized but are construed by the caregiver as referring to different elements. Here, Chera (aged 18 months) is holding out her cup toward a bottle of milk while looking at the caregiver and saying [mɔ]. This gesture was coded as a GIMME rather than an ATTENTION FOCUS gesture because Chera's fingers are outstretched at the side of the cup as she reaches toward the bottle of milk. It is as though she would pick up the bottle too, if she could. The caregiver assumes this and responds by asking “more?”. Chera maintains her position with the cup held out as though in anticipation of more milk filling it.

(11) *Speech and gesture (different element)*

[18 months]

Function	Time	Strategy	Communication	Element
			gaze on caregiver	
express desire for more milk	04:23	utterance	[mɔ]	more
index milk as target	04:24	GIMME	reaches toward bottle of milk while holding cup	bottle of milk
	04:25:5	cg response	“more?”	
	04:25.5	child response	maintains position with cup held out	

Example (11) indicates that in the month immediately following Chera's use of speech and gesture synchrony with the same element, she is able to use speech and gesture to refer to different elements. Chera has integrated speech and gesture and is now able to use speech to refer to one element, in this case the desire for more, and gesture to index the target that she wants more of, in this case milk.

The sequence of development in the above examples suggests that, in line with the unfolding developmental progress posited by Goldin-Meadow and colleagues, children's gesture-speech synchrony signals a child's readiness to begin communicating about more than one element at once. Chera began to use speech and gesture to communicate about more than one element within an event only after she had synchronized the use of gesture and speech to indicate the same element.

### *Fiona*

Fiona was using gestures with words at the onset of recording. However, from the earliest recording of Fiona combining gesture and speech, her path of development is in alignment with the other children's development. In (12), she uses speech and gesture to indicate the same element but her communications in using each modality occurred asynchronously.

(12) *Speech and gesture asynchrony (same element)*

[13 months]

Function	Time	Strategy	Communication	Element
			gaze toward box she is holding	
showing box	13:04	ATTENTION FOCUS	lifts box up to her own eye level	box
	13:05		holds out toward caregiver while moving gaze from box to caregiver	
indicate box	13:06	utterance	[bos] (uttered as hand is lowered from full extension)	box
	13:06	cg response	rapid head nod	
	13:06.5	child response	lowers box to lap	

At the stroke of Fiona's gesture in (12), she is silent. She lowers her arm slightly after the stroke while still holding the gesture and she says [bos] (box). The gesture is held for a further 1.5 seconds, post-stroke, before Fiona retracts her hand. The caregiver nods her head and then Fiona lowers the box to her lap, letting it fall from her now outstretched fingers. She uses speech and gesture to indicate the same object (the box) but at this stage she does not use them simultaneously. Synchronous use of the two gestures occurs nearly four months later, as illustrated in (13).

(13) *Speech and gesture asynchrony (same element)*

[17 months]

Function	Time	Strategy	Communication	Element
			gaze moves toward caregiver, then moves to jug of water	
indicate water	55:35	utterance	[chus] (Fiona apparently thinks the water is juice)	water
index water	55:35	POINT	toward food cart with jug of water on it	water
	55:36	cg response	head nod	
	55:36.5		"water"	
	55:37.5		looks at caregiver, then water, then away	

In the example in (13) above, Fiona sees the water on the table and says [chus] (apparently thinking it is juice) while pointing toward the jug. Her pointing gesture is fairly rapid and the stroke is timed exactly with the utterance and the caregiver responds with a slightly mixed message of nodding her head and saying "water". At this stage she has not used gesture and speech together to communicate about more than one element within an event. However, as we can see when we look at example (14), within two weeks of the first recording of gesture + speech synchrony, Fiona combines the two modalities while communicating about different elements.

(14) *Speech and gesture (different element)*

[17 months]

Function	Time	Strategy	Communication	Element
			gaze toward caregiver	
indicate diaper	06:36	utterance	[aipər]	diaper
index Lette	06:37	POINT	to another child, Lette	Lette
	06:38	cg response	head nod + “Lette’s gonna have her diaper changed”	
	06:40	child response	moves away from diaper table	

The example presented above is an interesting instance of a child combining more than one element within an event. Fiona has just had her diaper changed. Typically the children have their diapers changed one after the other. While it is not clear what Fiona intends to communicate when she points to Lette and says [aipər] she is clearly linking Lette with something to do with a diaper. Whether she is indicating that she thinks it is Lette’s turn next or is trying to convey something else altogether, is not clear. Regardless, the caregiver treats the communication as though Fiona is expressing more than one element with respect to a single event, by saying “Lette’s gonna have her diaper changed”. Fiona responds by moving away from the diaper table.

*Lette*

Words accompanying Lette’s early gestures were fairly difficult to understand. Lette’s earliest utterances were not clearly articulated. However, by the age of about 17-18 months, words that she used regularly (*no, my, that, Brailey, book, ball, juice, milk*) became clearer and easier to understand (or at least the onsets and stressed syllables of these words became clearer).

(15) *Speech and gesture asynchrony (same element)*

[12 months]

Function	Time	Strategy	Communication	Element
			gaze toward caregiver, who is busy with another child	
index food as target	10:42	POINT	gaze moves toward plate of food (very limp gesture)	plate of food
indicate food	10:44	utterance	[dæ]	plate of food
	10:44.5	cg response	glances toward Lette (slight head movement)	
	10:44.5		walks away	

In example (15), Lette is near a table where another child is being fed. She does a POINT gesture toward the plate of food but receives no response. While her arm is still in a post-stroke holding position, she says [dæ], but then is immediately distracted by a noise on the other side of the room and walks away without waiting to see if she will receive a response.

Five months later, Lette again combines speech and gesture with respect to the same element. However, as shown in example (16), she uses the two modalities at the same time.

(16) *Speech and gesture synchrony (same element)*

[17 months]

Function	Time	Strategy	Communicate	Element
			gaze toward woman (Lani) outside the porch railing (i.e. outside the children's space); gaze moves to caregiver	
indicate Lani	36:50	utterance	[yani] (Lani)	Lani

index Lani as target	36:50	POINT	to Lani while maintaining eye gaze with the caregiver	Lani
	36:51	adult response	“Hi Lette”	
	36:51.5	child response	retracts hand	

In (16) Lette sees Lani arrive at the daycare center. Lette runs to the railing and points toward Lani while at the same time saying her name [yani] in a loud voice, as though calling to her (Lani responds by calling out “Hi Lette”) and Lette retracts her hand. Lette can now temporally coordinate speech and gesture and is ready to move on to using speech to communicate about one element and gesture to communicate about something different within an event. In example (17), which occurred three weeks after the example in (16), Lette combines modalities across different elements.

(17) *Speech and gesture (same element)*

[18 months]

Function	Time	Strategy	Communication	Element
		gaze	at caregiver who is putting bibs on the children	
indicate bib	53:25	GIMME	reaching toward bib	bib
indicate possession	53:28	utterance	“my”	Lette (self)
	53:31	cg response	“your bib” put the bib on Lette	
	53:32.5	child response	smile	

In (17) Lette uses her speech to refer to herself while a pointing gesture indicates the bib that she apparently wants. It is only now, one month after she has shown an ability to integrate a word and gesture temporally, that she can use them to refer to more than one element within the same speech event, in an example similar to several seen in the literature. The caregiver construes this as meaning *my bib* and responds by putting a bib on Lette saying “*your bib*”. Lette responds with a wordless smile.

### Implications of gesture-word synchrony

Children move through a path of development from using asynchronous gesture-word combinations to using speech and gesture synchronously after months of asynchronous use. This study indicates that children progress in similar ways through each stage of development, and, as shown here, this includes a stage in which their combined modalities are responded to as though indexing a single element prior to being considered to index multiple elements. The progression of these stages and the time children take to move through them allows them to deal with the cognitive complexity of moving on to the next skill. Let us consider why. When children reach a developmental stage in which they can integrate the two modalities as part of a routinized action of communication the routinization of use of these modalities requires very little effort. When the gesture plus speech strategy has become well-established and highly routinized this potentially frees up mental space, enabling new communications. This allows the child to move beyond the communicative strategy of expressing one element about an event and affords the child the mental space or capacity to express a second element within the same communication, prior to moving on to expressing verbal multi-word and multi-event constructions.

In a sense, each stage works as a scaffold for the following stage and the caregiver responses help the child differentiate between what information the gesture is doing (often the work of what will later be a verb, Kelly, 2006) and what information is carried by the word. Suggestions by Kelly (2011) and Pizzuto et al. (2005) regarding a range of possible meanings in a child's multi-element constructions raise the issue of the semantic value of these constructions and the question of whose perspective we base this value on. These researchers argue for a child-based perspective in assessing construction meaning. However, on the basis of caregiver responses to children's communications in this study, Butcher and Goldin-Meadow's (2000) and Özçalışkan and Goldin-Meadow's (2005) model for determining semantic content in the communication via the meanings assessed by the analyst still holds, suggesting that the range of possible meanings construed by the caregiver may not differ from those construed by the analyst. For each of the children in this study, the multimodal combinations and their temporal synchronization preceded moving on to the new developmental stage of using speech and gesture to communicate information construed by caregivers as being about different elements of a scene.

### Concluding remarks

This longitudinal study of how individual children's interaction skills unfold from communications across a single modality to multi-modal synchronized communications has highlighted the meanings caregivers construe across the sequential

structure of these interactions. Each stage of temporal coordination across modalities is a necessary pre-requisite for the next because there is a need for automaticity of execution in order for the child to be able to free enough mental space to be able to express a second element through a different communicative strategy. The yet-to-be-tested prediction that follows from this is that children will not use simultaneous words and gestures to communicate about different elements until *after* they have synchronized the use of a word plus a gesture to communicate about the same element.

## References

- Abry, C., Vilain, A., & Schwartz J. L. Introduction: Vocalize to localize? A call for better crosstalk between auditory and visual communication systems researchers. In C. Abry, A. Vilain, & J.L. Schwartz (Eds.), *Vocalize to localize* (pp. 1–12). Amsterdam: John Benjamins.
- Bates, E. *Language and context: The acquisition of pragmatics* (pp. 313–325). New York, NY: Academic Press.
- Butcher, C., & Goldin-Meadow, S. Gesture and the transition from one- to two-word speech: When hand and mouth come together. In D. McNeill (Ed.), *Language and gesture* (pp. 235–257). Cambridge: CUP. DOI: 10.1017/CBO9780511620850.015
- Capirci, O., Caselli, M.C., Iverson, J.M., Pizzuto, E., & Volterra, V. Gesture and the nature of language in infancy: The role of gesture as a transitional device en route to two-word speech. In D.F. Armstrong, M.A. Karchmer, & J.V. Van Cleve (Eds.), *The study of signed languages: Essays in honor of William C. Stokoe* (pp. 213–246). Washington DC: Gallaudet University Press.
- Capirci, O., Iverson, J.M., Pizzuto, E., & Volterra, V. Gestures and words during the transition to two-word speech. *Journal of Child Language*, 23, 645–673. DOI: 10.1017/S030500090008989
- Clark, E.V. *The lexicon in acquisition*. Cambridge: CUP. DOI: 10.1017/CBO9780511 554377
- Clark, E.V., & Estigarribia, B. 2011. Using speech and gesture to inform young children about unfamiliar word meanings. *Gesture*, 11(1), 1–23. DOI: 10.1075/gest.11.1.01cla
- De Laguna, G. 1927. *Speech: Its function and development*. New Haven, CT: Yale University Press.
- Estigarribia, B., & Clark, E.V. 2007. Getting and maintaining attention in talk to young children. *Journal of Child Language*, 34, 799–814. DOI: 10.1017/S0305000907008161
- Filipi, A. 2009. *Toddler and parent interaction: The organisation of gaze, pointing and vocalization*. Amsterdam: John Benjamins. DOI: 10.1075/pbns.192
- Goldin-Meadow, S. 1999. The role of gesture in communication and thinking. *Trends in Cognitive Science*, 3, 419–429. DOI: 10.1016/S1364-6613(99)01397-2
- Goldin-Meadow, S., & Butcher, C. 2003. Pointing toward two-word speech in young children. In S. Kita (Ed.), *Pointing: Where language, culture, and cognition meet* (pp. 85–106). Mahwah, NJ: Lawrence Erlbaum Associates.
- Iverson, J.M., & Goldin-Meadow, S. 2005. Gesture paves the way for language development. *Psychological Science*, 16, 367–371. DOI: 10.1111/j.0956-7976.2005.01542.x
- Kelly, B.F. 2003. The emergence of argument structure from gesture to speech. Unpublished PhD dissertation. University of California, Santa Barbara.

- Kelly, B.F. 2006. The development of constructions through gesture. In E.V. Clark & B.F. Kelly (Eds.), *Constructions in acquisition* (pp. 11–25). Stanford CA: CSLI.
- Kelly, B.F. 2011. A new look at redundancy in children's gesture and word combinations. In A. Inbal & E.V. Clark (Eds.), *Experience, variation and generalization: Learning a first language* (pp. 73–90). Amsterdam: John Benjamins.
- Kendon, A. 1982. The study of gesture: Some remarks on its history. *Recherches Semiotique/Semiotic Inquiry*, 2, 45–62.
- Masur, E.F. 1983. Gestural development, dual-directional signaling, and the transition to words. *Journal of Psycholinguistic Research*, 12, 93–109. DOI: 10.1007/BF01067406
- McNeill, D. 1992. *Hand and mind*. Chicago, IL: University of Chicago Press.
- Morford, M., & Goldin-Meadow, S. 1992. Comprehension and production of gesture in combination with speech in one-word speakers. *Journal of Child Language*, 19, 559–580. DOI: 10.1017/S0305000900011569
- Özcalışkan, Ş., & Goldin-Meadow, S. 2005. Gesture is at the cutting edge of early language development. *Cognition*, 96(3), B101–B113. DOI: 10.1016/j.cognition.2005.01.001
- Pizzuto, E., Capobianco, M., & Devescovi, A. 2005. Gestural-vocal deixis and representational skills in early language development. *Interaction Studies*, 6, 223–252. DOI: 10.1075/is.6.2.05piz
- Werner, H., & Kaplan, B. 1963. *Symbol formation: An organismic developmental approach to language and the expression of thought*. New York, NY: Wiley.

# **Shared attention, gaze and pointing gestures in hearing and deaf children**

Aliyah Morgenstern  
Sorbonne Nouvelle University

This chapter illustrates the richness of pointing and gaze as integral elements of spontaneous oral interactions both in signing and speaking mother-child dyads. These attention-sharing behaviors help infants interpret their caregivers' productions. The children will then use them as first communication tools.

But they have a particular function for signing children since they are fully integrated into the formal linguistic system of sign language. A comparison between the use of pointing and gaze in the longitudinal data of one deaf signing and one deaf speaking little girl from eight months to two, shows that the deaf child uses gaze and pointing more frequently and with more diversified functions than the hearing child who combines visual and auditory means.

## **A personal note**

Eve Clark has always been at the heart of my scientific life. I cannot remember how and when I met her in person. She seems to have always been there for me, first as a conversational partner in my imaginary dialogues with her as I kept reading her books and her papers, or wrote my own; more recently, as a god-mother, a mentor, a friend, who regularly comes to Paris and with whom I share intense conversations, good meals and great affection. She is constantly telling me to stop doing so much, to sit down and WRITE. So here I am, writing for you Eve, a brief respite.

Even though children have innate biological and cognitive capacities, they need to learn linguistic conventions and formal patterns from the language in their environment. They gradually become fully developed interacting speakers building on such cognitive and social skills as the ability to follow the other's gaze, to draw and maintain their attention, to read others' intentions, to make analogies, to categorize and to symbolize.

Eve Clark has made innumerable in-depth scientific contributions to show that nativist views of language development might not be adequate, and that a

complete account of language development must carefully consider children's social experience and in particular adult-child interactions. She has patiently collected and neatly assembled data to show again and again, in the most elegant manner, the wonderful "richness of the stimulus". This rich input includes gaze and pointing, and I shall try to illustrate in this chapter their importance as integral elements of language, both in signing and speaking children. Gestures, verbal productions, signs, gaze, facial expressions, postures, are all part of our socially learned, inter-subjective communicative system. Human beings, with all their representational skills, combine modalities in order to share meaning, to refer to present and absent entities and events, to express their projects, their desires and their inner feelings. As McNeill pointed out, we might need to "broaden our concept of language" (1992: 2). Thanks to combinations of experimental and naturalistic studies, to video recordings, to a variety of specialized software, international databases, theoretical approaches that include multimodality and multiple levels of analyses, and rich collaborations among experts of several scientific fields related to language development, we now have the tools that help us create new methods to do so.

## Introduction

Social interaction in infancy is dependent on the interplay between infants' affects, their neural learning processes, their perceptual and motor skills and the structure of their social and affective environment (Cole & Cole, 1996). Social information helps infants decipher the meaning of others' language acts (Tomasello, 1999). Their drive to attend to the same objects helps infants enter the language community.

Infants' capacity for attention sharing therefore plays a crucial role in their communication with older children and adults (Brazelton, Koslowki, & Main, 1974; Bruner, 1983; Baldwin, 1993; Tomasello, 1999). The notion of shared or joint attention<sup>1</sup> is used for a whole range of skills that include gaze following, request gestures and postures, and especially pointing. By 9-12 months, even though there are great individual differences in their rhythm of development, most typical infants follow adults' gaze and pointing gestures and learn to discriminate what is important for them in their environment, based on the attention shown and the feedback given by more expert caregivers.

It is the interaction and complementarity between basic perceptual, cognitive and affective processes that seem to trigger and guide the emergence of shared

---

1. Some authors have different definitions of these two notions but I will use them interchangeably.

attention, which will then lead the child into symbolic communication (Deák & Triesch, 2006). Adults and older children constantly recruit infants' attention. Attention-sharing behaviors used by adults such as gaze and pointing gestures first help infants interpret their caregivers' productions. But the children will then use these attention-sharing processes themselves as first communication tools with reinforcement from adults. Children can thus request the adults' attention and monitor it skillfully as they grow up: "gaze and gesture are the early means through which the child can take part in conversation and maintain participation across sequences of talk" (Filipi, 2009: 2).

### *Gaze in hearing and deaf children*

However important gaze is both in the input and in children's communicative behavior, from their very first weeks in the world, all sorts of sounds and visual cues can compete for infants' attention in their ecological environment. Even in experimental studies, adults use verbal cues more often than gesture or gaze as attention getters (before the children look at the objects), especially with young children under 1;6 (Estigarribia & Clark, 2007). They will then rely on the children's gaze to know if they are attending.

Object manipulation is also of high importance during the first year (You, Deák, & Jasso, 2005). According to Deák, Jasso, Krasno and Triesch (2006), infants almost never follow caregivers' gaze shifts unless the adult also manipulates the object or points. In their study in quasi-naturalistic conditions, infants followed mostly combinations of manual actions and gaze shifts. During the play sessions, adults then complemented most of their speech with gestures and constantly referred verbally to the toys they were manipulating.

Hearing children can therefore learn visual skills very early on, but they rely on a combination of semiotic cues used by their adult interlocutors to attract and maintain their attention. They also use a combination of multimodal means themselves to imitate those behaviors. What then is the situation of deaf children who cannot rely on the auditory modality?

Like all children, deaf children learn about the world thanks to shared attention with their communication partners. But they enter language only through visual attention<sup>2</sup>. While hearing children can be looking around them and listening to their caregivers at the same time, deaf children explore their environment

2. I do not want to dismiss the other senses. Touch and to some degree Taste (in the shape of mouthing, one of the babies' ways of exploring the world) certainly play an important role in child development, and contribute to language development even more in the case of blind infants, but I will not discuss them in this study.

AND perceive all their language input thanks to the visual mode. How does the use of one single modality impact on their joint attention skills?

Limousin (2011) has shown, in a longitudinal study of a French deaf girl's linguistic development, the French deaf signing parents' incredible skills at adjusting to their daughter's visual development as she grew up. They constantly used different means to attract and maintain her attention, tapping her lightly, waving their hands towards her, or tapping their foot on the floor to establish eye contact. They could also notice when her concentration was weaker and she needed some respite. They bent down to sign right in her visual field, as she was much smaller than them. They sometimes signed directly in front of her eyes. As she grew up, they trusted her visual skills and capacity to focus more and started signing with fewer repetitions, more quickly, with smaller movements – thus less saliently (i.e. their prosody became less emphatic, as in hearing Child Directed Speech) and without adopting special positions. All these specific strategies are referred to as "Infant Directed Sign" or "Child Directed Sign" (Masataka, 1996; Mayberry & Squires, 2006). Just as Child Directed Speech scaffolds hearing children's use of oral language, Child Directed Sign helps deaf children acquire the same conversational skills but adapted to a visual language. The facilitative strategies enable deaf children to acquire sign language, by attracting and holding their attention. When deaf children do not benefit from those strategies (and only a minority of deaf children have deaf signing parents), their linguistic, emotional and social development can be hindered (Courtin, 2000; Meristo, et al., *in press*).

Deaf children who benefit from Child Directed Sign develop skills that enable them to engage in joint attention from a very young age. Deaf two-year-olds constantly shift their gaze back and forth and therefore connect the speech addressed to them to the objects referred to. Such precocious meaningful gaze shifting is not used as much and with as much control by hearing children at the same age in their natural environment. They reflect specific skills developed from perceiving all of language visually and resorting to one modality where hearing children could also use their voice and their ears.

Deaf children therefore learn to control their own eye gaze in order to alternate attention between signers and objects. If they do not learn those skills quickly enough, they do not "see" enough language, communication is "incomplete" and they do not fully benefit from their input. Hearing children can stop concentrating, get engrossed in their activity, but they still hear verbal productions.

Eye gaze then becomes an integral part of sign language itself since it is "grammaticalized" and serves different linguistic functions. Eye gaze is, for example, what enables signers to distinguish between narrative – gaze is not on the interlocutor – and dialogue – gaze is on the interlocutor (Cuxac, 2000). It also plays a syntactic role to mark pronominal reference (differentiating second and third

person, either complementing or replacing pointing gestures). Eye gaze is therefore an important element of sign productions and must be finely controlled and monitored by children for them to become expert signers. They can only benefit from early focus from their caregivers on the development of the complex abilities necessary to master visual interaction.

### *Pointing in hearing and deaf children*

Pointing, especially with an outstretched and aligned arm, hand and finger (or other parts of the body) is a very natural and salient way to direct another's attention to a new object (Butterworth & Itakura, 2000). Caregivers and infants produce pointing gestures to direct each other's attention (Bates, Camaioni, & Volterra, 1975). Pointing is used by adults once they move away from things that are in "the immediate vicinity of the infant" and which they can handle, show them and hand to them. Thanks to very fine motor adjustments between hand and eye, the children will then pick the things up themselves, handle them, show them, give them to the adult (Clark, 1978: 93). Children then also take up those pointing gestures directed at them, and will reciprocate by showing, commenting or requesting objects that are out of reach with the same means. Of course pointing gestures are rare by comparison to gaze shifting, but they are more salient for infants because of the movement they involve. They also imply more 'effort' and children may think the target is even more interesting when pointing is used. Nine-month-old children follow gaze much more reliably when it is accompanied by a pointing gesture (Flom, Deák, Phill, & Pick, 2003).

Pointing gestures play an important role in the language acquisition process. They are grounded in joint attention, they trigger interaction, and they may also facilitate children's entry into word combination and syntax (Iverson and Goldin-Meadow 2005; Kelly 2011). Numerous studies have tackled this issue in the development of spoken language. The "founding fathers" of the study of child development and language had great intuitions about the importance of gestures and their relation to language. In his notes on his son's development, Darwin (1877) stresses the importance of observing the transition from uncontrolled body movements to intentional gestures. Stern (1924) considers pointing in particular as a precursor of intentional marking. For Werner and Kaplan (1963), pointing represents children's ability to discriminate between external objects and their own person. Communicational pointing then becomes the basis for referential behavior and reciprocity established in common activities between children and their parents (Bruner, 1975). As Tomasello et al. underline, "pointing may thus represent a key transition, both phylogenetically and ontogenetically, from nonlinguistic to linguistic forms of human communication" (2007: 720).

The issue of continuity between gesture and language is quite challenging in the case of sign language, since pointing gestures are fully integrated in the linguistic system (Hoiting & Slobin, 2007), just as gaze is grammaticalized (as explained above). The analysis of pointing in sign language acquisition is a unique occasion to observe the possible continuity or discontinuity between gesture and sign in the Saussurian sense.

When children first produce pointing gestures both in speaking or signing environments, they designate a place, an object, a person or sometimes an event. But for the child who is surrounded by sign language, those pointing gestures are progressively incorporated into her formal linguistic system and used for demonstrative and personal reference among other functions in combination with gaze.

Continuity between pointing gestures and language is questioned by Bellugi & Klima (1982) and Petitto (1986), based on their observations of a time-line discontinuity in the production of pointing gestures. The deaf signing children they observed ceased using points and when they started using them again, there were instances of pronominal inversions. According to these authors, children's pre-linguistic gestures are different from signs despite the same hand-shape and may correspond to two distinct categories of pointing gestures: some indexical and some symbolic (Tomasello, 2003). While discontinuity between pointing gestures and points used as personal pronouns was thus illustrated in the acquisition of American Sign Language, the same phenomenon was not confirmed in Italian Sign Language with longitudinal data (Pizzutto & Capobianco, 2005), nor in the longitudinal recordings of deaf children using French Sign Language (Morgenstern, 1997; Limousin, 2011) which showed no interruption of pointing toward persons and no pronominal reversal. In other longitudinal studies of children using ASL, though occasional substitutions were found for reference to first and second person, they were not systematic (Jackson, 1989), just as in most cases of "pronominal reversal" in hearing children (Morgenstern, 2012).

The problem is that the pointing *signs* used for personal pronouns and demonstratives do not look very different from the pointing pre-linguistic or co-verbal *gestures* used by hearing children. Because "points are considered linguistic in the adult system, it is tempting to consider the child's points as linguistic" (Schick, 2003: 221). How are we to decide whether the nature of pointing is linguistic or "non-linguistic"/"pre-linguistic"? Most sign language researchers assume that these pointing signs are pronouns, but this assumption is challenged by Evans and Levinson (2009) and Cormier (In press). Pointing signs do not look different on the surface from pointing in non-signers (Kendon, 2004; Kita, 2003). In both cases, points index locations of objects, persons, events in the deictic space. Some studies, however, have focused on specific features of pointing in deaf children and on their ability to use different forms and types of pointing for different functions

(with the combinatorial dimensions of finger, wrist and arm configuration, movement, intensity and speed). Not only do deaf children use an impressive number of pointing gestures from very early on, but the functions of these points are “integrated into the process of conventionalization of gesture and control of the signing space” (Hoiting, 2009: 84).

In order to tackle the differences between French and French Sign language, I will compare the role of shared attention, gaze and pointing gestures in the interactions between two little girls, one deaf signing and one hearing-speaking, and their parents in their natural environment. We will not try to categorize points as being either “linguistic” or “gestural”. Symbolic gestures are part of our broad, integrative view of language as a multimodal interactive system and we will focus on the development of pointing within the communication systems that children develop, be they spoken or signed. Two studies conducted in the framework of the CoLaJE project financed by the French National Agency (Morgenstern, et al., 2010) are presented here in order to make an attempt at establishing comparisons between two longitudinal studies in two different situations. The researchers<sup>3</sup> created the coding system together in order to conduct comparable analyses<sup>4,5</sup>.

### Shared attention, gaze and pointing in two longitudinal follow-ups

#### *The data*

Madeleine<sup>6</sup> is a French monolingual hearing girl with two hearing monolingual parents, and a sister twelve years older<sup>7</sup>. She lives in Paris in an upper-middle class family. She was taken care of by a nanny until she entered kindergarten. Martine

3. The team on pointing gestures included at various stages of the project Emmanuelle Mathiot, Fanny Limousin, Marion Blondel, Dominique Boutet and Aliyah Morgenstern, the team on personal reference included Stéphanie Caët, Fanny Limousin, Marion Blondel and Aliyah Morgenstern.

4. The coding of Charlotte’s data was conducted by Fanny Limousin with the help of Stéphanie Caët and discussed during regular working sessions with Aliyah Morgenstern. The coding for Madeleine’s sessions was done by Marie Collombel-Leroy, Emmanuelle Mathiot and Aliyah Morgenstern.

5. Despite the use of video data, and of rich coding systems with video and transcription alignment using CLAN and ELAN, some specific analyses could not be conducted. We were not able to study the parents’ gaze for instance: the camera operator did not always film the adults with enough attention when the child was producing gestures, words or signs.

6. The data is part of the *Paris corpus* financed by the French Research Agency (ANR) in the framework of the *Léonard Project* directed by Aliyah Morgenstern and is available on CHILDES <<http://childe.psych.cmu.edu>>.

7. A brother was born after the end of the data analyzed in this paper.

Sekali filmed her for one hour nearly once a month from the age of ten months to the age of seven (Morgenstern & Parisse, 2012; Sekali, 2012).

Charlotte is a deaf girl raised by deaf middle-class parents who both use French Sign Language. She is their first child. She was filmed for one hour once a month from the age of seven months to three. Charlotte lives in Paris and attended a day-care center at the time with one deaf educator. She was filmed exclusively by Fanny Limousin, a deaf signing junior researcher<sup>8</sup>.

The two girls were quite precocious in their linguistic development and could be considered as quite comparable in the two modalities used. Their data has been analyzed for various studies focusing on prosody and morpho-syntax which gives us more insight into their linguistic development (Morgenstern & Sekali, 2009; Morgenstern, 2009; Limousin, 2011; Morgenstern & Parisse, 2012; Sekali, 2012). For this study, we used the data up to two years old.

### *Comparing children's acquisition of LSF and French*

We can observe the important differences between the two languages, due to the modalities used, in the interactions between Madeleine and Charlotte and their mothers.

One of the major differences between the mother/ child dyads we have studied lies in the use of GAZE. The eyes of Charlotte's mother are her essential link with her child and enable her to check how safe and well she is at all times. Her visual field is therefore wider than that of the hearing mother. However, when Charlotte wants to draw her mother's attention while she points, she makes intense movements using her head, legs and chest. She can also repeat the pointing gesture. Another of her strategies is to become totally still and gaze fixedly at her mother. Therefore the amplification or absence of movement is a marked form as opposed to normal gestures. Those strategies demonstrate how much Charlotte is aware of her mother's attention (or lack of it) and how she is able to manage and monitor gaze quite skillfully.

Another difference is that in oral language, sound, gesture and gaze are all extremely important in early communication, whereas in sign language, gesture and gaze are predominant. A child exposed to sign language might therefore be even more sensitive to gesture, and deaf adults are going to interpret (and over-interpret) their children's first gestures much earlier, just like hearing adults might do with babbling.

Because gaze and gestures are so central to communication, the articulation between daily activities and language is totally different. In the hearing dyad, the

---

8. See Limousin, 2011 for a detailed account of Charlotte's data.

mother is often doing other things while she talks to her child. In the deaf dyad, it is more difficult to communicate in sign language while you do a manual activity: you cannot sign with both your hands and your whole body, as required for Child Directed Sign, and efficiently cook, change a diaper, clean, sort papers at the same time. The interactional mode is therefore quite specific. The EYES replace the EARS and the mother is constantly « visually listening » to her child. The language moments are intense; both participants are focused on the other. Language is mostly a mono-activity. Of course, when children grow up, their interlocutors can rely on their child's experience in sign and use one hand or smaller movements to sign with them. The children themselves will acquire the ability to act and sign simultaneously.

There is another main difference between vocal and sign languages that seems to be an advantage for the explicit role of adults in the child's learning process: it is easier for a parent to modify the child's manual mode by acting on their hands, shaping and modeling them, than it is to rectify the oral mode. You cannot act on a child's vocal tract.

However, there are also apparent disadvantages in the use of visual languages: when the child is not focusing on interactions going on around her, she is cut off from non-addressed speech. We have numerous examples of Charlotte's mother signing to the observer out of Charlotte's visual field (behind her for example). In those cases, which are quite frequent, Charlotte has no access to this non-addressed speech whereas Madeleine constantly hears the adults speaking to each other and therefore has very frequent access to non-addressed speech.

#### *Talkativeness in the two children*

In order to measure whether these differences have an impact on the children's density of Speech/Sign, we performed a quantitative analysis of turn taking in the signing dyad and the speaking dyad. This showed that the number of turns exchanged in the hearing dyad per hour is twice the number of turns exchanged in the signing dyad in two different videos (Table 1).

**Table 1.** Number of turns produced by the children in an hour at 1;7 and 2;0 years.<sup>9</sup>

	1;7	2;0
Charlotte	134	152
Madeleine	285	395

9. Taken from Morgenstern, et al. (2010).

Madeleine's number of turns is greater than Charlotte's both at 1;07 and 2;0. Charlotte's turns increase by 13% in this period, whereas Madeleine's increase by 39%.

There seems to be an important impact of the different conversational styles and the constraints linked to the use of a visual language. With those differences in mind, we will now turn to the two girls' use of pointing.

### Pointing and gaze in Madeleine and Charlotte's data

#### *Quantitative use of pointing gestures in Madeleine and Charlotte's data*

One of our hypotheses was that Charlotte would produce more pointing gestures than Madeleine because of the nature of their input (French and gestures in Madeleine's case, Sign Language and gestures in Charlotte's case). In order to draw a comparison between the two girls, we extracted the total number of pointing gestures per one-hour session in our data (Figure 1).

For each child, the number of pointing gestures is very variable from one session to another according to the situations. Nevertheless, Charlotte produces pointing gestures three months earlier and more frequently in the 18 recordings (1187 in total) than Madeleine (465). The sessions when Madeleine uses a lot of pointing gestures (1;09 and 1;10) contain a lot of book-reading situations. The frequency of pointing gestures Charlotte produces increases irregularly but gradually between 7 months and 2 years.

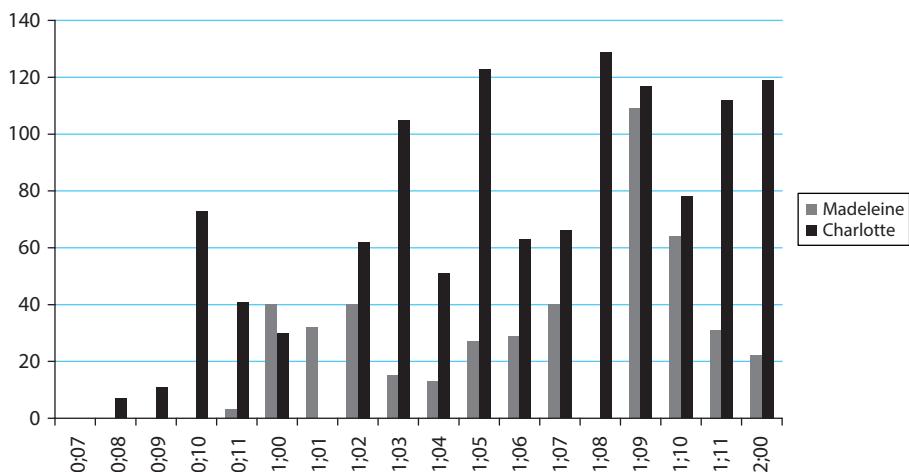


Figure 1. Number of pointing gestures produced in an hour for the two children according to age.

Since Charlotte's "communicative" turns are two times less frequent than Madeleine's, this high number of occurrences of pointing gestures in the same sessions might be considered surprising. The quantity of pointing gestures per turn is four to five times denser in Charlotte's data. But this could be linked to the high frequency of points and their grammaticalization in the input language, French Sign Language. In Charlotte's data, the adults use an average of 200 points per hour session, whereas in Madeleine's session, there is an average of 40 points per hour session. The overall number of points per hour actually decreases in her input when Madeleine gets older and stays stable in Charlotte's input.

#### *Use of gaze during pointing events in Charlotte and Madeleine's data*

Here, we do not analyze the gestures on their own, but consider the combination of pointing gestures and all the complementing elements (gaze, vocal productions and words/signs) as *pointing events* (Leroy, Mathiot, & Morgenstern, 2009)<sup>10</sup>.

Gaze towards the adult is generally considered an essential sign of the child's intention to communicate. The name given in the literature for this phenomenon is "visual checking". In an experimental situation, Franco and Butterworth (1996) have observed that the association between the children's pointing gestures and gaze towards the adult develops with age. At 10 months, there is pointing with gaze on the object, then at 12 months, the child points then gazes at the adult, finally at 15 months, the child gazes at the adult and then points. But in their natural environment, there are situations when children do not look at the adults because their attention is already focused, for example when the child and the mother are looking at a book together. When Madeleine is 1;01, she and her mother point one after the other at different pictures in a book, but the child does not gaze at her mother at all and even has her back to her.

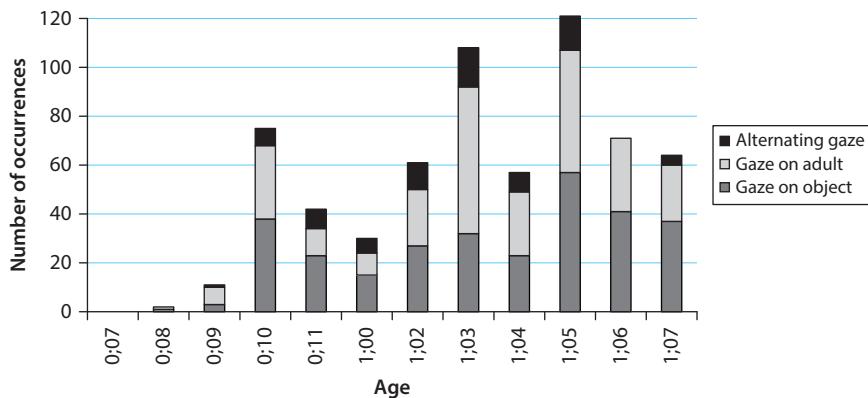
We also noticed that when Madeleine points without looking at her mother, she very often makes vocal productions, and even associates the pointing gesture to a recognizable two syllable word [œga] that seems to be a reproduction of her mother's very frequent use of "regarde!" (look) when she points. The combination of the auditory and the visual modality might not require the added use of visual checking as often. Charlotte does not use vocal productions with her points, at least not intentionally. Only one modality is involved. Since the two girls do not

<sup>10</sup>. The basic unit of pointing events is the pointing gesture. They always include gaze on the target or the adult. They often include synchronous vocal productions, words or other gestures/signs (sometimes produced with the other hand or even combined with the pointing gesture when the child is old enough to produce those complex combinations).

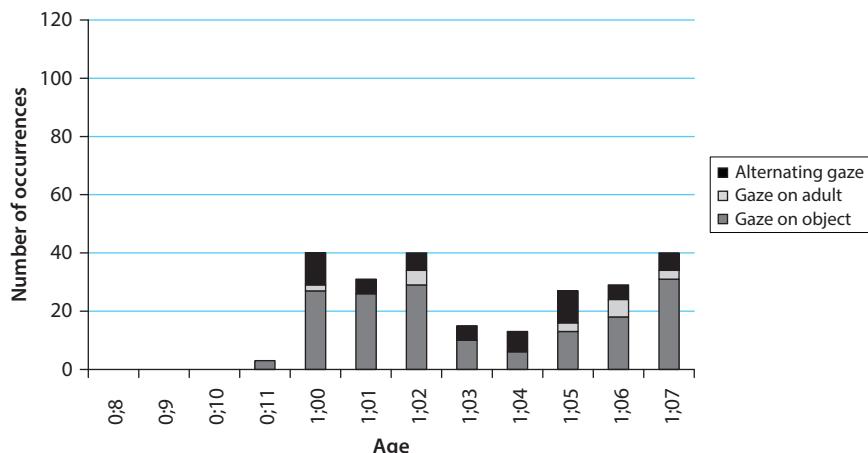
have the same attention getting tools at their disposal, will their use of gaze during pointing events be the same?

To answer that, we coded all the gazes occurring with pointing events in our longitudinal data from the beginning of the data to 1;07 (for technical reasons, we couldn't code the use of gaze in certain sessions after 1;07, see Figures 2 and 3).

Gaze and visual checking are used more by Charlotte (Figure 2), than by Madeleine (Figure 3). The overall proportion of gazes is shown in Table 2.



**Figure 2.** Number of occurrences of gaze on object, gaze on adult and alternating gaze during Charlotte's pointing events.



**Figure 3.** Number of occurrences of gaze on object, gaze on adult and alternating gaze during Madeleine's pointing events.

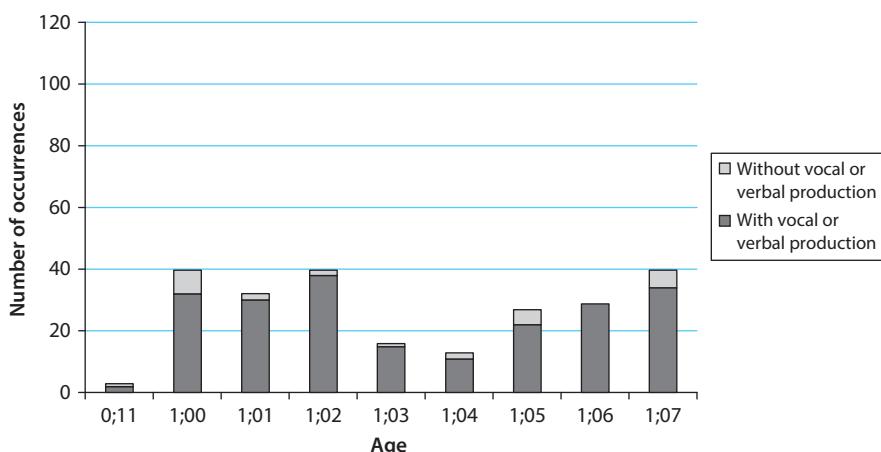
**Table 2.** Number of gaze on object, gaze on adult and alternating gaze in Madeleine and Charlotte's pointing events.

	Gaze on object	Gaze on adult	Alternating gaze	Total
Charlotte	252 (39%)	290 (44%)	110 (17%)	643
Madeleine	175 (73%)	19 (8%)	44 (17%)	238

The proportion of gazes on the adult in Madeleine's pointing events is significantly lower than in Charlotte's pointing events ( $\chi^2(1) = 57.6, p < 0.00001$ ) and conversely the proportion of gazes on the object in Madeleine's pointing events is significantly higher ( $\chi^2(1) = 25.3, p < 0.00001$ ). We note that for Charlotte, gazes on objects and adults occur in roughly equal proportions (39% vs. 44%) whereas most of Madeleine's gazes are on objects. Alternating gazes are used in the same proportions by both girls ( $\chi^2(1) = 0.09, p = 0.76$ ). Madeleine, on the other hand, can use another modality to catch the adult's attention: the auditory modality.

#### *Use of vocal and verbal productions during pointing events in Madeleine's data*

When we counted the number of pointing gestures that were produced simultaneously with vocal productions, we found very high frequencies as shown in Figure 4 (in dark grey).



**Figure 4.** Number of pointing events with and without vocal productions in Madeleine's data.

It is quite clear that silent points are rare in Madeleine's data, they represent 13% of all pointing events: in fact they often occur as a reaction to an adult's question to localize an object. Vocalizing while pointing seems to be the rule for Madeleine and it may well serve the aim of getting and sharing attention as much as adding supplementary or complementary proto-words or words to the pointing gesture. For example Madeleine at 1;05 says [sisis] as she points to the CD player (meaning *musique/music*) or says [vave] (*laver/wash*) as she points at a stain on her doll's head. The association between pointing and vocalization or verbalization is found 100% of the time in experimental situations according to Franco and Butterworth (1996). Guidetti (2003) compared pointing with other types of gesture, like "bye bye" and observed that more of these types of gesture are produced alone whereas pointing gestures are massively used with vocal or verbal productions.

We also observed that as she gets older, Madeleine uses fewer pointing gestures when she produces deictics, as shown in Figure 5.

Apart from locating an object in space in a very precise way, which she still does at four years old (she points at a stain and says "je me suis tâchée là" /I stained myself here), Madeleine progressively replaces referential pointing gestures by deictics<sup>11</sup>. She avoids redundant information but also talks more often about absent objects or events she can't directly point at. In French Sign Language, points

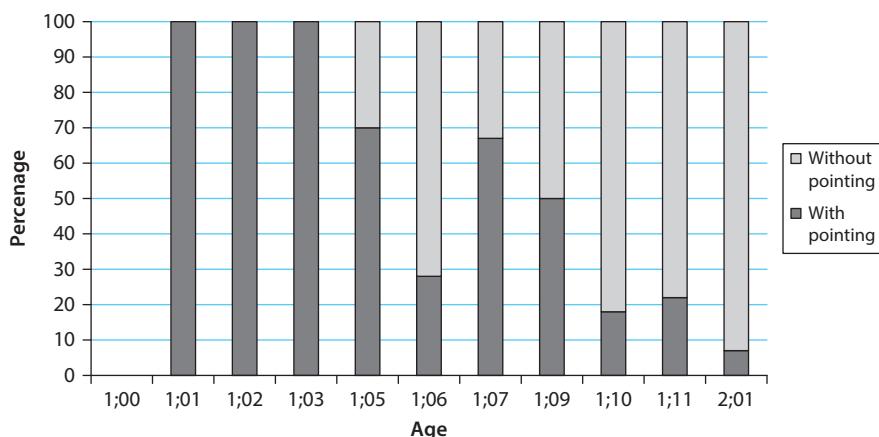


Figure 5. Number of deictics used with and without pointing gestures in Madeleine's data.

11. However, Madeleine continues to use co-verbal pointing gestures with a whole range of diversified functions as her communicative skills become more and more complex. She demonstrates excellent mastery of the location, the orientation, the motion of her pointing gestures, which enables her to mark subtle differentiation of their functions. She uses pointing for example to refer to time-spans or to attenuate, to suspend the predication she is making in speech.

are grammaticalized into deictics (personal and demonstratives pronouns in particular) and those grammatical markers are used just as frequently as their equivalent in French.

## Conclusions

We found that in the hearing child's data, gaze on the object could be considered as being non-marked and gaze on the adult as being marked. The hearing child only uses the latter when in the situation pointing plus vocal production are not enough to attract the adult's attention. Our hypothesis is that in oral languages, at least in French, vocal and verbal productions are the dominant modality to attract attention. Vocal productions are thus used to ATTRACT attention, pointing gestures enable the child to DIRECT attention, and gaze is only used to CHECK that the attention is there when there is a clear doubt or if the target is extremely important for the child. In the case of the deaf child in a signing environment, she will use a combination of gaze + pointing in order to ATTRACT, DIRECT and CHECK the adult's attention. Gaze is therefore used much more frequently and more expertly both to check and direct the adult's attention. Deaf signing children get precocious specific training in the use of gaze. In their daily communication, whenever they ask a question, for example, deaf children must keep their gaze on their interlocutor if they want to see the reply. It is not the case of the hearing child who can go back to her activity and still catch the language addressed (or not addressed) to her.

Because of its visual nature, in sign language acquisition, gaze and pointing are clearly privileged. At two years old, Charlotte has had extensive expertise in using gaze with her adult interlocutors. She is fully aware of the adults' attention and goes near them if she clearly wants to communicate with them, whereas Madeleine can also increase the intensity of her voice to attract adults' attention. Charlotte even sometimes lifts her arms in order to reach the adults' visual field or bends to sign for her dolls as adults do with her. She also uses a much greater number of points than Madeleine who progressively replaces certain pointing gestures by words (deictics, nouns, predicates).

Monitoring and checking are basic ingredients of communicative acts. Before they monitor, check and repair their speech (Clark, 1982), children learn to monitor and check adults' attention. This is already in place at a very early age thanks to pointing and gaze, which are two of the main instruments adults and children use to manage the attention of the interlocutor. Hearing children use vocal productions to complement pointing and gaze and therefore seem to rely on them less than deaf children.

The comparison between deaf and hearing children gives us some insights into the nature/nurture debate. Deaf signing children do not benefit from the use of hearing and do not develop the same subtle use of their vocal tract. But they develop an expertise at using the visual modality, which hearing children also have the capacity to develop, but do not need to depend solely on, since they also resort to the auditory modality. Unfortunately, not all deaf children are surrounded by expert users of the visual modality. A lot do not benefit from the necessary model and scaffolding early enough to learn through the use of gaze and pointing to share attention and make links between objects, persons, events, affects and language. But those who do benefit from it then have a solid base on which to become expert signers.

Despite great individual differences, and many possible paths, deaf and hearing children alike use all their capacities and all the input at their disposal, to master the greatest of social arts – language.

## References

- Baldwin, D. (1993). Infants' ability to consult the speaker for clues to word reference. *Journal of Child Language*, 20, 395–419. DOI: 10.1017/S0305000900008345
- Bates, E., Camaioni L., & Volterra, V. (1975). The acquisition of performatives prior to speech. *Merrill-Palmer Quarterly*, 21, 205–226.
- Bellugi, U., & Klima, E.S. (1982). The acquisition of three morphological systems in American Sign Language. *Papers and Reports on Child Language Development*, 21, 1–35.
- Brazelton, T.B., Koslowski, B., & Main, M. (1974). The origins of reciprocity; The early mother-infant interaction. In M. Lewis & L. Rosenblum (Eds.), *The effect of the infant on its caregiver* (pp. 49–76). New York, NY: Wiley-Interscience.
- Bruner, J.S. (1975). From communication to language: A psychological perspective. *Cognition*, 3, 255–287. DOI: 10.1016/0010-0277(74)90012-2
- Bruner, J. (1983). *Child's talk: Learning to use language*. New York, NY: Norton.
- Butterworth G., & Itakura, S. (2000). How the eyes, head and hand serve definite reference. *British Journal of Developmental Psychology*, 18(1), 25–50. DOI: 10.1348/026151000165553
- Clark, E.V. (1978). From gesture to word: On the natural history of deixis in language acquisition. In J.S. Bruner & A. Garton (Eds.), *Human growth and development: Wolfson College lectures 1976* (pp. 85–120). Oxford: OUP.
- Clark, E.V. (1982). Language change during language acquisition. In M.E. Lamb & A.L. Brown (Eds.), *Advances in developmental psychology*, Vol. 2 (pp. 173–197). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Cole, M., & Cole, S.A. (1996). *The development of children*. New York, NY: Scientific American Books.
- Cormier, K. (2010) Pronouns and pointing: Where do sign languages fit in? In the Proceedings of *Between you and me: Local pronouns across modalities*. Nijmegen: Radboud University.
- Courtin, C. (2000). The impact of sign language on the cognitive development of deaf children. The case of theories of mind. *Journal of Deaf Studies and Deaf Education*, 5(3), 266–276. DOI: 10.1093/deafed/5.3.266

- Cuxac, C. (2000). *La langue des signes française (LSF). Les voies de l'iconicité*. Paris: Ophrys.
- Darwin, C. (1877). A biographical sketch of an infant. *Mind*, 2, 285–294. DOI: 10.1093/mind/os-2.7.285
- Deák, G., Jasso, H., Krasno, A., & Triesch, J. (2006). Watch the hands: Infants learn gaze-following from parents' pointing and manual action. Paper presented at the XVth Biennial International Conference on Infant Studies, Westin Miyako, Kyoto, Japan.
- Deák, G.O., & Triesch, J. (2006). Origins of shared attention in human infants. In K. Fujita & S. Itakura (Eds.), *Diversity of cognition* (pp. 331–363). Kyoto: University of Kyoto Press.
- Estigarribia, B., & Clark, E.V. (2007). Getting and maintaining attention in talk to young children. *Journal of Child Language*, 34, 799–814. DOI: 10.1017/S0305000907008161
- Evans, N., & Levinson, S.C. (2009). The myth of language universals: Language diversity and its importance for cognitive science. *Behavioral and Brain Sciences*, 32, 429–492. DOI: 10.1017/S0140525X0999094X
- Filipi, A. (2009). *Toddler and parent interaction: The organization of gaze, pointing and vocalization*. Amsterdam: John Benjamins. DOI: 10.1075/pbns.192
- Flom, R., Deak, G., Phill, C.G., & Pick, A.D. (2003). Nine-month-olds' shared visual attention as a function of gesture and object location. *Infant Behavior and Development*, 27, 181–194. DOI: 10.1016/j.infbeh.2003.09.007
- Franco F., & Butterworth, G. (1996). Pointing and social awareness: Declaring and requesting in the second year. *Journal of Child Language*, 23, 307–336. DOI: 10.1017/S0305000900008813
- Guidetti, M. (2003). *Pragmatique et psychologie du développement: Comment communiquent les jeunes enfants*. Paris: Belin.
- Hoiting, N. (2009). *The myth of simplicity. Sign language acquisition by Dutch deaf toddlers*. Unpublished doctoral dissertation. University of Groningen.
- Hoiting, N., & Slobin, D.I. (2007). From gestures to signs in the acquisition of sign language. In S.D. Duncan, J. Cassell, & E.T. Levy (Eds.), *Gesture and the dynamic dimension of language: Essays in honor of David McNeill* (pp. 51–65). Amsterdam: John Benjamins.
- Iverson, J., & Goldin-Meadow, S. (2005). Gesture paves the way for language development. *Psychological Science*, 16(5), 367–371. DOI: 10.1111/j.0956-7976.2005.01542.x
- Kelly, B.F. (2011). A new look at redundancy in children's gesture and word combinations. In I. Arnon & E.V. Clark (Eds.), *Experience, variation, and generalization: Learning a first language* (pp. 73–90). Amsterdam: John Benjamins.
- Kendon, A. (2004). *Gesture: Visible action as utterance*. Cambridge: CUP.
- Kita, S. (Ed.). (2003). *Pointing: Where language, culture and cognition meet*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Leroy, M., Mathiot, E., & Morgenstern, A. (2009) Pointing gestures and demonstrative words: Deixis between the ages of one and three. In J. Zlatev, M.J. Falck, C. Lundmark, & M. Andrén (Eds.), *Studies in language and cognition* (pp. 386–404). Newcastle upon Tyne: Cambridge Scholars.
- Limousin, F. (2011). *Acquisition de la référence personnelle en LSF : Analyse longitudinale des pointages, des formes nulles et des noms signés chez une enfant sourde de parents sourds*. Unpublished doctoral dissertation. Université Paris 8.
- Masataka, N. (1996). Perception of motherese in a signed language by 6-month-old deaf infants. *Developmental Psychology*, 32, 874–879. DOI: 10.1037/0012-1649.32.5.874
- Mayberry, R.I., & Squires, B. (2006). Sign language: Acquisition. In E. Lieven (Ed.), *Language acquisition, Vol. 11: Encyclopedia of language and linguistics, 2nd ed.* Oxford: Elsevier.

- McNeill, D. (1992). *Hand and mind: What gestures reveal about thought*. Chicago, IL: University of Chicago Press.
- Meristo, M., Morgan, G., Geraci, A., Iozzi, L., Hjelmquist, E., Surian, L., & Siegal, M. (In press). Belief attribution in deaf and hearing infants. *Developmental Science*.
- Morgenstern, A. (1997). L'enfant sourd énonciateur-signeur: l'Auto-désignation chez l'enfant en Langue des Signes Française. *LIDIL*, 15, 119–140.
- Morgenstern, A. (2009). *L'enfant dans la langue*. Paris: Presses de la Sorbonne Nouvelle.
- Morgenstern, A., Caet, S., Collombel-Leroy, M., Limousin, F., & Blondel, M. (2010). From gesture to sign and from gesture to word: Pointing in deaf and hearing children. *Gesture*, 10, 172–202. DOI: 10.1075/gest.10.2-3.04mor
- Morgenstern, A., & Parisse, C. (2012). The Paris Corpus. *Journal of French Language Studies* 22(1), 7–12. DOI: 10.1017/S095926951100055X
- Morgenstern, A., & Sekali, M. (2009). What can child language tell us about prepositions? A contrastive corpus-based study of cognitive and social-pragmatic factors. In J. Zlatev, M.J. Falck, C. Lundmark, & M. Andrén (Eds.), *Studies in language and cognition* (pp. 261–271). Newcastle upon Tyne: Cambridge Scholars.
- Petitto, L.A. (1986). *From gesture to symbol: the relationship between form and meaning in the acquisition of personal pronouns in American Sign Language*. Bloomington, IN: Indiana University Linguistics.
- Pizzutto, E., & Capobianco, M. (2005). The link and differences between deixis and symbols in children's early gestural-vocal systems. *Gesture*, 5(1), 175–195.
- Schick, B. (2003). The development of ASL and manually-coded English systems. In M. Marschark & P.E. Spencer (Eds.), *Oxford handbook of deaf studies, language, and education*. Oxford: OUP.
- Sekali, M. (2012). The emergence of complex sentences in a French child's language from 0;10 to 4;01: Causal adverbial clauses and the concertina effect. In M. Sekali (Ed.), First language acquisition of French grammar (from 10 months to 4 years old), special issue of *Journal of French Language Studies* 22(1), 115–141. DOI: 10.1017/S0959269511000615
- Stern, W. (1924). *Psychology of early childhood*, London: Unwin. (German edition 1914).
- Tomasello, M. (1999). *The cultural origins of human cognition*. Cambridge, MA: Harvard University Press.
- Tomasello, M. (2003). *Constructing a language: A usage-based theory of language acquisition*. Cambridge, MA: Harvard University Press.
- Tomasello, M., Carpenter, M., & Liszkowski, U. (2007). A new look at infant pointing. *Child Development*, 78, 705–722. DOI: 10.1111/j.1467-8624.2007.01025.x
- You, Y., Deak, G., Jasso, H., & Teuscher, D. (2005). Emergence of shared attention from 3 to 11 months of age in naturalistic infant-parent interactions. Presented at the *Biennial meeting of the Society for Research in Child Development*, Atlanta, GA.
- Werner, H., & Kaplan, B. (1963). *Symbol formation*. Hillsdale, NJ: Lawrence Erlbaum Associates.

# How gesture helps children learn language\*

Susan Goldin-Meadow  
University of Chicago

Children gesture before they begin to speak and continue gesturing throughout the language learning process. This chapter focuses on those gestures, and explores the role they play in language learning. We find that children's early gestures not only precede, but also predict, the onset of a number of linguistic milestones—nouns, nominal constituents, simple and complex sentences. Gesturing may thus play a causal role in language learning, and could do so in two ways: (1) Gesturing gives children the opportunity to practice expressing ideas in a preverbal form. (2) A child's gestures offer parents and other communication partners insight into the child's linguistic level, thus giving the partners the opportunity to provide input tailored to that level.

## A personal note

Eve Clark has been one of the most eloquent proponents of the view that language is learned on the ground—that is, in day-to-day interactions with parents and peers. She, more than any other researcher, has convinced the field that children learn language in interaction, and she has done so by describing in exquisite detail the properties, and the implications, of those interactions. Happily for the field of gesture researchers, gesture is one of the behaviors to which Eve has turned her attention. And she has found that gesture is a behavior adults use to inform young children about new word meanings. My focus in this chapter is on the gestures that

---

\* The work described in this paper was supported by a program project grant from NICHD (P01 HD 40605). I thank my fellow PIs on this project, Susan Levine, Janellen Huttenlocher, Stephen Raudenbush, Larry Hedges, and Steve Small, for the many contributions they have made to the project and to my thinking; Cynthia Butcher, Erica Cartmill, Dedre Gentner, Whitney Goodrich, Dea Hunsicker, Jana Iverson, Eve Sauer LeBarton, Susan Levine, Marolyn Morford, Şeyda Özçalışkan, and Meredith Rowe for being such terrific collaborators; Kristi Schonwald, Jodi Khan, and Jason Voigt for their unwavering administrative and technical support; and the children and their families who participated in the studies for welcoming us into their homes and lives.

children themselves use,<sup>1</sup> and whether they too can play a role in language learning. We will see that children's gestures not only precede, but also predict, their acquisition of linguistic milestones and, more importantly from the perspective that Eve has taken, children's gestures often let parents know exactly what input they need to hear to take the next step toward a linguistic milestone. As Eve would predict, gesture is one of the behaviors that can help children learn language on the ground.

### **Gesture selectively predicts different linguistic milestones**

At a time in development when children are limited in the words they know and use, gesture offers a way to extend their communicative range. Children typically begin to gesture between 8 and 12 months (Bates, 1976; Bates, Benigni, Bretherton, Camaioni, & Volterra, 1979), first producing deictic gestures (pointing at objects, people, and places in the immediate environment, or holding up objects to draw attention to them), and later producing iconic gestures that capture aspects of the objects, action or attributes they represent (e.g., flapping arms to refer to a bird or to flying, Iverson, Capirci, & Caselli, 1994). The fact that gesture allows children to communicate meanings that they do not yet express in speech opens up the possibility that gesturing itself facilitates language learning. If so, changes in gesture should not only predate, but they should also predict, changes in language. And they do, both for words and for sentences (see Table 1).

**Table 1.** Examples of linguistic constructions preceded and predicted by gesture.

Type of construction	Preceding Construction Containing Gesture	Subsequent Construction Containing Speech Alone
<u>Nouns</u>	point at bottle	“bottle”
<u>Nominal constituents</u>	point at dog + “dog”	“the dog”
<u>Simple sentences</u>		
Argument + Argument (entity-location)	“daddy” + point at dirt on floor	“mommy in kitchen”
Argument + Predicate (agent-act)	“you” + HIT gesture	“I paint”
Argument + Predicate (act-patient)	“drive” + point at car	“ride horsie”
<u>Complex sentences</u>		
Predicate + Predicate	“I like it” + EAT gesture	“help me find”

1. I focus here on the spontaneous gestures that children produce, rather than the codified gestures, called *baby signs*, that many parents teach their children before they are able to speak (Johnston, Durieux-Smith, & Bloom, 2005; Kirk, Howlett, Pine, & Fletcher, 2013).

### *Words: nouns and verbs*

The more a child gestures early on, the more words are likely to be present in the child's vocabulary later in development (Acredolo & Goodwyn, 1988; Rowe, Özçalışkan, & Goldin-Meadow, 2008; Rowe & Goldin-Meadow, 2009a; Colonna, Stams, Koster, & Noom, 2010). Even more compelling, we can predict which particular nouns will enter a child's verbal vocabulary by looking at the objects that child indicated using deictic gestures several months earlier (Iverson & Goldin-Meadow, 2005). For example, a child who does not know the word "dog," but communicates about dogs by pointing at them is likely to learn the word "dog" within three months (Iverson & Goldin-Meadow, 2005). Gesture paves the way for children's early nouns.

However, gesture does not appear to pave the way for early verbs – although we might have expected iconic gestures that depict actions to precede, and predict, the onset of verbs, they do not. Özçalışkan, Gentner, and Goldin-Meadow (2013) observed spontaneous speech and gestures in 40 English-learning children from age 14 to 34 months, and found that the children produced their first iconic gestures 6 months *later* than their first verbs. The onset of iconic gestures conveying action meanings thus followed, rather than preceded, children's first verbs<sup>2</sup>. But iconic gestures did increase in frequency at the same time that verbs did and, at that time, children used these action gestures to convey specific verb meanings that they were not yet expressing in speech. Children thus do use gesture to expand their repertoire of verb meanings, but only *after* they have begun to acquire the verb system underlying their language.

### *Onset of sentences*

Even though they treat gestures like words in some respects, children very rarely combine their gestures with other gestures, and if they do, the phase tends to be short-lived (Goldin-Meadow & Morford, 1985). But children do often combine their gestures with words, and they produce these gesture + speech combinations well before they produce word + word combinations. Children's earliest gesture + speech combinations contain gestures that convey information that *complements* the information conveyed in speech; for example, pointing at a ball while saying "ball" (Capirci, Iverson, Pizzuto, & Volterra, 1996; de Laguna, 1927; Greenfield & Smith, 1976; Guillaume, 1927; Leopold, 1939-49). Soon after,

2. Estigarribia and Clark (2007) have found that pointing gestures attract and maintain attention in talk differently from iconic (or, in their terms, demonstrating) gestures, which may account for the fact that pointing gestures predict the onset of nouns, but iconic gestures do not predict the onset of verbs.

children begin to produce combinations in which gesture conveys information that is different from and *supplements* the information conveyed in the accompanying speech; for example, pointing at a ball while saying “here” to request that the ball be moved to a particular spot (Goldin-Meadow & Morford, 1985; Greenfield & Smith, 1976; Masur, 1982, 1983; Morford & Goldin-Meadow, 1992; Zinober & Martlew, 1985).

As in the acquisition of words, we find that changes in gesture (in this case, changes in the relationship gesture holds to the speech it accompanies) predict changes in language (the onset of sentences). The age at which children first produce *supplementary* gesture + speech combinations (e.g., point at cup + “drink”) reliably predicts the age at which they first produce two-word utterances (e.g., “drink cup”) (Goldin-Meadow & Butcher, 2003; Iverson, Capirci, Volterra, & Goldin-Meadow, 2008; Iverson & Goldin-Meadow, 2005). The age at which children first produce *complementary* gesture + speech combinations (e.g., point at cup + “cup”) does not. Moreover, supplementary combinations selectively relate to the syntactic complexity of children’s later sentences. Rowe and Goldin-Meadow (2009b) observed 52 children from families reflecting the demographic range of Chicago and found that the number of supplementary gesture + speech combinations the children produced at 18 months reliably predicted the complexity of their sentences (as measured by the IPSyn, Scarborough, 1990) at 42 months, but the number of different meanings they conveyed in gesture at 18 months did not. Conversely, the number of different meanings children conveyed in gesture at 18 months reliably predicted their spoken vocabulary (as measured by the PPVT, Dunn & Dunn, 1997) at 42 months, but the number of supplementary gesture + speech combinations they produced at 18 months did not. Gesture is thus not merely an early index of global communicative skill, but is a harbinger of specific linguistic steps children will soon take—early gesture words predict later spoken vocabulary, and early gesture sentences predict later spoken syntax.

### *Onset of different constructions*

Gesture does more than open the door to sentence construction—the particular gesture + speech combinations children produce predict the onset of corresponding linguistic milestones. Özçalışkan and Goldin-Meadow (2005) observed 40 of the children in the Rowe and Goldin-Meadow (2009b) sample at 14, 18, and 22 months, and found that the types of supplementary combinations the children produced changed over time and, critically, presaged changes in their speech. For example, the children began producing “two-verb” complex sentences in gesture + speech combinations (“I like it” + EAT gesture) several months before they produced complex sentences entirely in speech (“help me find it”). Supplementary

gesture + speech combinations thus continue to provide stepping-stones to increasingly complex linguistic constructions.

As mentioned earlier, the age at which children first produce *complementary* gesture + speech combinations in which gesture indicates the object labeled in speech (e.g., point at cup + “cup”) does not reliably predict the onset of two-word utterances (Iverson & Goldin-Meadow, 2005), reinforcing the idea that it is the specific way in which gesture is combined with speech, rather than the ability to combine gesture with speech *per se* that signals the onset of future linguistic achievements. The gesture in a complementary gesture + speech combination has traditionally been considered redundant with the speech it accompanies but, as Clark and Estigarribia (2011) show, gesture typically locates the object being labeled and, in this sense, has a different function from speech. Complementary gesture + speech combinations have, in fact, recently been found to point forward—but to the onset of nominal constituents rather than the onset of sentential constructions. If children are using nouns to classify the objects they label (as recent evidence suggests infants do when hearing spoken nouns; Parise & Csibra, 2012), then producing a complementary point with a noun could serve to specify an instance of that category. In this sense, a pointing gesture could be functioning like a determiner. Cartmill, Hunsicker, and Goldin-Meadow (2014) analyzed all of the utterances containing nouns produced by 18 children in the Rowe and Goldin-Meadow (2009b) sample, and focused on (1) utterances containing an unmodified noun combined with a complementary pointing gesture (e.g., point at cup + “cup”), and (2) utterances containing a noun modified by a determiner (e.g., “the/a/that cup”). They found that the age at which children first produced complementary point + noun combinations selectively predicted the age at which the children first produced determiner + noun combinations. Not only did complementary point + noun combinations precede and predict the onset of determiner + noun combinations in speech, but these point + noun combinations also decreased in number once children gained productive control over determiner + noun combinations. When children point to and label an object simultaneously, they appear to be on the cusp of developing an understanding of nouns as a modifiable unit of speech.

Gesture has also been found to predict changes in narrative structure later in development. In Demir, Levine, and Goldin-Meadow (under review), 38 children from Rowe and Goldin-Meadow (2009b) were asked to retell a cartoon at age 5 and then again at ages 6, 7, and 8. Although their narrative structure continued to improve over the 4-year period, the children showed little evidence of framing their narratives from a character’s perspective in speech even at age 8. However, at age 5, many of the children were able to take a character’s viewpoint into account in gesture. For example, to describe a woodpecker’s actions, a child moves her upper body and head back and forth, thus assuming the perspective of the bird (as

opposed to moving a beak-shaped hand back and forth and thus taking the perspective of someone looking at the bird, cf. McNeill, 1992). Moreover, the children who produced character-viewpoint gestures at age 5 were more likely than children who did not produce character-viewpoint gestures to produce well-structured stories in the later years. Gesture thus continues to act as a harbinger of change as it assumes new roles in relation to discourse and narrative structure.

### *Fleshing out constructions*

Gesture does not, however, always predict transitions in language learning. Gesture precedes and predicts linguistic developments when those developments involve *new* constructions, but not when the developments involve fleshing out existing constructions. For example, Özçalışkan and Goldin-Meadow (2009) observed 40 children in the Rowe and Goldin-Meadow (2009b) sample and found that the children produced combinations in which speech conveyed a predicate and gesture conveyed an argument (e.g., WASH gesture + “hair” = predicate in gesture, object in speech) several months before they produced predicate + argument combinations entirely in speech (e.g., “popped this balloon” = predicate, object, both in speech). However, once the basic predicate + argument construction had been acquired in speech, the children did not rely on gesture to add arguments to the predicate frame. Thus, the children produced their first predicate + 2 argument combinations in speech (e.g., “I want the Lego” = agent, predicate, object, all in speech) and in gesture + speech (point at father + “have food” = agent in gesture, predicate and object in speech) at the same age (Özçalışkan & Goldin-Meadow, 2009).

Why does early gesture selectively predict later spoken vocabulary size and sentence complexity? At the least, gesture reflects two separate abilities (word learning and sentence making) on which later linguistic abilities can be built. Expressing many different meanings in gesture early in development is a sign that the child is going to be a good vocabulary learner, and expressing many different types of gesture + speech combinations is a sign that the child is going to be a good sentence learner. The early gestures children produce thus reflect their cognitive potential for learning particular aspects of language. But early gesture could be doing more—it could be helping children realize their potential. In other words, the act of expressing meanings in gesture could be playing an active role in helping children become better vocabulary learners, and the act of expressing sentence-like meanings in gesture + speech combinations could be playing an active role in helping children become better sentence learners. The next section explores this possibility.

## The mechanisms underlying gesture's role in language learning

### *Gesture provides opportunities to practice conveying meanings*

Child gesture could have an impact on language learning in at least two ways. First, gesture gives children an opportunity to practice producing particular meanings by hand at a time when those meanings are difficult to express by mouth. We know, for example, that early gesture use is related to later vocabulary size. In a mediation analysis, Rowe and Goldin-Meadow (2009a) found that the relatively large vocabularies children from high SES families display at 54 months can be partially explained by child gesture use at 14 months. In turn, child gesture use at 14 months can be explained by parent gesture use at 14 months, even when parent speech is controlled. Importantly, parent gesture does not appear to have a direct effect on subsequent child spoken vocabulary—the effect is mediated through child gesture, suggesting that it is the act of gesturing on the part of the child that is critical.

Although these findings suggest that child gesture is playing a causal role in language learning, we need to manipulate gesture to be certain of this claim. Previous work has found that telling 9- and 10-year-old children to gesture when explaining how they solved a math problem does, in fact, make them particularly receptive to subsequent instruction on that problem—the gesturing itself appears to be responsible for their improved performance after instruction (Broaders, Cook, Mitchell, & Goldin-Meadow, 2007). As another example more relevant to language learning, LeBarton, Raudenbush, and Goldin-Meadow (in press) studied 15 toddlers (beginning at 17 months) in an 8-week at-home intervention study (6 weekly training sessions plus follow-up 2 weeks later) in which all children were exposed to object words, but only some were told to point at the named objects. Before each training session and at follow-up, children interacted naturally with their parents to establish a baseline against which changes in communication were measured. Children who were told to gesture increased the number of gesture meanings they conveyed not only when interacting with the experimenter during training, but also when later interacting with their parents. Critically, these experimentally-induced increases in gesture led to larger spoken repertoires at follow-up. The findings suggest that gesturing can play an active role in word learning, perhaps because gesturing to a target picture in the context of labeling focuses children's attention to objects in the environment, to the labels, or to the object-label relation (Goldfield & Reznick, 1990; Werner & Kaplan, 1963). Children's active engagement in the bidirectional labeling context when told to gesture may draw their attention to gesture's communicative function, which could also have beneficial consequences for vocabulary development (Csibra & Gergely, 2009;

Tomasello, Carpenter, & Liszkowski, 2007; Woodward & Guajardo, 2002; Yoon, Johnson, & Csibra, 2008).

Although we know that encouraging children to point at objects enhances word learning, there have been no studies to date encouraging children to produce supplementary gesture + speech combinations. We thus know only that early supplementary gesture + speech combinations reflect the child's readiness to produce two-word utterances. More work is needed to determine whether these combinations play an active role in bringing about the onset of two-word utterances.

### *Gesture elicits timely speech from listeners*

The second way in which child gesture could play a role in language learning is more indirect--child gesture could elicit timely speech from listeners (see, for example, Kishimoto, Shizawa, Yasuda, Hinobayashi, & Minami, 2007). Because gesture seems to reflect a child's readiness for acquiring a particular linguistic structure, it has the potential to alert listeners (parents, teachers, clinicians) to the fact that a child is ready to learn that word or sentence. Listeners who pay attention to those gestures and can "read" them, might then adjust their talk, providing just the right input to help the child learn the word or sentence. Consider a child who does not yet know the word "rabbit" and refers to the animal by pointing at it. His obliging mother responds, "yes, that's a rabbit," thus supplying him with just the word he is looking for. Or consider a child who points at her mother while saying the word "hat." Her mother replies, "that's mommy's hat," thus translating the child's gesture + word combination into a simple sentence.

Just as mothers are sensitive to whether their children are familiar with the words they present, adjusting their strategies to make the word comprehensible (e.g., linking the new word to related words, offering terms that contrast with it directly, situating it by appealing to past experiences, Clark, 2010), mothers are sensitive to their children's gestures (Golinkoff, 1986; Masur, 1982). Mothers translate into their own words not only the single gestures that children produce (e.g., "that's a bird," produced in response to the child's point at a bird), but also the gestures that children produce in combination with words conveying different information, that is, supplementary gesture + speech combinations ("the bird's taking a nap," produced in response to the child's point at bird + "nap") (Goldin-Meadow, Goodrich, Sauer & Iverson, 2007). Interestingly, mothers produce longer sentences in response to their children's supplementary gesture + speech combinations (point at bird + "nap") than to their complementary gesture + speech combinations (point at bird + "bird"). Moreover, mothers' sentences tend to be longest when they pick up on information conveyed in child speech *and* gesture (e.g., "the bird's taking a nap"), despite the fact that they could easily have produced sentences

that are just as long when they pick up on information conveyed only in the child's speech ("It's time for your *nap*") or only in the child's gesture ("It's just like grandma's *bird*") or when they ignore the child's utterance entirely ("Let's read another book") (Goldin-Meadow et al., 2007).

If child gesture is playing an instrumental role in language learning, mothers' translations ought to be related to later word- and sentence-learning in their children--and they are (Goldin-Meadow et al., 2007). In terms of word-learning, when mothers translate the gestures that their children produce into words, those words are more likely to quickly become part of the child's vocabulary than words for gestures that mothers do not translate. In terms of sentence-learning, children whose mothers frequently translate their child's gestures into speech tend to be first to produce two-word utterances. The age at which children produce their first two-word utterance is highly correlated with the proportion of times mothers translate their child's gestures into speech, suggesting that mothers' targeted responses to their children's gestures might be playing a role in helping the children take their first steps into multiword combinations. Because they are finely-tuned to a child's current state (cf. Vygotsky's, 1986, zone of proximal development), adult responses of this sort could be particularly effective in teaching children how an idea is expressed in the language they are learning.

### Gesture as a diagnostic tool for language delay

Children with early right hemisphere lesions have been found to display delays in gesture early in development (Bates et al., 1997; Marchman et al., 1991). Are the children who exhibit delays in gesture the same children who exhibit delays in vocabulary development? We might expect delays in gesture use to go hand-in-hand with delays in language simply because, as we have seen here, gesture and language form an integrated system in typically developing children at the early stages of language learning. If the gesture-language system is robust in the face of early unilateral brain injury, children whose language development is proceeding at a typical pace should display typical gesture, and children whose language is delayed should display delays in gesture. Moreover, early gesture should predict subsequent language development, as it does in typically developing children. If this is the case, child gesture has the potential to serve as an early diagnostic tool, identifying which children will exhibit subsequent language delays, and which will catch up and thus fall within the normative range.

Children with pre- or perinatal unilateral brain lesions (PL) exhibit marked plasticity for language functions. Even when their lesions affect classic language areas, children with PL typically do not exhibit the aphasias that adults with similar

lesions display (e.g., Bates & Dick, 2002; Feldman, 2005). However, children with PL often exhibit delays in both productive and receptive language and these delays are transient for some children with PL, but persistent for others (Bates, Thal, & Janowsky, 1992; Feldman, Holland, Kemp, & Janosky, 1992). Can early child gesture be used to predict subsequent vocabulary development in children with PL?

Sauer, Levine and Goldin-Meadow (2010) categorized 11 children with PL into two groups based on whether their gesture use at 18 months was within or below the range for typically developing (data from 53 children were used to establish the range): (1) Children in the LOW group ( $n = 5$ ) fell below the 25th percentile for gesture production at 18 months in the typically developing group. (2) Children in the HIGH group ( $n = 6$ ) fell above the 25th percentile. Sauer and colleagues also charted the number of different words that the children with PL produced, again in relation to the number produced by the typically developing children. As a group, the children with PL produced more and more different words over time. However, there was a great deal of variability within the group. The question is whether this variability can be related to gesture use at 18 months.

Children with PL whose gesture use was within the range for typically developing children at 18 months, the HIGH group, developed a productive vocabulary at 22 and 26 months that was within the normative range, indeed close to the mean. In contrast, children with PL whose gesture use was below the range for the typically developing children at 18 months, the LOW group, remained outside of (and below) the normative range at both 22 and 26 months. The children displayed a similar pattern for receptive vocabulary (PPVT administered at 30 months). There was a significant correlation between the number of gesture types a child produced at 18 months and that child's PPVT score at 30 months. Importantly, speech at 18 months could not be used to predict children's later PPVT scores simply because there was very little variation in the number of speech types the PL children produced at 18 months—all of the children produced very few. These findings suggest that early gesture can predict subsequent spoken vocabulary, both productive and receptive, not only for children who are learning language at a typical pace, but also for those who are exhibiting delays.

Özcalışkan, Levine, and Goldin-Meadow (2013) found similar effects with respect to the onset of different types of sentence constructions in 11 children with PL, compared to 30 typically developing children. On average, children with PL produced their first instance of a two-argument sentence in gesture + speech ("mama" + point at stairs) four months later than the typically developing children, and the children with PL were comparably delayed in their first instance of a two-argument sentence conveyed entirely in speech ("turtle in truck"). The children with PL displayed the same pattern for argument + predicate sentences, producing them first in gesture + speech ("drink" + point at juice) and only later

entirely in speech (“pour the tea”), both at a 4-month delay relative to the typically developing children. However, the children with PL (unlike the typically developing children) did not reliably produce predicate + predicate constructions in gesture + speech (“I see” + GIVE gesture) before producing them entirely in speech (“I get zipper and zip this up”). Overall, the gesture-speech system appears to be a robust feature of language learning for simple – but not complex – sentence constructions, acting as a harbinger of change in language development even when that language is developing in an injured brain.

The paucity of particular gesture constructions in children with PL has the potential not only to serve as a diagnostic for later language delay, but also to contribute to those delays. First, the fewer gestures that children with PL produce, the fewer opportunities they have to practice communicating ideas they cannot yet express in speech. Second, the relatively small number of single gestures and gesture + speech combinations that children with PL produce provides parents with fewer opportunities to “translate” gesture into words, and thus fewer opportunities to tailor their input to what’s currently on the child’s mind. A good strategy for interacting with children with PL (and even typically developing children) is to encourage them to gesture as they talk. Augmenting child gesture will increase opportunities for child practice and for parents to fine-tune their input to the child’s state. Moreover, encouraging children to gesture has the potential to bring out previously unspoken ideas--once in the child’s production repertoire, those ideas can take hold and facilitate learning (cf. Broaders et al., 2007).

## Conclusion

We know that children come to language learning with ideas about how communication ought to be structured--for example, deaf children who are not exposed to a usable model of a conventional language nevertheless invent communication systems that have many of the fundamental properties of natural language (Goldin-Meadow, 2003; 2005). The challenge for children who *are* exposed to a conventional language is to mesh the ideas they bring to language learning with the details of the particular language that they are actually learning. It is here where gesture can play a seminal role. Child gesture gives children the opportunity to practice expressing ideas in a preverbal form, ideas that are packaged in words or in sentences. Equally important, child gesture gives parents the opportunity to (literally) see ideas that the child is on the cusp of expressing in speech--parents can, for example, see when their child is ready to acquire sentences by watching the child’s hands while listening to her words. By translating the gestures children produce into speech, parents can provide linguistic input that is tailored to the child’s state,

thus taking advantage of a “teachable” moment. In the tradition of Eve Clark’s research, gesture thus provides another example of how language is learned in the details of interaction (cf. Clark & Amaral, 2010).

## References

- Acredolo, L.P., & Goodwyn, S.W. (1988). Symbolic gesturing in normal infants. *Child Development*, 59, 450–466. DOI: 10.2307/1130324
- Bates, E. (1976). *Language and context: The acquisition of pragmatics*. New York, NY: Academic Press.
- Bates, E., Benigni, L., Bretherton, I., Camaioni, L., & Volterra, V. (1979). *The emergence of symbols: Cognition and communication in infancy*. New York, NY: Academic Press.
- Bates, E., & Dick, F. (2002). Language, gesture, and the developing brain. *Developmental Psychobiology*, 40, 293–310. DOI: 10.1002/dev.10034
- Bates, E., Thal, D., & Janowsky, J. (1992). Early language development and its neural correlates. In S. Segalowitz & I. Rapin (Eds.), *Handbook of neuropsychology*, Vol. 7: *Child neuropsychology* (pp. 69–110). Amsterdam: Elsevier.
- Bates, E., Thal, D., Trauner, D., Fenson, J., Aram, D., Eisele, J., & Nass, R. (1997). From first words to grammar in children with focal brain injury. *Developmental Neuropsychology*, 13, 275–343. DOI: 10.1080/87565649709540682
- Broaders, S., Cook, S.W., Mitchell, Z., & Goldin-Meadow, S. (2007). Making children gesture brings out implicit knowledge and leads to learning. *Journal of Experimental Psychology: General*, 136(4), 539–550. DOI: 10.1037/0096-3445.136.4.539
- Capirci, O., Iverson, J.M., Pizzuto, E., & Volterra, V. (1996). Communicative gestures during the transition to two-word speech. *Journal of Child Language*, 23, 645–673. DOI: 10.1017/S0305000900008989
- Cartmill, E., Hunsicker, D., & Goldin-Meadow, S. (2013). Pointing and naming are not redundant: Children use gesture to modify nouns before they modify nouns in speech. *Developmental Psychology*. 2014, DOI: 10.1037/a0036003
- Clark, E.V. (2010). Adult offer, word-class, and child uptake in early lexical acquisition. *First Language*, 30, 250–269. DOI: 10.1177/0142723710370537
- Clark, E.V., & Amaral, P.M. (2010). Children build on pragmatic information in language acquisition. *Language & Linguistics Compass*, 4(7), 445–457. DOI: 10.1111/j.1749-818X.2010.00214.x
- Clark, E.V., & Estigarribia, B. (2011). Using speech and gesture to introduce new objects to young children. *Gesture*, 11(1), 1–23. DOI: 10.1075/gest.11.1.01cla
- Colonnesi, C., Stams, G.J.J.M., Koster, I., & Noom, M.J. (2010). The relation between pointing and language development: A meta-analysis. *Developmental Review*, 30, 352–366. DOI: 10.1016/j.dr.2010.10.001
- Csibra, G., & Gergely, G. (2009). Natural pedagogy. *Trends in Cognitive Sciences*, 13, 148–153. DOI: 10.1016/j.tics.2009.01.005
- de Laguna G. (1927). *Speech: Its function and development*. Bloomington, IN: Indiana University Press.

- Demir, O.E., Levine, S., & Goldin-Meadow, S. (Under review). A tale of two hands: Development of narrative structure in children's speech and gesture.
- Dunn, L.M., & Dunn, L.M. (1997). *Peabody picture vocabulary test, 3rd ed.* Circle Pines, MN: American Guidance Service.
- Estigarribia, B., & Clark, E.V. (2007). Getting and maintaining attention in talk to young children. *Journal of Child Language*, 34, 799–814. DOI: 10.1017/S0305000907008161
- Feldman, H.M. (2005). Language learning with an injured brain. *Language Learning and Development*, 1(3–4), 265–288.
- Feldman, H.M., Holland, A.L., Kemp, S.S., & Janosky, J.E. (1992). Language development after unilateral brain injury. *Brain and Language*, 42(1), 89–102. DOI: 10.1016/0093-934X(92)90058-M
- Goldfield, B.A., & Reznick, J.S. (1990). Early lexical acquisition — Rate, content, and the vocabulary spurt. *Journal of Child Language*, 17, 171–183. DOI: 10.1017/S0305000900013167
- Goldin-Meadow, S. (2003). *The resilience of language: What gesture creation in deaf children can tell us about how all children learn language.* New York, NY: Psychology Press.
- Goldin-Meadow, S. (2005). What language creation in the manual modality tells us about the foundations of language. *Linguistic Review*, 22, 199–225. DOI: 10.1515/tlir.2005.22.2-4.199
- Goldin-Meadow, S., & Butcher, C. (2003). Pointing toward two-word speech in young children. In S. Kita (Ed.), *Pointing: Where language, culture, and cognition meet.* Hillsdale, NJ: Lawrence Erlbaum Associates.
- Goldin-Meadow, S., Goodrich, W., Sauer, E., & Iverson, J. (2007). Young children use their hands to tell their mothers what to say. *Developmental Science*, 10, 778–785. DOI: 10.1111/j.1467-7687.2007.00636.x
- Goldin-Meadow, S., & Morford, M. (1985). Gesture in early child language: Studies of deaf and hearing children. *Merrill-Palmer Quarterly*, 31, 145–176.
- Golinkoff, R. (1986). I beg your pardon?: the preverbal negotiation of failed messages. *Journal of Child Language*, 13, 455–476. DOI: 10.1017/S0305000900006826
- Greenfield, P., & Smith, J. (1976). *The structure of communication in early language development.* New York, NY: Academic Press.
- Guillaume P. (1927). Les débuts de la phrase dans le langage de l'enfant. *Journal de Psychologie*, 25, 1–25.
- Iverson, J.M., Capirci, O., & Caselli, M.S. (1994). From communication to language in two modalities. *Cognitive Development*, 9, 23–43. DOI: 10.1016/0885-2014(94)90018-3
- Iverson, J.M., Capirci, O., Volterra, V., & Goldin-Meadow, S. (2008). Learning to talk in a gesture-rich world: Early communication of Italian vs. American children. *First Language*, 28, 164–181. DOI: 10.1177/0142723707087736
- Iverson, J.M., & Goldin-Meadow, S. (2005). Gesture paves the way for language development. *Psychological Science*, 16, 368–371. DOI: 10.1111/j.0956-7976.2005.01542.x
- Johnston, J.C., Durieux Smith, A., & Bloom, K. (2005). Teaching gestural signs to infants to advance child development: A review of the evidence. *First Language*, 25(2), 235–251. DOI: 10.1177/0142723705050340
- Kirk, E., Howlett, N.R., Pine, K.J., & Fletcher, B. (2013). To sign or not to sign? The impact of encouraging infants to gesture on infant language and maternal mind-mindedness. *Child Development*.

- Kishimoto, T., Shizawa, Y., Yasuda, J., Hinobayashi, T., & Minami, T. (2007). Do pointing gestures by infants provoke comments from adults? *Infant Behavior and Development*, 30, 562–567. DOI: 10.1016/j.infbeh.2007.04.001
- LeBarton, E.S., Raudenbush, S., & Goldin-Meadow, S. (in press). *Journal of Cognition and Development*. Experimentally-induced increases in early gesture lead to increases in spoken vocabulary.
- Leopold, W.F. (1939–1949). *Speech development of a bilingual child: A linguist's record*, Volumes 1–4. Evanston, IL: Northwestern University Press.
- Marchman, V., Miller, R., & Bates, E. (1991). Babble and first words in children with focal brain injury. *Applied Psycholinguistics*, 12, 1–22. DOI: 10.1017/S0142716400009358
- Masur, E.F. (1982). Mothers' responses to infants' object-related gestures: Influences on lexical development. *Journal of Child Language*, 9, 23–30. DOI: 10.1017/S0305000900003585
- Masur, E.F. (1983). Gestural development, dual-directional signaling, and the transition to words. *Journal of Psycholinguistic Research*, 12, 93–109. DOI: 10.1007/BF01067406
- McNeill, D. (1992). *Hand and mind: What gestures reveal about thought*. Chicago, IL: University of Chicago Press.
- Morfitt, M., & Goldin-Meadow, S. (1992). Comprehension and production of gesture in combination with speech in one-word speakers. *Journal of Child Language*, 19, 559–580. DOI: 10.1017/S0305000900011569
- Özçalışkan, S., Gentner, D., & Goldin-Meadow, S. (2013). Do iconic gestures pave the way for children's early verbs? *Applied Psycholinguistics*. DOI: 10.1017/S0142716412000720.
- Özçalışkan, S., & Goldin-Meadow, S. (2005). Gesture is at the cutting edge of early language development. *Cognition*, 96(3), B101–B113. DOI: 10.1016/j.cognition.2005.01.001
- Özçalışkan, S., & Goldin-Meadow, S. (2009). When gesture-speech combinations do and do not index linguistic change. *Language and Cognitive Processes*, 24, 190–217. DOI: 10.1080/01690960801956911
- Özçalışkan, S., Levine, S., & Goldin-Meadow, S. (2013). Gesturing with an injured brain: How gesture helps children with early brain injury learn linguistic constructions. *Journal of Child Language*, 40(5), 69–105.
- Parise, E., & Csibra, G. (2012). Electrophysiological evidence for the understanding of maternal speech by 9-month-old infants. *Psychological Science*, 23(7), 728–733. DOI: 10.1177/0956797612438734
- Rowe, M., & Goldin-Meadow, S. (2009a). Differences in early gesture explain SES disparities in child vocabulary size at school entry. *Science*, 323, 951–953. DOI: 10.1126/science.1167025
- Rowe, M., & Goldin-Meadow, S. (2009b). Early gesture selectively predicts later language learning. *Developmental Science*, 12(1), 182–187. DOI: 10.1111/j.1467-7687.2008.00764.x
- Rowe, M., Özçalışkan, S., & Goldin-Meadow, S. (2008). Learning words by hand: Gesture's role in predicting vocabulary development. *First Language*, 28(2), 185–203. DOI: 10.1177/0142723707088310
- Sauer, E., Levine, S.C., & Goldin-Meadow, S. (2010). Early gesture predicts language delay in children with pre- or perinatal brain lesions. *Child Development*, 81, 528–539. DOI: 10.1111/j.1467-8624.2009.01413.x
- Scarborough, H.S. (1990). Index of productive syntax. *Applied Psycholinguistics*, 11, 1–22. DOI: 10.1017/S0142716400008262
- Tomasello, M., Carpenter, M., & Liszkowski, U. (2007). A new look at infant pointing. *Child Development*, 78, 705–722. DOI: 10.1111/j.1467-8624.2007.01025.x

- Vygotsky, L. (1986). *Thought and language* (A. Kozulin, translation). Cambridge, MA: The MIT Press. (Original work published 1934).
- Werner, H., & Kaplan, B. (1963). *Symbol formation*. New York, NY: Wiley.
- Woodward, A., & Guajardo, J. (2002). Infants' understanding of the point gesture as an object directed action. *Cognitive Development*, 17, 1061–1084. DOI: 10.1016/S0885-2014(02)00074-6
- Yoon, J., Johnson, M., & Csibra, G. (2008). Communication-induced memory biases in preverbal infants. *Proceedings of the National Academy of Sciences of the United States of America*, 105, 13690–13695. DOI: 10.1073/pnas.0804388105
- Zinober, B., & Martlew, M. (1985). Developmental changes in four types of gesture in relation to acts and vocalizations from 10 to 21 months. *British Journal of Developmental Psychology*, 3, 293–306. DOI: 10.1111/j.2044-835X.1985.tb00981.x



PART 3

**Pragmatic forces in language  
learning (six papers)**



# Referential pacts in child language development

Gemma Stephens and Danielle Matthews

Department of Psychology, University of Sheffield

Referential pacts are temporary conventions created by interlocutors for the duration of a conversation. They occur when a speaker somewhat arbitrarily chooses between multiple possible referring expressions (e.g., saying “*the spotty dog*” rather than “*the muddy dog*” for a dog that is both spotty and muddy). A listener will expect the speaker to be consistent in this choice of expression from then on, but would not necessarily expect a new conversational partner to make the same choice. We review the adult psycholinguistic literature on referential pacts and present evidence from children. We consider pacts as a prime example of the Principles of Contrast and Convention and suggest further research is warranted into the origin of these assumptions about language use.

It's hard to think how one could begin to explain language acquisition without appeal to the Principles of Contrast and Convention (E.V. Clark, 1987, 2007). Children quickly come to assume that speakers will use the same linguistic form to convey the same meaning across time and that any departure from this consistency marks a change in intended meaning. These assumptions are incredibly powerful not just in word learning (E.V. Clark, 1993; Diesendruck, 2005) but for language use and language learning across the board (E.V. Clark, 2007). In celebration of this fundamental insight, and of Eve Clark's pioneering work in giving pragmatics a central role in language acquisition, the following chapter explores one example of children's early awareness of the conventional nature of linguistic interaction: children's adherence to referential pacts.

## Referential pacts

When adults enter into dialogue, they implicitly collaborate in converging upon stable referring expressions that pick out things in the world from a certain

perspective. To illustrate, imagine two friends were conversing about constructing an item of flat-packed furniture. Confronted with typically ambiguous instructions containing only pictorial diagrams to represent the array of disconnected pieces in the box, it is safe to assume that if one friend was to refer to a particular piece as “the heavy circle”, the other would too, thus creating a referential pact (still implicit) between the two, stipulating that they refer to the object using only this term. Now imagine that one of the partners suddenly started to refer to this same piece as “the marble disc” for no contextually justifiable reason. Experiments show that the violation of the other’s expectation of pact adherence would result in significantly longer reaction times on the part of the listener (Barr & Keysar, 2002; Brennan & H. H. Clark, 1996; Krönmüller & Barr, 2007), but of course more is at stake. Breaking referential pacts erratically is fundamentally uncooperative and in our flat-packed furniture scenario, certainly grounds for an argument!

The wider phenomenon of entrainment, that of two distinct but proximal entities converging in some or other physical process, is famously ubiquitous in the natural world, and humans, like Huygen’s original synchronising pendula, really do seem to obey a similar principle of tacit co-ordination during communication (Huygen, 1893). Observe any two people in conversation and there is a strong possibility that they will converge on all sorts of linguistic and paralinguistic levels, for example in their choice of sounds (Fowler, Brown, Sabadini & Weihsing, 2003; Goldinger, 1998), referring expressions, (Brennan & H.H. Clark, 1996; Garrod & Anderson, 1987), syntactic structures (Pickering & Ferreira, 2008; Levelt & Kelter, 1982), and they will even mirror each other’s gestures (Richardson & Dale, 2005). Indeed, explicit and deliberate ‘non-convergence’ is often strikingly dissonant and stylistically powerful, as in the example from Danet (1980) in which opposing sides in a legal trial *refuse* to entrain upon a term for an aborted child during dialogue, one sticking with ‘foetus’ and the other with ‘baby’. However, away from the artifice of contrived court-room speech (and thus from the constructed and monologic), in our everyday exchanges, we are naturally highly co-ordinated language users and we work together to develop shared names for referents within conversations that we entrain upon and adhere to (H.H. Clark & Wilkes-Gibbs, 1986; Schober & H.H. Clark 1989).

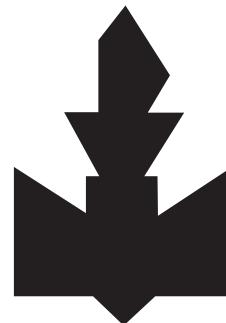
Only recently however, have researchers turned to questions of the ontogeny of entrainment phenomena. It is well established that children can use words to take a perspective on the world from very early on (E.V. Clark, 1997), but more recently, studies have shown that once a given expression has been used to take such a perspective, children rather inflexibly expect this term to be adhered to (Graham, Sedivy, & Khu 2013; Matthews, Tomasello, & Lieven, 2010). In what follows, we will argue that exploring children’s adherence to pacts is useful for addressing theoretical questions in two domains. First, children’s adherence to pacts,

and their response when others break them, can tell us a lot about their understanding of the conventional nature of language. By studying the cognitive underpinnings of children's understanding of pacts, we can establish how, with still developing social and cognitive skills, such complex communicative behaviour can be achieved so young. Second, by tracing the roots of pact formation, we are in a good position to inform debates about the adult psycholinguistic architecture (see Barr & Keysar, 2002; E.V. Clark, 1990; Pickering & Garrod, 2004). Thus, we will argue that exploring pre-schoolers' sensitivity to pacts has the potential to unravel their cognitive bases for children and adults alike. The chapter will first detail the main evidence of referential pacts in adult dialogue, and then outline the major theories that have been posited to explain such phenomena. These can be loosely divided into "rich social knowledge" and "lean memory" based interpretations of the underlying mechanisms which in their extreme forms represent opposing ends of a continuum of egocentricity in dialogue processing, differing as to the extent to which they invoke complex modelling of the interlocutor (see Brennan, Galati, & Kuhlen, 2010 for a good overview). Next, we explore the developmental context by looking at the referential behaviour of children in a general sense before describing two key experimental works with children and referential pacts, and discussing what these studies can add to the wider debate.

### Referential pacts in adults

Consider the following from Stellman and Brennan, taken from a 1993 experiment:

- "A bat"
- "The candle"
- "The anchor"
- "The rocket ship"
- "The Olympic torch"
- "The Canada symbol"
- "The symmetrical one"
- "Shapes on top of shapes"
- "The one with all the shapes"
- "The bird diving straight down"
- "The airplane flying straight down"
- "The angel upside down with sleeves"
- "The man jumping in the air with bell bottoms on"



**Figure 1.** Object descriptions from Stellman & Brennan (1993).

The diverse list of expressions to the left of Figure 1 represents the way that participants chose to refer to the depicted novel shape. Indeed, most objects in the physical world can feasibly be described in heterogeneous, and often idiosyncratic

ways. Even in more mundane dialogues, the wide variety of lexical options dictates that we must *co-ordinate* our attention on things that can be conceptualised from multiple perspectives. For example, any one of us could be described in a multitude of ways dependent upon what is most pertinent to the speaker and to the context, so that we may be variously, a woman, a student, a flautist, a driver or a vegetarian (see E.V. Clark, 1997; Geeraerts, Grondelaers, & Bakema, 1994). The conceptual perspectives that a speaker judges most contextually appropriate must then be accepted and ratified by the conversational partner for a referential pact to be formed – this implicitly stating that the perspective is to be the one maintained throughout the dialogue in the absence of contextual shifts. Once advanced, acceptance is not guaranteed (H.H. Clark & S.E. Brennan, 1991), and very often partners refuse to entrain until a term more mutually acceptable to the dyad is proposed like in the following example from H.H. Clark and Wilkes-Gibbs (1986), in which B rejects A's initial conceptualization for the novel shape and offers another which B then accepts:

- A. Um, third one is the guy reading with, holding his book to the left.
- B. Okay, kind of standing up?
- C. Yeah.
- D. Okay.

(H. H. Clark & Wilkes-Gibbs, 1986: 11)

Once they *are* formed however, adherence to referential pacts demonstrably facilitates referent identification in conversation (Barr & Keysar, 2002; Brown-Schmidt, 2009; Metzing & Brennan, 2003). Indeed, our dialogues so depend upon them that in a classic experiment, Brennan and H.H. Clark (1996) showed that interlocutors are willing to disregard referential parsimony in order to maintain a pact with their partner. They showed that participants formed a pact to suit a certain context (e.g., settling on 'the loafer' to describe a certain shoe in an array featuring different types of shoes), adhered to it consistently throughout the dialogue, and even maintained the pact if the array changed so that 'the loafer' style shoe was the only 'shoe-like' item left thereby being over-informative (as now, 'the shoe' would suffice). As Van der Wege (2009) points out, "maintaining and contrasting a history of referential conceptualizations is more important than being as concise as the reference set allows" (Van der Wege, 2009: 461). Indeed, pact adherence does seem to outweigh strict rules of informativeness in certain circumstances (for example, see Bortfield & Brennan (1997) in which English speakers maintained idiosyncratic and highly unconventional pact expressions in dialogues with non-native speakers). In spite of this perseverance, if a pact is strongly 'grounded' between two speakers quite often it will become compressed over time and will thereby expedite processing. For example, confronted once more with a novel shape, participants in the H.H. Clark and Wilkes-Gibbs (1986) study went from

the one that “looks like a person who’s ice skating” to “the ice skater” (H.H. Clark & Wilkes-Gibbs, 1986:12; but also see Krauss & Weinheimer, 1966; Schober, 1993), although presumably the conceptualisation remains the same and only the expression is abridged.

As much as *pact formation* facilitates our dialogue, *pact defeasibility* is just as important. For example, when a new speaker enters a dialogue, they are likely to bring with them their own unique conceptualisation, and so a pact established with a previous speaker would have to be abandoned to take this into account. Metzing and Brennan (2003) devised an experiment in which participants moved around novel objects in response to established pacts with confederate speakers. A pact was then broken (i.e., switching from “shiny cylinder” to “silver pipe”), either by the original confederate or by a new speaker who shared no referential history with the participant. Participants were significantly slower to react when their original partners broke a pact than when a new speaker brought with them a new term, and hence a fresh perspective. They concluded that the representations of referent expressions during dialogue must encode information about the particular interlocutor that we are interacting with and that this information enables an established pact to be abandoned (see also, Barr & Keysar, 2002; Brennan & Hanna, 2009; Garrod & Anderson, 1987; Krönmüller & Barr, 2007).

In sum, as adults we exhibit both flexible (partner specific defeasibility) and inflexible (over-informative) referential behaviour in dialogue depending upon the current situational demands. Several theoretical positions have emerged over recent years that attempt to explain these observations and we go on to outline the main contention in the next section by discussing them as either positing ‘rich social knowledge-based’ or ‘lean memory-based’ accounts of the behaviour.

### Theoretical accounts of referential pacts phenomena

Referential pacts are the example par excellence of Eve Clark’s principle of Contrast, which states that “speakers assume that any difference in form signals a difference in meaning” (E.V. Clark, 2009: 133). For example, imagine that the Stellman and Brennan (1993) novel shape featured above (A) is put into an array with an equally novel shape (B) and two interlocutors have previously established that (A) is “the ice-skater”, but have no such pact for item (B). Two interlocutors operating under the principle of Contrast and wishing to refer to item (B) would successfully avoid referential chaos, and communicative breakdown, by adhering to the following process:

1. This shape (A) is called “the ice-skater”.
2. If my partner means to refer to “the ice-skater”, she will ask for “the ice-skater”.

3. She did not, she asked for “the bird diving straight down”.
4. She must mean an alternative referent within this array (in this case, shape (B)).

Note, the principle of Contrast neatly accommodates any partner specificity effects (as found by Metzing & Brennan, 2003), because hitherto established pacts are legitimately defeasible at stage (2) if the conversational partner is new and is therefore bringing with them a new perspective on surrounding referents. Almost all theoretical accounts recognise that we are capable of reasoning in this way but where contemporary accounts tend to differ is on whether or not we *always* have to process pacts in this way and whether there may be speedier mechanisms that we can use in the rapid back and forth of conversation. This debate has arisen because it has been argued to be difficult to come up with a mechanistic account of processes that involve a full invocation of common ground – a communicative scenario in which interlocutors model each other’s minds, thoughts, beliefs and knowledge states (see Schelling, 1980; Schiffer, 1972) so that the intent involved in these exchanges amounts to a sort of recursive ‘I know that you know that I know that you know’ mechanism, possibly ad infinitum. Indeed, as Levinson remarks, the “threat of infinite regress has not endeared the idea of mutual knowledge to those interested in plausible models of psychological processes” (Levinson, 1995: 229). Those theorists concerned about the cognitive cost of common ground have proposed accounts based on heuristics that will handle pacts initially without reference to common ground (although full processing of the type proposed above is almost always seen as a recovery mechanism) and they form a group of accounts that we term lean memory based. For example, Shintel and Keysar (2007, 2009) have attempted to account for partner specificity effects within such an account by using ‘overhearing’ paradigms. Here, participants are allowed to observe a communicative exchange between an experimenter and another person in which pacts are formed. Critically, participants are hidden from the interlocutors while observing them and, subsequently, are not allowed to communicate to the experimenter that they bore witness to the pact formations. Despite knowing that they do not have true *shared* knowledge (the kind of recursive ‘I know that you know’ situation that constitutes true common ground), with the experimenter, participants tend to use the pacts when they play the game with the experimenter themselves a short while later. Furthermore, hesitation occurs when the experimenter breaks the pact in spite of the fact that it wasn’t specifically *their* pact (i.e., negotiated, ratified and fully ‘common grounded’). They conclude that listeners “merely encode the identity of the speaker into the memory representation” (Shintel & Keysar, 2007: 367 see also Krönmüller & Barr, 2007; Pickering & Garrod, 2004).

This lean view of pacts has met with other work suggesting attention to common ground even at the very earliest stages of processing, forming an opposing

group of rich social knowledge accounts based on making complex inferences about the intentions of conversational partners. For example, Brown-Schmidt (2009) showed partner-specificity effects as detectable early on during on-line comprehension of pacts in a referential communication task, but showed that they were absent in a 'non-interactive' condition in which, instead of taking the form of a conversation with an experimenter, participants merely heard recordings of the object descriptions. Indeed, findings that entrainment behaviour seems to differ *qualitatively* according to who or what a speaker perceives their audience to be lends further support to the idea that dialogue is at least partly based on socio-pragmatic inferences. Van der Wege (2009) suggests that just as Clark's principle of Contrast guides word learning by contrasting form and meaning within a lexical community, it may also guide interlocutor conceptualizations of referents during dialogue but may vary according to the salience of the audience at hand. Using a picture matching paradigm with audience design conditions that gradually decreased proximity to a real addressee, she showed that speakers addressing a real interlocutor relied much more heavily on lexical entrainment and differentiation than speakers asked to record instructions for an imagined interlocutor, or speakers simply naming pictures with no addressee with which to collaborate (Van der Wege, 2009; see also, Lockridge & Brennan, 2002; Brown & Dell, 1987; Galati & Brennan, 2010 for evidence of audience design in interaction).

The main conclusion to be drawn from this debate is that, as adults, we are certainly adept at using common ground in conversation but there appear to be multiple ways in which we can do so, some of which have a more heuristic nature than others. When faced with picking apart how such a complex system would work, it can be helpful to turn to developmental studies. By investigating conversational skills when language is still developing, we can reveal telltale errors or differences that indicate how conversational processes – those which are generally fast and seamless in adulthood – are being put together during childhood. As the richer accounts of entrainment involve complex inferential reasoning, they would generally predict that children would behave differently than adults in entraining upon referential pacts and would presumably have difficulty in knowing when to justifiably renege upon a pact. However, before we look at the experimental evidence with preschoolers, we present a brief summary of their general referential behaviour so as to contextualise the findings.

### Children's understanding of reference

There is abundant evidence that children follow the principles of Contrast and Convention from very early on, and use these to guide processes like word learning

(e.g., E.V. Clark, 1988; Diesendruck & Markson, 2001; Scofield & Behrend, 2007). There is even evidence to suggest that 3-year-olds take such principles a little too far, becoming *hyper-conventional*, and have been shown to reject alternative names for objects instead insisting that, for example, a rabbit cannot be a rabbit *and* a bunny at the same time (Doherty, 2000; Perner, Rendl & Garnham, 2007; Perner, Stummer, Sprung, & Doherty, 2002). However, that is not to say that children are entirely inflexible, as there are many accounts of infants handling multiple expressions for a single referent perfectly well (even productively, see Deák, Yen, & Pettit, 2001). This sits well with the fact that by the age of 3, children can successfully manage sophisticated perspective shifts (E.V. Clark & Svaib, 1997), accepting that a picture of a dog dressed up as a sailor could be felicitously referred to as 'dog' and 'sailor'. Also along similar lines, they appear adept at transferring the qualities of one object to another, for example during 'pretend' play, they are able to use a bowl as a hat whilst using it later that day in its more conventional role, i.e., to hold food (E.V. Clark, 1997; see also Rakoczy, Warneken, & Tomasello, 2008; Wyman, Rakoczy, & Tomasello, 2009, for further evidence of conventionality and flexibility in pretend play).

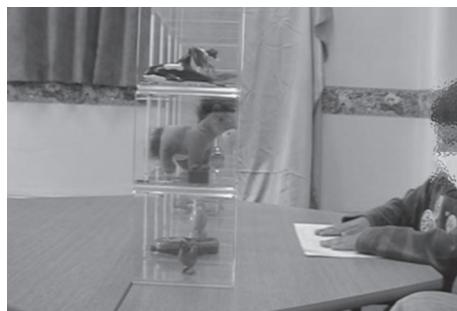
Children under the age of 3 also show evidence of taking partner-specific information into account when producing referring expressions (Matthews, Butcher, Lieven, & Tomasello, 2012), when hearing them (Fisher, Church & Chambers, 2004) and when learning them for the first time (Baldwin, 1991). Thus, they show flexibility in their judgement on the *acceptability* of sources of word learning. For example, 3- and 4-year-old children reacted differentially in a mutual exclusivity paradigm according to whether a speaker had previously marked themselves as an unreliable labeller (Sobel, Sedivy, Buchanan, & Hennessy, 2012). If the labeller had been historically unreliable, they were more likely to choose a familiar object as the referent for a novel label (a direct reversal of their behaviour with a reliable speaker; see also Nurmsoo, Robinson, & Butterfill, 2010; Sabbagh & Baldwin, 2001). There is even evidence that 17-month-old infants will take into account a speaker's false belief when learning new words (Southgate, Chevallier, & Csibra, 2010).

In sum, there is evidence that young children approach linguistic interactions equipped with the ability to make sophisticated pragmatic inferences. Yet, in the preschool years, non-adult like interactions demonstrate that these pragmatic skills are still very much in development. Therefore experiments into entrainment in infancy have a great deal to contribute to both the debate into the nature of very early communication and simultaneously, to the nature of the underlying mechanisms behind language processing.

### Children's sensitivity to referential pacts

To date, only a handful of studies have looked at children's sensitivity to referential pacts. Matthews, Tomasello and Lieven (2010) adapted the Metzing and Brennan (2003) paradigm to determine just whether children showed the same expectations as adults in terms of partner specificity. Three- and five-year-olds were invited to participate in a game which involved moving objects to different locations inside a perspex box in order to match a photograph. An experimenter held the photograph and asked the child if they could help her by putting the objects in the right place. In a warm up phase, pacts for the objects were established through scripted dialogue produced by the first experimenter (E1) (for example, using 'tree' or 'car' for certain items). In the test phase the same objects were moved about but what varied were the expressions used to refer to them and the identity of the experimenter. In the original speaker condition, E1 referred to four test items alternating between using the original expression (e.g., car) and a new expression (e.g., bush, instead of tree). The new speaker condition was exactly the same except that the child interacted with a different experimenter (E2) who had not been privy to the original pact. Children's reaction times were measured in each case from the video angle illustrated in Figure 2.

Results showed that the children displayed sensitivity to partner-specificity in using pacts and were particularly slow to respond to a pact break from an original partner as compared with a new one. However, several additional observations made the interpretation of reaction times more complex. First, unlike adults, the children slowed down when anyone used a new term, they just did it more with the original partner than with a new partner. Second, the partner-specific effect



**Figure 2.** Child participating in a study of referential pacts. Reaction times were measure from the onset of the critical referring expression to the frame in which the child's hand entered the Perspex grid.

was only evident in the first test trial. That is, if a pact had been broken once, all bets were off and processing slowed across the board. This finding has been replicated in adults (Brown-Schmidt, 2009) and in line with richer ‘social knowledge’ based accounts, could be taken as evidence that referential pacts are not just the product of a simple partner-specific memory trace. That is, if children were just quicker with an original partner as they had encoded pacts with that partner’s voice in mind, then the advantage of the original partner’s voice should hold over several trials, but it did not.

In addition to slowing down, children frequently protested at the use of a new term (saying, for example, “It’s not a pony it’s a horse!”) Given the fully counterbalanced design, it was clear that these protests were based on whatever E1 had said in the warm up phase, not on some prior preference on the part of the child. Only the 5-year-olds appeared to be aware of the source of their preference (saying, for example “*she said* it was a pony”). However, children of both ages protested with both new and old partners, which seems to suggest that children harbour a general reluctance to abandon referential pacts once they are formed, even in pragmatically motivated circumstances like the ‘new speaker’ condition.

Graham, Sedivy, and Khu (2013) extended these findings along two dimensions. First, they used adjective-noun combinations and broke pacts by using a new adjective (e.g., spotty dog, fluffy dog). This is important as it was possible that children in the Matthews et al. (2010) study were protesting that the adult had misunderstood what the object actually was (not that it had just been referred to in an unexpected way). Secondly, they added an implicit measure using eye-tracking to capture any manifestations of sensitivity not captured by reaction times. Four-year-olds were asked to look at and point to objects that had two or more salient attributes. On test trials either the original experimenter or a new one used either the original expression (e.g., the *fluffy* dog) or a new one (e.g., the *spotted* dog). Children were less likely to point to the correct object if a new expression was used, regardless of speaker. Given this was a simple task for 4-year-olds, this highlights how strong the effect of a pact can be (children were essentially willing to override normal semantic processing in favour of the pact).

Evidence for partner-specific effects came from the eye-tracking measures, specifically the target advantage score.<sup>1</sup> While no advantage was found before or whilst the adjective was unfolding (e.g., *fluffy*), for the period whilst the noun unfolded (e.g., *dog*), there was a significant expression by partner interaction. This was driven by the fact that children showed a particularly strong, positive target advantage score when the original partner used the original term. When the original partner used a new term, the target advantage score became negative. In

---

1. The probability of looking to target minus the probability of looking to the competitor object.

contrast, in the new partner conditions, there was no effect of expression on the scores, which were low but positive in both cases.

It is clear then that children's overt responses (in this case pointing) did not match the more implicit measure (in this case eye-tracking). Taken together, these findings suggest children place a lot of value on the maintenance of pacts and that they attend to partner-specific information when processing referring expressions but that, at least at 4 years of age and under, they find it difficult to square these constraints when they come into conflict. Thus, while children are less surprised by a new term if it is produced by a new speaker, they are still reluctant to adopt it. This is in line with the assumption that multiple cognitive processes underwrite referential pact phenomena. Future studies need to pin down exactly how children manage to bring the principles of Convention and Contrast into balance in such complex settings where conventions on different levels (speakers and communities, for example) come into conflict.

### Future directions

Given that children show sensitivity to pacts on some measures but not on others, a logical next step would be to explore the robustness of children's understanding and check for further evidence of defeasibility. When we piloted introducing an element to an array that would justify breaking a pact, we saw no evidence of defeasibility. In this case, a wooden doll (which was the only doll in an array) was first referred to as the 'little girl'. Another doll, this time plastic, was then added to the array and the original doll was requested by referring to it as 'the wooden doll'. Children picked up the *plastic* doll in response to this instruction despite having previously understood the word 'wooden' on a standardised vocabulary test. This suggests that one of the heuristics that children use is that people will be consistent in their use of referring expressions no matter what (and indeed adults shape the language they produce for children right in line with this expectation; see Callanan & Sabbagh, 2004). We expect that children will develop multiple heuristics, some more quick and dirty (i.e., expect the same object to be referred to in the same way), and some more complex and addressee focused, and it would be expected that they trade these off online depending on the specific constraints of the situation. Further exploration of the limits on children's pact comprehension is bound to help establish the cognitive processes that underlie these heuristics and inferences, and this will almost certainly rely on using more implicit measures (Nilsen, Graham, Smith, & Chambers, 2008).

On the other hand though, it would be interesting to study how children create and maintain pacts in their own *productions*. Preliminary findings in a study by

Köyman, Schmerse, Lieven, & Tomasello (submitted), suggest that by 6 years of age, children form pacts with their peers and maintain over-informativeness in a contextual shift similar to the Brennan and H. H. Clark (1996) findings with adults. Fremery and Grassmann (in prep) have also recently studied production and find that 4-year-olds are much more likely to adopt a term that is non-canonical for them if it is a synonym (boat>ship) or if it denotes a material (house>lego) than if it is a super-ordinate term (apple-fruit). However, in this study there was little evidence of partner specificity. Thus, although we are beginning to have a far fuller picture of children's pragmatic expectations and abilities in this domain, the picture is complex, with signs of partner-specific adaptations being picked up mostly in more implicit measures of comprehension.

To sum up, while we still have much to learn about how children handle referring expressions in complex physical and social contexts under time pressure, it is clear that children as young as 3 years of age actively strive to meet, and even to maintain, the principles of Convention and Contrast. It appears that trying to do so sometimes leads the young language learners to be inflexible (to reject alternative terms even when they are justified) and to expect conventions to apply across the board, whereas at other times it results in quite remarkable speaker-specific adaptations. The developmental challenge is to get all these expectations into balance and up and running for real-time conversation. Future research should go beyond the lean/rich debate to explain how children achieve what they do given the cognitive resources and experience they have. Doing this clearly requires postulating multiple mechanisms that most often converge to the same solution in adults, but can give rise to apparently self-contradictory behaviour early in development. Fully accounting for this developmental process will require a new focus on the principles of Convention and Contrast and perhaps lead us to explore how children arrive at them in the first place.

## References

- Baldwin, D.A. (1991). Infants' contribution to the achievement of joint reference. *Child Development*, 62(5), 875–890. DOI: 10.2307/1131140
- Barr, D.J., & Keysar, B. (2002). Anchoring comprehension in linguistic precedents. *Journal of Memory and Language*, 46(2), 391–418. DOI: 10.1006/jmla.2001.2815
- Bortfeld, H., & Brennan, S.E. (1997). Use and acquisition of idiomatic expressions in referring by native and non-native speakers. *Discourse Processes*, 23(2), 119–147. DOI: 10.1080/01638537709544986
- Brennan, S.E., & Clark, H.H. (1996). Conceptual pacts and lexical choice in conversation. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 22(6), 1482–1493. DOI: 10.1037/0278-7393.22.6.1482

- Brennan, S.E., Galati, A., & Kuhlen, A.K. (2010). Two minds, one dialog: Coordinating speaking and understanding. *Psychology of Learning and Motivation*, 53, 301–344. DOI: 10.1016/S0079-7421(10)53008-1
- Brennan, S.E., & Hanna, J.E. (2009). Partner-specific adaptation in dialog. *Topics in Cognitive Science*, 1(2), 274–291. DOI: 10.1111/j.1756-8765.2009.01019.x
- Brown, P.M., & Dell, G.S. (1987). Adapting production to comprehension: The explicit mention of instruments. *Cognitive Psychology*, 19(4), 441–472. DOI: 10.1016/0010-0285(87)90015-6
- Brown-Schmidt, S. (2009). Partner-specific interpretation of maintained referential precedents during interactive dialog. *Journal of Memory and Language*, 61(2), 171–190. DOI: 10.1016/j.jml.2009.04.003
- 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(5), 746–763. DOI: 10.1037/0012-1649.40.5.746
- 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. (1988). On the logic of contrast. *Journal of Child Language*, 15(2), 317–335. DOI: 10.1017/S0305000900012393
- Clark, E.V. (1990). On the pragmatics of contrast. *Journal of Child Language*, 17(2), 417–431. DOI: 10.1017/S0305000900013842
- Clark, E.V. (1993). *The lexicon in acquisition*. Cambridge: CUP. DOI: 10.1017/CBO9780511554377
- Clark, E.V. (1997). Conceptual perspective and lexical choice in acquisition. *Cognition*, 64(1), 1–37. DOI: 10.1016/S0010-0277(97)00010-3
- Clark, E.V. (2009). *First language acquisition*. Cambridge: CUP. DOI: 10.1017/CBO9780511806698
- Clark, E.V. (2007). Conventionality and contrast in language and language acquisition. *New Directions for Child and Adolescent Development*, 115, 11–23. DOI: 10.1002/cd.179
- Clark, E.V., & Svaib, T.A. (1997). Speaker perspective and reference in young children. *First Language*, 17(51), 57–74.
- Clark, H.H., & Wilkes-Gibbs, D. (1986). Referring as a collaborative process. *Cognition*, 22(1), 1–39. DOI: 10.1016/0010-0277(86)90010-7
- Clark, H.H., and Brennan, S.A. (1991). Grounding in communication. In L.B. Resnick, J.M. Levine, & S.D. Teasley (Eds.). *Perspectives on socially shared cognition* (pp. 127–149). Washington: APA Books
- Danet, B. (1980). 'Baby' or 'fetus'? Language and the construction of reality in a manslaughter trial. *Semiotica*, 32(3–4), 187–220. DOI: 10.1515/semi.1980.32.3-4.187
- Deák, G.O., Yen, L., & Pettit, J. (2001). By any other name: When will preschoolers produce several labels for a referent? *Journal of Child Language*, 28(3), 787–804. DOI: 10.1017/S0305000901004858
- Diesendruck, G. (2005). The principles of conventionality and contrast in word learning: An empirical examination. *Developmental Psychology*, 41(3), 451–463. DOI: 10.1037/0012-1649.41.3.451
- Diesendruck, G., & Markson, L. (2001). Children's avoidance of lexical overlap: A pragmatic account. *Developmental Psychology*, 37(5), 630–641. DOI: 10.1037/0012-1649.37.5.630
- Doherty, M.J. (2000). Children's understanding of homonymy: Metalinguistic awareness and false belief. *Journal of Child Language*, 27(2), 367–392. DOI: 10.1017/S0305000900004153
- Fisher, C., Church, B.A., & Chambers, K.E. (2004). Learning to identify spoken words. In D.G. Hall & S. Waxman (Eds.), *Weaving a lexicon* (pp. 3–40). Cambridge, MA: The MIT Press.

- Fowler, C.A., Brown, J.M., Sabadini, L., & Weihing, J. (2003). Rapid access to speech gestures in perception: Evidence from choice and simple response time tasks. *Journal of Memory and Language*, 49(3), 396–413. DOI: 10.1016/S0749-596X(03)00072-X
- Fremery, J., & Grassman, S. (In prep). Do 4-year-olds choose referential expressions according to speaker-specific referential pacts?
- Galati, A., & Brennan, S.E. (2010). Attenuating information in spoken communication: For the speaker, or for the addressee? *Journal of Memory and Language*, 62(1), 35–51. DOI: 10.1016/j.jml.2009.09.002
- Garrod, S., & Anderson, A. (1987). Saying what you mean in dialogue: A study in conceptual and semantic co-ordination. *Cognition*, 27(2), 181–218. DOI: 10.1016/0010-0277(87)90018-7
- Geeraerts, D., Grondelaers, S., & Bakema, P. (1994). *The structure of lexical variation: Meaning, naming, and context*, Vol. 5. Berlin: Mouton de Gruyter. DOI: 10.1515/9783110873061
- Goldingen, S.D. (1998). Echoes of echoes? An episodic theory of lexical access. *Psychological Review*, 105(2), 251–279. DOI: 10.1037/0033-295X.105.2.251
- Graham, S.A., Sedivy, J., & Khu, M. (2013). That's not what you said earlier: Preschoolers expect partners to be referentially consistent. *Journal of Child Language*, FirstView, 1–17.
- Huygens, C., & der Wetenschappen, H.M. (1893). *Oeuvres complètes de Christiaan Huygens*. Martinus Nijhoff.
- Köymen, B., Schmerse, D., Lieven, E., & Tomasello, M. (submitted). Young children create partner-specific referential pacts with peers.
- Krauss, R.M., & Weinheimer, S. (1966). Concurrent feedback, confirmation, and the encoding of referents in verbal communication. *Journal of Personality & Social Psychology*, 4(3), 343–346. DOI: 10.1037/h0023705
- Kronmüller, E., & Barr, D.J. (2007). Perspective-free pragmatics: Broken precedents and the recovery-from-preemption hypothesis. *Journal of Memory and Language*, 56(3), 436–455. DOI: 10.1016/j.jml.2006.05.002
- Levelt, W.J.M., & Kelter, S. (1982). Surface form and memory in question answering. *Cognitive Psychology*, 14(1), 78–106. DOI: 10.1016/0010-0285(82)90005-6
- Levinson, S.C. (1995). Interactional biases in human thinking. In E.N. Goody (Ed.), *Social intelligence and interaction: Expressions and implications of the social bias in human intelligence* (pp. 221–260). Cambridge: CUP. DOI: 10.1017/CBO9780511621710.014
- Lockridge, C., & Brennan, S. (2002). Addressees' needs influence speakers' early syntactic choices. *Psychonomic Bulletin & Review*, 9(3), 550–557. DOI: 10.3758/BF03196312
- Matthews, D., Butcher, J., Lieven, E., & Tomasello, M. (2012). Two- and four-year-olds learn to adapt referring expressions to context: Effects of distractors and feedback on referential communication. *Topics in Cognitive Science*, 4(2), 184–210. DOI: 10.1111/j.1756-8765.2012.01181.x
- Matthews, D., Lieven, E., & Tomasello, M. (2010). What's in a manner of speaking? Children's sensitivity to partner-specific referential precedents. *Developmental Psychology*, 46(4), 749–760. DOI: 10.1037/a0019657
- Metzing, C., & Brennan, S.E. (2003). When conceptual pacts are broken: Partner-specific effects on the comprehension of referring expressions. *Journal of Memory and Language*, 49(2), 201–213. DOI: 10.1016/S0749-596X(03)00028-7
- Nilsen, E.S., Graham, S.A., Smith, S., & Chambers, C.G. (2008). Preschoolers' sensitivity to referential ambiguity: Evidence for a dissociation between implicit understanding and explicit behavior. *Developmental Science*, 11(4), 556–562. DOI: 10.1111/j.1467-7687.2008.00701.x
- Nurmsoo, E., Robinson, E., & Butterfill, S. (2010). Children's selective learning from others. *Review of Philosophy and Psychology*, 1(4), 551–561. DOI: 10.1007/s13164-010-0043-y

- Perner, J., Rendl, B., & Garnham, A. (2007). Objects of desire, thought, and reality: Problems of anchoring discourse referents in development. *Mind & Language*, 22(5), 475–513. DOI: 10.1111/j.1468-0017.2007.00317.x
- Perner, J., Stummer, S., Sprung, M., & Doherty, M. (2002). Theory of mind finds its Piagetian perspective: Why alternative naming comes with understanding belief. *Cognitive Development*, 17(3–4), 1451–1472. DOI: 10.1016/S0885-2014(02)00127-2
- Pickering, M.J., & Ferreira, V.S. (2008). Structural priming: A critical review. *Psychological Bulletin*, 134(3), 427–459. DOI: 10.1037/0033-2909.134.3.427
- Pickering, M.J., & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences*, 27(2), 169–190.
- Rakoczy, H., Warneken, F., & Tomasello, M. (2008). The sources of normativity: Young children's awareness of the normative structure of games. *Developmental Psychology*, 44(3), 875–881. DOI: 10.1037/0012-1649.44.3.875
- Richardson, D.C., & Dale, R. (2005). Looking to understand: The coupling between speakers' and listeners' eye movements and its relationship to discourse comprehension. *Cognitive Science*, 29(6), 1045–1060. DOI: 10.1207/s15516709cog0000\_29
- 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(4), 1054–1070. DOI: 10.1111/1467-8624.00334
- Schelling, T.C. (1980). *The strategy of conflict*. Cambridge, MA: Harvard University Press.
- Schiffer, S.R. (1972). *Meaning*. Oxford: Clarendon Press.
- Schober, M.F. (1993). Spatial perspective-taking in conversation. *Cognition*, 47(1), 1–24. DOI: 10.1016/0010-0277(93)90060-9
- Schober, M.F., & Clark, H.H. (1989). Understanding by addressees and overhearers. *Cognitive Psychology*, 21(2), 211–232. DOI: 10.1016/0010-0285(89)90008-X
- Scofield, J., & Behrend, D.A. (2007). Two-year-olds differentially disambiguate novel words and facts. *Journal of Child Language*, 34(4), 875–889. DOI: 10.1017/S0305000907008100
- Shintel, H., & Keysar, B. (2007). You said it before and you'll say it again: Expectations of consistency in communication. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 33(2), 357–369. DOI: 10.1037/0278-7393.33.2.357
- Shintel, H., & Keysar, B. (2009). Less is more: A minimalist account of joint action in communication. *Topics in Cognitive Science*, 1(2), 260–273. DOI: 10.1111/j.1756-8765.2009.01018.x
- Sobel, D.M., Sedivy, J., Buchanan, D.W., & Hennessy, R. (2012). Speaker reliability in preschoolers' inferences about the meanings of novel words. *Journal of Child Language*, 39(1), 90–104. DOI: 10.1017/S0305000911000018
- Southgate, V., Chevallier, C., & Csibra, G. (2010). Seventeen-month-olds appeal to false beliefs to interpret others' referential communication. *Developmental Science*, 13(6), 907–912. DOI: 10.1111/j.1467-7687.2009.00946.x
- Stellmann, P., & Brennan, S. (1993). Flexible perspective-setting in conversation. Paper presented at the 34th Annual Meeting of the Psychonomic Society, Washington, DC.
- Van Der Wege, M.M. (2009). Lexical entrainment and lexical differentiation in reference phrase choice. *Journal of Memory and Language*, 60(4), 448–463. DOI: 10.1016/j.jml.2008.12.003
- Wyman, E., Rakoczy, H., & Tomasello, M. (2009). Young children understand multiple pretend identities in their object play. *British Journal of Developmental Psychology*, 27(2), 385–404. DOI: 10.1348/026151008X322893



## “We call it as puppy”

### Pragmatic factors in bilingual language choice

Medha Tare and Susan A. Gelman

University of Maryland and University of Michigan

In this chapter we examine a pragmatic task that bilingual speakers face every day: determining which language to use in which conversational context (also known as “pragmatic differentiation”). We propose that pragmatic differentiation is a gradually emerging skill that requires multiple skills, including theory of mind reasoning. Two studies are described, examining pragmatic differentiation in bilingual Marathi-English language learners under varying communicative goals and contexts. These variations reveal that children’s pragmatic differentiation is dynamic and developing over the preschool years, rather than an all-or-nothing ability. Performance on these tasks seems to be influenced by concurrently developing social insights, such as theory of mind, but may also be affected by group interaction and parent talk about language differentiation.

**Acknowledgements:** This research was supported by dissertation research grants from the University of Michigan and the journal *Language Learning* to Medha Tare, and NICHD grant HD36043 to Susan Gelman.

We wish to start with a story, told from the perspective of the second author. In the fall quarter of 1980, I started graduate school at Stanford and enrolled in Eve Clark’s language acquisition class. This was truly a transformative experience. In a manner both gentle and brilliant, Eve discussed and probed every aspect of child language: phonology, semantics, morphology, syntax, pragmatics, sociolinguistics, metalinguistics. Her knowledge was dauntingly encyclopedic. The coursepack of assigned readings was massive. The lectures were crammed full of ideas, information, insights, scientific experiments, linguistic theory, analyses of far-ranging languages, and well-chosen observations. The assigned projects, requiring that we round up actual children and record their speech in different settings and with different tasks, were thrilling and intense. I learned how to listen – really listen – to a child’s speech. I learned why transcribing a half-hour of conversation can take

upwards of several hours. I learned to appreciate the “dance” between parent and child. I learned to appreciate the charm, humor, sophistication, and layered logic of child language. This course was just the beginning of a friendship that spans more than 30 years. So thank you, Eve, for inspiring me to study child language, and for reminding us all that the language learner is a whole child within a social context, striving for meaning.

## Introduction

As Eve Clark’s research demonstrates so well, pragmatic language skills are central to language use and emerge early in a child’s development (Clark, 2004). Even infants are highly attuned to the intentions of their conversational partners and the people around them (Clark & Amaral, 2010), and by two years of age children take into account others’ language knowledge when interpreting new information, such as object labels (Akhtar, 2005; Diesendruck, 2005; Jaswal, 2004; Sabbagh & Baldwin, 2001). The development of pragmatic understanding is thus foundational to language use, from gesture to word learning to grammatical elements (Clark & Amaral, 2010).

One pervasive phenomenon that is steeped in pragmatic understanding is that of adjusting speech as a function of addressee. All speakers face the problem of determining speaker’s intent and making choices in their language use. For example, monolingual children as young as four years old adjust their speech according to their addressee, using simpler sentences when speaking to children younger than themselves and more complex sentences when speaking to adults (Shatz & Gelman, 1973). Yet pragmatic modifications are even more dramatic in bilingual<sup>1</sup> speakers who switch *between* languages to accommodate their audience. Thus, the bilingual context provides a valuable window onto the development of pragmatic language ability and its relationship to other developing metacognitive abilities.

In this chapter we focus on *pragmatic differentiation*, or the capacity of bilingual children to use their two languages appropriately with interlocutors who speak different languages. Pragmatic differentiation is one strand of a broader capacity that multilingual children have to distinguish their languages, syntactically, lexically, and phonologically (Mehler, et al., 1988; Paradis & Genesee, 1996; Pearson, Fernandez, & Oller, 1995). Bilingual children have shown an early capacity to engage in pragmatic differentiation (De Houwer, 1990; Deuchar & Quay,

---

1. Throughout the paper we use the word “bilingual” to refer to both bilingual and multilingual children.

1999; Koppe & Meisel, 1995; Nicoladis & Genesee, 1996; Quay, 2008). For example, in one foundational study, Nicoladis and Genesee (1996) conducted a longitudinal analysis of the speech of four bilingual children from age 1;7 to 3;0. The children were recorded interacting in free play with their parents. Each of their parents was dominant in either French or English and used that language primarily with the child. Each child's choice of language was analyzed relative to his or her language proficiency in the languages; that is, pragmatic differentiation was measured by examining the child's ability to use the context-appropriate language (speaking French to the French-speaking parent; speaking English to the English-speaking parent) more frequently than his/her normal language production. The children showed this differentiation, with the first demonstration of this capacity ranging in age from 1;9 to 2;4.

Importantly, further evidence suggests that young children's appreciation of appropriate language choice may extend to unfamiliar interlocutors. In one of the few extant studies on children's pragmatic differentiation with unfamiliar others, Genesee, Boivin, and Nicoladis (1996) studied four English-French bilingual children's language accommodation to strangers, again relative to the children's language dominance. Three of the four children ( $M$  age = 2;2) made accommodations to the stranger by using relatively more of the stranger's language during the free play sessions than they would normally; however, only one child used a majority of the stranger's language during her free play session. This case study suggests that children have not merely learned specific rules for particular, familiar individuals (in effect, "I talk this way with Mama and that way with Papa"), but instead possess a more general representation of when to use which of two languages.

Altogether, this prior work provides valuable evidence of bilingual children's early sensitivity to language context. At the same time, these studies leave many questions unanswered. Although, as described above, 2-year-olds are capable of using the context-appropriate language more frequently relative to their normal language production, they do not consistently produce the "appropriate" language the majority of the time. Clearly, pragmatic differentiation undergoes significant change past age two. Yet surprisingly little work has focused on the *development* of pragmatic differentiation, namely, when, how, or why these skills develop in early childhood. Relatedly, the component skills of this pragmatic understanding and their relation to other developing capacities (including metacognitive skills) have not been examined. The ability to pragmatically differentiate language use is part of a more general set of social-pragmatic skills required for appropriate language use, and involves understanding the communicative intentions of the people in one's environment (Tomasello, 2001).

In the sections that follow, we examine the developmental course of pragmatic differentiation in early childhood. We propose that bilingual children's pragmatic

understanding of when to use which language entails a complex set of skills that undergo considerable change as children gain greater understanding of their social world. Some of these skills include theory of mind reasoning about others' intentions. Cognitive developmental research has amply demonstrated that theory of mind is a layered, multi-faceted ability that undergoes continual modifications throughout childhood (Wellman & Liu, 2004). Although success on the classic theory of mind task (false belief) does not emerge until roughly four or five years of age in most children, other aspects of theory of mind emerge by two years of age or even earlier. Pragmatic differentiation shows a similar trajectory in that it undergoes a protracted development, with different components coming online at different points. To examine this complex process, we conducted two studies with young children, roughly 2–5 years of age (see Tare & Gelman, 2010, 2011, for more details). Study 1 examined children's pragmatic differentiation in three distinct contexts, and found that children's performance differs dramatically as a function of task requirements. Study 2 examined parental cues regarding language differentiation, and found that parents provide a rich set of metalinguistic cues for children at a young age. We conclude with a discussion of what these data suggest regarding the process of pragmatic differentiation in young bilingual children, as well as the mechanisms underlying this development.

### **Study 1: Pragmatic differentiation in three contexts**

If pragmatic differentiation is a single insight that emerges full-blown in toddlerhood, then children should display appropriate language choice across tasks and settings. However, if pragmatic differentiation is a gradually emerging capacity, reflecting children's budding social cognitive understanding, then performance should differ across contexts, as a function of the demands and/or scaffolding provided. We therefore designed a study in which children were asked to engage in pragmatic differentiation in three different contexts. The children were bilingual speakers of English and Marathi, an Indo-Aryan language primarily spoken in the Maharashtra state of India by nearly 96 million people (Wali, 2005). They lived in the U.S. in predominantly English-speaking communities. All children lived with two parents who each were also bilingual in English and Marathi. Parents in our study had moved to the U.S. from India on average seven years prior to the study, and reported using Marathi with their children more than 70% of the time. In order to examine developmental change, we compared participants from two different age groups (younger group: 2;7 to 3;10,  $N = 14$ , 9 girls; and older group: 4;1 to 4;11,  $N = 14$ , 4 girls). The average ratio of reported English:Marathi vocabulary knowledge, as measured by the MacArthur Communicative Development

Inventory, for the younger age group was 1.19:1 words. The average ratio of reported English:Marathi vocabulary knowledge for the older age group was 1.37:1 words. We also included metacognitive measures (e.g., theory of mind), to assess how developing cognitive capacities relate to pragmatic ability.

The three contexts, described below, were Free Play, Object Naming, and Bystander. The Free Play task was most similar to that used in prior case studies examining pragmatic differentiation (e.g., Nicoladis & Genesee, 1996). In this task, each of two researchers played with the child, speaking a single language. One researcher spoke English; one spoke Marathi (in counterbalanced order). The Free Play task provided a naturalistic, conversational context in which there was a single interlocutor, thus minimizing processing demands. Furthermore, the child received direct and repeated information regarding the interlocutor's language choice, received implicit feedback regarding his/her language choice (i.e., the researcher responded only to the "correct" language), and had complete discretion as to the topic of conversation and how to express it. Thus, we expected optimal performance in this setting.

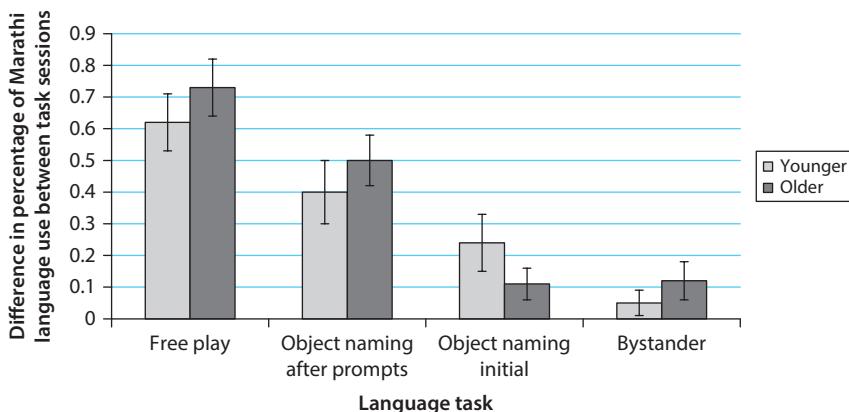
In the Object Naming task, an experimenter asked the child to name a set of familiar objects (for which the child knew both translation equivalents, as determined by a parental checklist). The child completed this task twice, once with the researcher who spoke only English and once with the researcher who spoke only Marathi (in counterbalanced order). The experimenter started by asking first, "What is this?" If the child provided a label in the wrong language (i.e., English in response to Marathi; Marathi in response to English), he/she received minimal conversational feedback in the form of three scripted prompts. As Clark (2004) writes, "if children can follow the speaker's intention, they should be able to deal with repairs to that intention" (Clark, 2004: 572). Thus, these prompts required pragmatic understanding of the speech act and indirectly requested a language switch. After a response in the wrong language, the experimenter responded with the *first prompt* indicating that she didn't understand: "What?" If the child again responded with the wrong language, the experimenter responded with the *second prompt*: "I don't know that word." If the child again used the wrong language, the experimenter responded with the *third prompt*: "Can you say it another way?" This context was predicted to provide a more demanding test than Free Play, due to the structured nature of the prompts and the requirement to produce labels in a particular language.

In the Bystander task, the child sat with one of his/her bilingual parents and looked at a book with colorful pictures of familiar objects, first alone and then in two sessions where we varied the third person (English- or Marathi-speaking researcher, in counterbalanced order) – a bystander who was present but did not converse with the parent or child. Pragmatic differentiation on this task required

sensitivity to the linguistic needs of the bystander, and thus was predicted to be the most difficult task of all.

We predicted that children would show the greatest language differentiation (i.e., using more of each speaker's language) in the Free Play task because the task had few constraints as to the content of the conversation required and provided conversational feedback in the appropriate language. In other words, the processing demands are minimal in the Free Play task, and children can direct the topic and flow of conversation. In contrast, we predicted that children would have some difficulty with Object Naming due to the cognitive demands of recognizing the pragmatic requirements of the situation and producing a label in a specific language with minimal feedback. This was predicted to increase the cognitive load and thus decrease children's ability to switch languages appropriately. Finally, for the Bystander task, we predicted that parents would accommodate their language use to the third party (using relatively more of the bystander's language), but that children would not make this accommodation because they were not speaking directly with the experimenter and would have difficulty recognizing the pragmatic requirement. Although parents were expected to accommodate their language by modulating the relative frequency of Marathi and English, they were also expected to continue using both languages, which we predicted would be too subtle a cue to shift children's language use.

Figure 1 provides an overview of how children in the two age groups performed across the three tasks (see Tare & Gelman, 2010, 2011, for details on the statistical analyses). For the Free Play and Bystander tasks, the figure shows the difference between the percentages of children's Marathi utterances in the Marathi



**Figure 1.** Difference in percentage of Marathi language use when speaking with the English versus Marathi monolingual interlocutors.

session and Marathi utterances in the English session, thus representing the degree of language switch between the two sessions. For the Object Naming task, the figure shows the difference between the percentages of trials where children gave a Marathi label in the Marathi session and in the English session both initially and after the researcher provided the prompts to switch.

As predicted, and consistent with prior research (Genesee, Boivin, & Nicoladis, 1996; Nicoladis & Genesee, 1996), children performed very well in the Free Play task; we found an interaction between the session (English or Marathi) and the language produced such that more English utterances were produced in the English session and more Marathi utterances were produced in the Marathi session ( $F(1,24) = 54.54, p < .01$ ). Furthermore, children in both age groups switched languages between the two sessions, which occurred just a few minutes apart, showing that children had a strong command of their conversational language use. We characterized individual children's response patterns based on whether they differentiated their languages, producing more English utterances (than Marathi) in the English session *as well as* more Marathi utterances (than English) in the Marathi session. This measure, a strong test as it requires children to produce a majority of correct-language utterances, revealed that 22 children out of 28 (79%) used their languages differentially. None of the children produced the reverse pattern.

In contrast, children had more difficulty with Object Naming, and used predominantly English labels with both speakers. After the initial question, only 4 children out of 28 (14%) differentiated the languages (producing more correct-language utterances than incorrect-language utterances), and even after the full series of prompts, only 43% of children did so. We did find developmental effects on the Object Naming task, with older children being more responsive to prompts to switch to Marathi after initially responding in English ( $t(26) = 2.55, p < .05$ ). Finally, as predicted for the Bystander task, parents were sensitive to the experimenter's presence and adjusted their language use accordingly ( $F(2, 51) = 7.15, p < .01$ ). Children, however, rarely adjusted their language use between sessions and there were no developmental effects. Only 7 children out of 28 (25%) were found to use the differentiating pattern.

These results demonstrate that pragmatic differentiation is not an all-or-none ability, but one with component skills that develop over the preschool years. As predicted, children display an early ability to select the appropriate language on the Free Play task, when processing demands are minimal and children can direct the topic and flow of conversation. In contrast, they showed only partial capacity for pragmatic differentiation on the Object Naming task. Children labeled appropriately in the English session of Object Naming but had more difficulty using Marathi labels, despite the fact that they knew the specific Marathi words that were elicited and were given repeated prompts. In contrast to performance on the

Free Play task where generating any talk in Marathi led to success, this task resulted in children's failure to produce the correct Marathi labels. Bilingual children may have difficulty with lexical access and production of specific words (Oller, Pearson, & Cobo-Lewis, 2007; Yan & Nicoladis, 2009), which in this case were lexical items in Marathi, generally not the children's dominant language. However, lexical retrieval in and of itself cannot account for children's difficulty on the Object Naming task, as children could have used Marathi to say "I forget" or "I only know how to say it in English". Instead, children often simply repeated the English label without demonstrating awareness of the discordant language choice. Finally, the Bystander task led to essentially no pragmatic differentiation whatsoever at this age, revealing a lack of sensitivity to someone who is listening to the conversation but not contributing to it. Note that the poor performance on the Bystander task was not due to lack of information regarding the researchers' languages. Each researcher introduced herself at the start of the session, speaking exclusively in one language (English or Marathi, depending on the researcher), and parents did appropriately modify their language use to more frequently include the bystander's language.

One important question concerns the factors that contributed to children's ability to produce the appropriate language. We propose that successful pragmatic differentiation requires metacognitive understanding, because children must realize the extent of their conversational partner's lack of language knowledge. The strongest support for this interpretation is that higher responsiveness to prompts on the Object Naming task correlated with children's scores on the theory of mind scale, even controlling for age (in months;  $r = .60$ ,  $N = 28$ ,  $p < .01$ ). Theory of mind did not correlate with performance on the Free Play or Bystander tasks, though this may be related to the fact these tasks showed less variability, with most children performing well on Free Play and poorly on Bystander. Although it is unclear from the present data what the exact relationship is between pragmatic differentiation and social cognition, we propose that increased metacognitive understanding contributes to children's ability to use their languages appropriately to communicate successfully. In particular, increased theory of mind might allow children to consider what it means for the Marathi experimenter's language ability when she says "I don't know that word" in response to an English label; this skill would be a key mechanism for success in this task. From the present study, it seems that theory of mind development positively influences bilingual children's pragmatic skills in the preschool years. We did not find a significant correlation between children's theory of mind score and a score of how balanced they were in their two vocabularies, suggesting that being bilingual per se did not account for their theory of mind abilities. This will be an important issue to address in future research.

### Study 2: Parental metalinguistic cues regarding language differentiation

Another source for pragmatic language learning may be parental input, specifically metalinguistic talk. Clark and Estigarribia (2011) examined the type of talk that monolingual parents use to introduce new objects to young children (aged 1;6 and 3;0), including how they “establish joint attention on an object, introduce a new label for the object, and situate that object conceptually” (1), and Clark (1978) noted that monolingual children refer to language itself around age three. We examine this type of metalinguistic talk, used to situate the labels in a bilingual context, in our Bystander task (see Tare & Gelman, 2011, for more details regarding the methodology and results).

In this study, we found that bilingual parents use a variety of metalinguistic strategies to highlight translations for object labels during everyday picture book discussion. Parents were not asked to discuss translations or provide multiple labels; nonetheless, the majority of parents used at least one of the metalinguistic strategies that we analyzed. Further, parents did not use more strategies when a researcher was present than when they were alone with the child, suggesting that metalinguistic talk occurs naturally. The primary strategies, which we expected to find, were that parents asked children to provide translations, provided translations themselves, and used the names of the two languages. However, these naturalistic data also provided the opportunity to examine parents’ other unexpected strategies, which we coded if used by more than one parent. These include parents requesting that children repeat the translations that were provided, repeating translations themselves, and even suggesting wrong translations to their children (so that children would recognize the right one when parents said it).

We coded the conversational turns to examine the ways in which parents scaffolded the pragmatics of bilingual word learning for their children. The following annotated excerpt from a conversation shows how a parent navigates eliciting labels for the pictures in the book, including using very explicit requests for translation.

- (1) Excerpt (Mother with male child, age 4;11):

MOT: *Hmm, ani hai kai?*  
 ‘Hmm, and what’s this?’

Coded as: Complete Marathi

CHI: Dog

Coded as: Complete English; Child’s First Label in English

MOT: *Ani dogla kai mhantow Marathit tu?*  
 ‘And what do you call “dog” in Marathi?’

- Coded as: Marathi with English Insertion (Quoting);  
Requests Translation; Uses Name of Language
- CHI: We call it as puppy.
- Coded as: Complete English
- MOT: *Nahi nahi, apan kai mhantow Marathit?*  
'No, no, what do we say in Marathi?'
- Coded as: Complete Marathi; Requests Translation;  
Uses Name of Language
- CHI: I don't know.
- Coded as: Complete English
- MOT: *Kutra*  
'Dog'
- Coded as: Complete Marathi; Provides Translation Equivalent
- MOT: *Kai mhantow?*  
'What is it called?'
- Coded as: Complete Marathi; Requests Repetition of Translation
- CHI: *Kutra*  
'Dog'
- Coded as: Complete Marathi; Repeats Translation

The excerpt shows the complex negotiation of label(s) for the picture of a dog, including how the parent asks for a Marathi label after the child's initial English response and how the child produces an alternate English label ("puppy") rather than a translation equivalent. The majority of parents (79%, or 22/28) used at least one metalinguistic strategy, and nearly half (43%, or 12/28) used such strategies on 10% or more of their utterances (as noted earlier, see Tare & Gelman 2011, for more details). However, there was no significant relation between parents' use of metalinguistic strategies and their children's ability to show sensitivity to the bystander's language knowledge.

The most frequently used strategy was the parent requesting a translation equivalent from the child. Also as predicted, parents often used the names of the two languages and provided translations. They also repeated translations themselves, and sometimes requested that the child repeat a translation. Children provided translations and repeated translations, but, surprisingly, did not request translations from their parents. Sixty-five percent of parents' translations were given in Marathi (as opposed to English). Eighty-four percent of children's translations were given in Marathi.

We also analyzed whether parents “marked” the translations that they provided to determine how clear the contrast would be to children that they were hearing a translation equivalent rather than a label for a new object. This was done by examining whether parents made an explicit comment regarding the translation prior to providing it (i.e., requesting a translation or using the name of a language) or if they stated it without making any comment, requiring children to realize themselves that it was a translation equivalent. The analysis showed that 60% of parents’ translations were marked (41% of translations to older children and 68% of translations to younger children). These results indicate that, similar to parents of monolingual children who explain when they provide multiple English labels for an item, such as part-terms (Callanan & Sabbagh, 2004), bilingual parents also tend to clarify the relationship between the labels in the two languages in this context. Parents also appear to adjust their metalinguistic talk to be developmentally appropriate as monolingual parents do (Clark & Estigarribia, 2011), as indicated by the greater use of marked translations with younger children.

To determine when children were likely to translate, we examined parents’ utterances preceding children’s translations. Twenty percent of children’s translations were provided spontaneously, without prompting; in contrast, 80% of children’s translations were given in response to parents’ translation requests. Of the translation requests made by parents, 67% explicitly used the name of the language they wanted the child to translate a label into. Thus, most of children’s translations were provided after an explicit prompt from their parent. Parents’ subtler prompts to translate came in different forms, including questions such as, “What else is it called?” or “What do we call it?”

Children’s initial labeling of a picture (“first label”) was more often in English ( $M = 37.64$ ,  $SD = 8.87$ ) than in Marathi ( $M = 5.43$ ,  $SD = 4.38$ ),  $t(27) = 14.43$ ,  $p < .01$ . Older children ( $M = 41.07$ ,  $SD = 7.99$ ) provided significantly more first labels in English than younger children ( $M = 34.21$ ,  $SD = 8.61$ ),  $t(26) = 2.18$ ,  $p < .05$ , and younger children provided significantly more first labels in Marathi ( $M = 7.21$ ,  $SD = 5.38$ ) than older children ( $M = 3.64$ ,  $SD = 2.02$ ),  $t(26) = -2.33$ ,  $p < .05$ . This developmental difference of older children shifting to English dominance was expected, given that the older children were more often enrolled in English-speaking preschools, compared to the younger children, who were more often still at home in a primarily Marathi-speaking household. Given the status of English as the majority language, we also predicted that the percentage of utterances after which parents used a translation strategy would be higher when children first labeled something in Marathi compared to English, as parents would have a bias toward ensuring that children knew the English label for an object. Thus, we predicted that parents would be particularly likely to use one of the metalinguistic strategies to elicit or provide the English label if the child gave the Marathi label for an object first. As predicted,

the percentage of Marathi first labels provided by children that was followed by a translation strategy was higher than the percentage of English first labels that was followed by a translation strategy ( $t(27) = -1.83, p < .05$ ; one-tailed).

These results show the nuances of parents' and children's talk about translations. Parents' explicitness when providing a label in another language suggests that they may be trying to help children organize their language knowledge. A potential consequence of this explicit talk may be that bilingual children develop language learning biases, such as mutual exclusivity, differently compared to monolingual children (Au & Glusman, 1990). Discussion about the nature of the two languages they speak may enhance preschoolers' understanding that the label for an item is not inextricably tied to the referent, an awareness which many researchers have examined in school-aged bilingual children (Bialystok, 1988; Rosenblum & Pinker, 1983).

Consistent with what other researchers have found (Kasuya, 1998; Lanza, 2001; Pan, 1995; Vedder, Kook, & Muysken, 1996), parents primarily used the minority language of Marathi (i.e., their native language) in our study, whereas children primarily used the majority language, English. The strategies often focused on translating into the minority language of Marathi, but in contrast to parental strategies examined in other studies (Kasuya, 1998; Lanza, 2001), they were not necessarily intended to shift children's language use away from the majority language of English. Parents showed some bias toward English in that they did not expect children to label pictures in Marathi and even seemed to ensure that children know the labels in English. Parents allowed a bilingual context (Lanza, 2001; Quay, 2012) and children exhibited some tendencies of "passive" bilinguals who understand both languages but mainly speak one of them (Kasuya, 1998). Thus, children's lack of pragmatic sensitivity in the Bystander and Object Naming tasks may also have resulted from their engaging in a typical conversational pattern of responding to adults' Marathi questions in English.

## Conclusions

These studies highlight the role of interaction in children's language learning and the nuanced development of pragmatic language use. As Siegal and Surian (2007) write, "a number of mechanisms underwrite the development of conversational understanding that involve an increasing sensitivity with age to conversational conventions and to linguistic and extra-linguistic contexts for the interpretation of meaning" (Siegal & Surian, 2007: 306). By varying the communicative goals and contexts in these studies, we examined bilingual children's linguistic responses to an open-ended conversational context (Free Play), to direct (though not explicit)

requests for use of a specific language (Object Naming), and to conversation with a familiar bilingual partner but with a monolingual bystander present (Bystander). These variations reveal that children's pragmatic differentiation is dynamic and developing over the preschool years, rather than an all-or-nothing ability. Performance on these tasks seems to be influenced by concurrently developing social insights, such as theory of mind, but may also be affected by group interaction and parent talk about language differences.

Using experiments to manipulate the cognitive and pragmatic demands of the situation, we teased apart children's current and developing competences. Importantly, complementary information can also be gleaned by examining transcripts in detail. In Study 2, we examined turn-by-turn exchanges to understand the nuances of how children's language choice elicited linguistically and metalinguistically informative responses from their parents. The results of this analysis, which revealed the common conversational pattern of parent questions in Marathi being answered in English by children, may help explain children's performance in the Study 1 Object Naming task. Our approach follows in the footsteps of Eve Clark, who has provided many brilliant examples of how close attention to children's language can provide insights into the developmental process (e.g., Clark, 2007; Clark & Bernicot, 2008; Clark, Gelman, & Lane, 1985; Clark & Sengul, 1978).

There are still many questions to be answered. Because bilingual populations are so heterogeneous, factors on many different levels (individual, familial, and community) may affect the development of children's pragmatic differentiation. Children in one-parent/one-language homes may experience very different input from children in homes with two bilingual parents in terms of the amount of household members' language mixing or competence in multiple languages. This could affect children's performance on language tasks; for example, if a child does not have much experience with monolingual speakers of each of their languages, he/she may be less successful on these pragmatic tasks than a child from a one-parent/one-language home. Altogether, these data suggest that rich input and experiences in multilingual settings as well as corresponding cognitive developments play an important role in supporting the development of mature bilingual responses. This area is ripe for future investigations that explore the nature of these developments in more detail.

## References

- Akhtar, N. (2005). The robustness of learning through overhearing. *Developmental Science*, 8, 199–209. DOI: 10.1111/j.1467-7687.2005.00406.x

- Au, T., & Glusman, M. (1990). The principle of mutual exclusivity in word learning: To honor or not to honor? *Child Development*, 61, 1474–1490. DOI: 10.2307/1130757
- Bialystok, E. (1988). Levels of bilingualism and levels of linguistic awareness. *Developmental Psychology*, 24, 560–567. DOI: 10.1037/0012-1649.24.4.560
- 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. DOI: 10.1037/0012-1649.40.5.746
- Clark, E.V. (1978). Awareness of language: Some evidence from what children say and do. In A. Sinclair, R. Jarvela, & W.J.M. Levelt (Eds.), *The child's conception of language* (pp. 17–43). New York, NY: Springer. DOI: 10.1007/978-3-642-67155-5\_2
- Clark, E.V. (2004). Pragmatics and language acquisition. In L.R. Horn & G. Ward (Eds.), *Handbook of pragmatics* (pp. 562–577). Oxford: Blackwell.
- Clark, E.V. (2007). Young children's uptake of new words in conversation. *Language in Society*, 36, 157–182. DOI: 10.1017/S0047404507070091
- Clark, E.V., & Amaral, P. (2010). Children build on pragmatic information in language acquisition. *Language and Linguistics Compass*, 4, 445–457. DOI: 10.1111/j.1749-818X.2010.00214.x
- Clark, E.V., & Bernicot, J. (2008). Repetition as ratification: How parents and children place information in common ground. *Journal of Child Language*, 35, 349–371. DOI: 10.1017/S0305000907008537
- Clark, E.V., & Estigarribia, B. (2011). Using speech and gesture to introduce new objects to young children. *Gestalt*, 11, 1–23. DOI: 10.1075/gest.11.1.01cla
- Clark, E.V., Gelman, S.A., & Lane, N. M. (1985). Noun compounds and category structure in young children. *Child Development*, 56, 84–94. DOI: 10.2307/1130176
- Clark, E.V., & Sengul, C.J. (1978). Strategies in the acquisition of deixis. *Journal of Child Language*, 5, 457–475. DOI: 10.1017/S0305000900002099
- De Houwer, A. (1990). *The acquisition of two languages: A case study*. Cambridge: CUP. DOI: 10.1017/CBO9780511519789
- Deuchar, M., & Quay, S. (1999). Language choice in the earliest utterances: A case study with methodological implications. *Journal of Child Language*, 26, 461–475. DOI: 10.1017/S0305000999003852
- Diesendruck, G. (2005). The principles of conventionality and contrast in word learning: An empirical examination. *Developmental Psychology*, 41, 451–463. DOI: 10.1037/0012-1649.41.3.451
- Genesee, F., Boivin, I., & Nicoladis, E. (1996). Talking with strangers: A study of bilingual children's communicative competence. *Applied Psycholinguistics*, 17, 427–442. DOI: 10.1017/S0142716400008183
- Jaswal, V.K. (2004). Don't believe everything you hear: Preschoolers' sensitivity to speaker intent in category induction. *Child Development*, 75, 1871–1885. DOI: 10.1111/j.1467-8624.2004.00822.x
- Kasuya, H. (1998). Determinants of language choice in bilingual children: The role of input. *International Journal of Bilingualism*, 2, 327–346.
- Koppe, R., & Meisel, J.M. (1995). Code-switching in bilingual first language acquisition. In L. Milroy & P. Muysken (Eds.), *One speaker, two languages: Cross-disciplinary perspectives on code-switching* (pp. 276–301). Cambridge: CUP. DOI: 10.1017/CBO9780511620867.013

- Lanza, E. (2001). Bilingual first language acquisition: A discourse perspective on language contact in parent-child interaction. In J. Cenoz & F. Genesee (Eds.), *Trends in bilingual acquisition* (pp. 201–230). Amsterdam: John Benjamins.
- Mehler, J., Jusczyk, P., Lambertz, G., Halsted, N., Bertoni, J., & Amiel-Tison, C. (1988). A precursor of language acquisition in young infants. *Cognition*, 29, 143–178. DOI: 10.1016/0010-0277(88)90035-2
- Nicoladis, E., & Genesee, F. (1996). A longitudinal study of pragmatic differentiation in young bilingual children. *Language Learning*, 46, 439–464. DOI: 10.1111/j.1467-1770.1996.tb01243.x
- Oller, D.K., Pearson, B.Z. & Cobo-Lewis, A.B. (2007). Profile effects in early bilingual language and literacy. *Applied Psycholinguistics*, 28, 191–230. DOI: 10.1017/S0142716407070117
- Pan, B.A. (1995). Code negotiation in bilingual families: “My body starts speaking English.” *Journal of Multilingual and Multicultural Development*, 16, 315–327. DOI: 10.1080/01434632.1995.9994610
- Paradis, J., & Genesee, F. (1996). Syntactic acquisition in bilingual children: Autonomous or interdependent? *Studies in Second Language Acquisition*, 18, 1–25. DOI: 10.1017/S02722263100014662
- Pearson, B.Z., Fernandez, S., & Oller, D.K. (1995). Cross-language synonyms in the lexicons of bilingual infants: one language or two? *Journal of Child Language*, 22, 345–368. DOI: 10.1017/S030500090000982X
- Quay, S. (2008). Dinner conversations with a trilingual two-year-old: Language socialization in a multilingual context. *First Language*, 28, 5–33. DOI: 10.1177/0142723707083557
- Quay, S. (2012). Discourse practices of trilingual mothers: Effects on minority home language development in Japan. *International Journal of Bilingual Education and Bilingualism*, 15, 435–453. DOI: 10.1080/13670050.2012.665828
- Rosenblum, T., & Pinker, S.A. (1983). Word magic revisited: Monolingual and bilingual children’s understanding of the word-object relationship. *Child Development*, 54, 773–780. DOI: 10.2307/1130064
- Sabbagh, M., & Baldwin, D. (2001). Learning words from knowledgeable versus ignorant speakers: Links between preschoolers’ theory of mind and semantic development. *Child Development*, 72, 1054–1070. DOI: 10.1111/1467-8624.00334
- Shatz, M., & Gelman, R. (1973). The development of communication skills: Modifications in the speech of young children as a function of listener. *Monographs of the Society for Research in Child Development*, 38, 1–37. DOI: 10.2307/1165783
- Siegel, M., & Surian, L. (2007). Conversational understanding in young children. In E. Hoff & M. Shatz (Eds.), *Blackwell handbook of language development* (pp. 304–323). Oxford: Blackwell. DOI: 10.1002/9780470757833.ch15
- Tare, M., & Gelman, S.A. (2010). Can you say it another way? Cognitive factors in bilingual children’s pragmatic language skills. *Journal of Cognition and Development*, 11, 137–158. DOI: 10.1080/15248371003699951
- Tare, M., & Gelman, S.A. (2011). Bilingual parents’ modeling of pragmatic language use in multiparty interactions. *Applied Psycholinguistics*, 32, 761–780. DOI: 10.1017/S0142716411000051
- Tomasello, M. (2001). Perceiving intentions and learning words in the second year of life. In M. Bowerman & S.C. Levinson (Eds.), *Language acquisition and conceptual development* (pp. 132–158). Cambridge: CUP. DOI: 10.1017/CBO9780511620669.007

- Vedder, P., Kook, H., & Muysken, P. (1996). Language choice and functional differentiation of languages in bilingual parent-child reading. *Applied Psycholinguistics*, 17, 461–484. DOI: 10.1017/S0142716400008201
- Wali, K. (2005). *Marathi*. Munich: Lincom.
- Wellman, H.M., & Liu, D. (2004). Scaling of theory of mind tasks. *Child Development*, 75, 523–541. DOI: 10.1111/j.1467-8624.2004.00691.x
- Yan, S., & Nicoladis, E. (2009). Finding *le mot juste*: Differences between bilingual and monolingual children's lexical access in comprehension and production. *Bilingualism: Language and Cognition*, 12, 323–335. DOI: 10.1017/S1366728909990101

# Learning words through probabilistic inferences about speakers' communicative intentions

Michael C. Frank

Department of Psychology, Stanford University

How do children learn the meanings of words? This chapter presents a probabilistic, communicative view of word learning that synthesizes insights from work on statistical learning and social learning. By describing the formal characteristics of models, it is possible to differentiate communicative models that make inferences about the speaker's intentions from associative models that treat social information as a signal of salience. In addition, the probabilistic communicative framework can be integrated with models of pragmatic reasoning, leading to insights into how Gricean principles can facilitate word learning.

**Acknowledgements:** Thanks to the editors of this volume for the opportunity to contribute, to Dave Kleinschmidt, Molly Lewis, Dan Yurovsky, Chigusa Kurumada, and an anonymous reviewer for helpful comments, and also to Noah Goodman, my collaborator in much of the work described here.

## Introduction

The linguistic world of young children is likely an overwhelming place. Even if they are not assaulted with James' (1890) "blooming buzzing confusion," it must be perplexing for infants to be surrounded constantly by sounds whose only interpretable meaning at first comes from the tone of voice in which they are uttered. Perhaps another metaphor is more apt: The infant is a traveler trying to negotiate a task in a foreign language, with the help of a sympathetic interlocutor (a parent or caregiver). The child and caregiver may share goals, or at least an understanding of the other's goals. But without knowing any words, only a few utterances can be decoded from context; and the fewer words that are known, the less leverage the

child has to infer the meanings of others. Even in the highly supportive contexts created by parents, some substantial portion of spoken language is likely to be incomprehensible to young infants.

How then do children begin to break into the vocabulary of their first language? Two broad proposals run throughout work on word learning, from historical sources to contemporary models: associative and intentional proposals. In associative accounts from Locke (1690/1964) onward, infants are hypothesized to match elements of their linguistic environment with the world around them, identifying the consistent mappings between words and other stimuli. In intentional accounts from St. Augustine (397/1963) onwards, in contrast, the attention of learners is on the speakers who produce words. These words are then mapped to the speakers' intended meanings – often instantiated by the speakers' intended referent in the current context. Although distinct evidence is often cited in support for one view or the other (e.g., Baldwin, 1993; Smith & Yu, 2008), these two views need not be in conflict. I will argue here that they can be integrated in a single framework.

I did not begin studying language acquisition with this view. When I entered graduate school, I was convinced that all of language acquisition, especially word learning, could be reduced to a process of pure statistical inference. This idea was more probabilistic than the classic generative story for syntactic acquisition, but otherwise quite similar: The child would passively observe some corpus of input. She would then carry out a defined learning procedure over this corpus, resulting in a grammar of the target language. Eve Clark's "First Language Acquisition" (2003) played a crucial role in my departure from this purely statistical viewpoint.

Reading Eve's account of early language development, illustrated with examples from the speech of actual children, made a substantial impression on me. Despite my interest in acquisition, and my presence at the same institution as Eve, I had unfortunately missed the opportunity to take her course on the topic. So it was an awakening to read how she wrote about children's learning situation – as consisting of a context, a communicative goal, and a creative drive to use language to accomplish that goal. This vision of children playing, exploring, and interacting at the same time as they learned also lined up nicely with new work suggesting this same sort of process is at work in children's causal learning (e.g., Schulz, Kushnir, & Gopnik, 2007). But most importantly, it felt true to the actual children I had met.

My goal in this chapter is to provide an overview of the broader framework that has emerged from my work on language learning in a communicative context. This framework is in some deep sense a computational instantiation of the view described in Eve's writings: It is fundamentally communicative in that it is oriented around representations of speakers' goals and intentions. But a further strength – I believe – is that it also captures a substantial amount of the core value of associative

accounts: In particular, it takes advantage of the graded, probabilistic nature of learning to aggregate information across multiple learning situations.

The probabilistic communicative framework fits well with our emerging understanding of children's cognitive development. The intentional aspects of the framework are congruent with research on children's social cognition (e.g., Csibra & Gergely, 2009; Onishi & Baillargeon, 2005; Vouloumanos, Onishi, & Pogue, 2012) and the framework supports graded probabilistic learning of the sort that is attested across development (Gopnik, 2012; Tenenbaum, Kemp, Griffiths, & Goodman, 2011). Most importantly, however, the work I describe here is grounded in the literature on pragmatic inference (H. Clark, 1996; Grice, 1975) and its role in language acquisition (Bloom, 2002; E. Clark, 2003; Tomasello, 2003). In this chapter, for reasons of space I will focus primarily on the computational literature with only occasional references to empirical issues, but I hope the debt is clear.

I use the word "framework" in this chapter rather than "model" because a framework to me refers to a set of principles, while a model refers to a particular instantiation of those principles within a working, implemented system from which concrete, quantitative predictions can be made. Although my collaborators and I have made a number of such models (which I reference here), all are specific to particular applications or kinds of data. I believe that there are some more general conclusions that can be drawn from these individual systems when they are examined together, however; these are the conclusions I focus on in this chapter.

As I describe this framework, I will also try to lay out some theoretical distinctions in more depth than is allowed in a strictly empirical report. I will first describe a taxonomy of computational models of word learning. In the next section, I will specify formal details corresponding to this taxonomy. The goal of this section will be to distinguish between those cross-situational models that merely use social information and those that make an intentional assumption about the nature of the learning situation. Finally, the last section will focus on the links between our intentional framework for word learning and our work on modeling pragmatic inference.

A note on terminology. When discussing word learning, I will be discussing how children learn open-class words like "rabbit," "white," or "run" rather than closed-class words like "the" or "no." The problem of inferring meanings for open-class words can be broken into two separate tasks: the task of words to referents in the moment (reference assignment), and the task of identifying the meaning or concept corresponding to the indicated referent – or to the aspect or property of that referent that the speaker wishes to highlight (concept learning). The distinction between these two problems is clear, but our terms for discussing them are often clumsy. The issue is often confused further by the use of simple noun learning as a case study, since solving the reference problem in that case comes close to

giving away the concept-learning problem (modulo a bias for basic-level categories; Markman, 1991). I'll be focusing initially on word-object mapping but I will note when the equivalence between mapping and concept learning is broken.

### A taxonomy of models of word learning

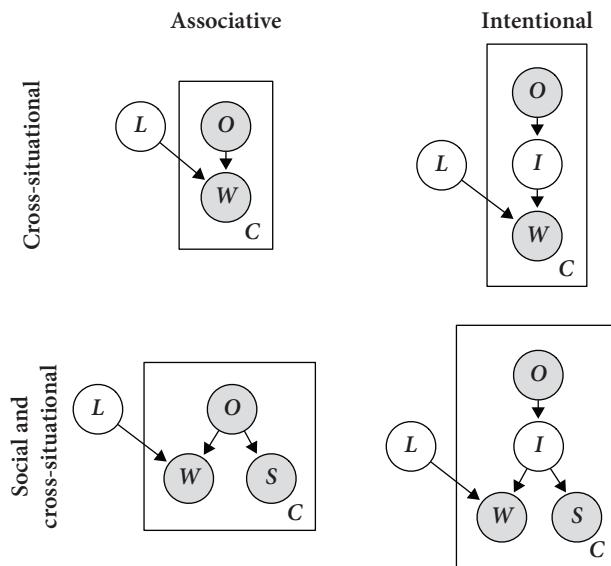
The strategy of learning words via co-occurrence has been labeled “cross-situational” word learning. Recent empirical work provides strong support for the idea that both adults and infants can learn word-object mappings by gathering consistent associations across multiple, ambiguous exposures (Smith & Yu, 2008; Vouloumanos, 2008; Vouloumanos & Werker, 2009; Yu & Ballard, 2007). A variety of theorists have asked about the utility of cross-situational learning in acquiring words of different types (Akhtar & Montague, 1999; Fisher, Hall, Rakowitz, & Gleitman, 1994; Gillette, Gleitman, Gleitman, & Lederer, 1999; Gleitman, 1990; Pinker, 1984). Though theoretically all are learnable (Siskind, 1996), there is likely to be a continuum from those words most learnable by co-occurrence, e.g., nouns and some property terms, to those least likely to be inferred from context alone. This latter category likely includes verbs – which refer to events that have multiple construals (e.g., “chase”/“flee”) – context-dependent adjectives, and function words. Because of the relative simplicity of learning basic-level object nouns from context, I begin by discussing word-object mapping models.

A taxonomy of models of word-object mapping is shown in Figure 1. All of these models are “cross-situational” in the sense that all consider evidence about the relationship between words and parts of the world across multiple observations. All of them also attempt to learn a lexicon: a set of consistent word-object mappings. They differ, however, in both the information sources they consider and the ways they use these information sources. Each of these models describes a set of variables, both observed (shaded) and unobserved (unshaded) as well as a set of causal dependencies between these variables. These dependencies define a generative process: a set of steps by which the learner assumes the observed data have been generated. Unobserved parts of this generative process can then be estimated using standard inference techniques.<sup>1</sup>

In our taxonomy, we differentiate models on two dimensions. The first – represented by the vertical axis in Figure 1 – is the information sources considered by the model. “Pure” cross-situational models consider only the co-occurrence between words and objects in establishing links between words and objects in the

---

1. Perfors, Tenenbaum, Griffiths, and Xu (2011) give a very nice tutorial introduction to this general style of modeling as applied to developmental questions.



**Figure 1.** A progression of possible models of the basic challenge faced by early word learners. L is the child's lexicon, C refers to the contexts in which utterances containing words W are observed, accompanied by object referents O and social cues S, as well as (unobserved) communicative intentions I.

lexicon. Social and cross-situational models consider also a set of social cues (envisioned here as signals like a point, a gesture, or a gaze towards an object, with some temporal connection to a particular utterance).

The second dimension is the way models represent the relationship between words and the world – shown on the horizontal axis in Figure 1. Associative models describe a generative process in which words are assumed to be generated directly by the presence of their associated objects (without the presence of an intervening speaker). In contrast, intentional models, as we define them, assume that words are generated via the intention of a speaker to refer to an object. Thus, in the sense used by theorists of social cognition, intentional models are fundamentally triadic: they define a relationship between the child, the speaker, and objects in the context (Baldwin, 1995; Carpenter, Nagell, & Tomasello, 1998).

In the current analysis, models are ideal observers: realizations of the assumptions that learners make and information sources they use, rather than the mechanisms by which these assumptions and information sources are processed (Geisler, 2003). Another way of putting this is that our analyses are at the computational, rather than algorithmic level, describing models as they represent the task faced by the child rather than as the child solves them (Marr, 1982).

There is a rich literature describing how such algorithmic constraints should be implemented in word learning models (Fazly, Alishahi, & Stevenson, 2010; Yu & Smith, 2012), and my own work in other domains has investigated these questions as well (Frank & Gibson, 2011; Frank, Goldwater, Griffiths, & Tenenbaum, 2010). In addition, there is a fascinating and growing literature investigating the ways that resource-bounded decision-making can relate to normative models (see e.g., Sanborn, Griffiths, and Navarro, 2010 for review and discussion). Yet I worry that implementational considerations often mask (rather than reveal) the learner's underlying assumptions about the learning situation, and I believe that there is value in considering the computational level independent from – and perhaps in parallel to – the algorithmic.

Related to this issue is a set of recent criticisms of the idea that cross-situational observation is a factor in learning word-object mappings (Medina, Snedeker, Trueswell, & Gleitman, 2011; Trueswell, Medina, Hafri, & Gleitman, 2013). The key empirical question underlying these criticisms is whether individual word learners represent multiple hypotheses about the meanings of words. The proposed alternative is that individuals make noisy, stochastic choices of individual hypotheses that they then test against data (a “propose but verify” strategy). The average of many such stochastic choices across items and individuals would then produce the pattern of gradual learning that is sensitive to the degree of cross-situational ambiguity that was observed in previous studies. This criticism is rooted in a long-standing debate about the nature of learning more generally, and whether it is typically gradual and associative in character or discrete, and hypothesis-based (Gallistel, 1990; Gallistel, Fairhurst, & Balsam, 2004).

Such concerns relate primarily to the mechanisms of learning applied by the learner in recovering the lexicon of their language, rather than to the underlying assumptions that guide this learning. Even if “propose but verify” learners do not retain previous data to test their hypotheses, they still must make the assumption that their hypotheses should in principle be consistent with previous data as well as future observations. In other words, even if learners are memoryless, their behavior is still consistent with an assumption of cross-situational statistical consistency. Therefore, in this chapter I will not discuss an important body of modeling work that investigates the consequences of algorithmic details of representation for modeling human performance in word learning (e.g., Li, Farkas, & MacWhinney, 2004; Regier, 2005; Yu & Smith, 2012).

In the next section, I describe computational details underlying the models in Figure 1. Providing these details allows for clarity about how an intentional assumption can be implemented, and additionally allows for comparison and integration with the models of Gricean pragmatics described below.

## A formal framework for cross-situational learning

Formalization allows us to consider what has previously been a somewhat slippery distinction: between *social cross-situational models*, which consider information generated by other people, and *intentional models*, which are based on a stronger assumption about the generating source of this information. Following the taxonomy in Shafto et al. (2012), I take the fundamental assumption underlying a communicative model to be that language is produced as a rational action to accomplish a goal.<sup>2</sup> The “language as rational action” assumption allows for stronger inferences from data than those possible under a view that considers social data but does not consider the communicative intentions that lead to those data being generated. This section walks through the formal mechanics of these ideas, describing the basics of cross-situational learning under the general family of models, laying out the key differences between associative and intentional models, and describing how social information is used in associative and intentional models.

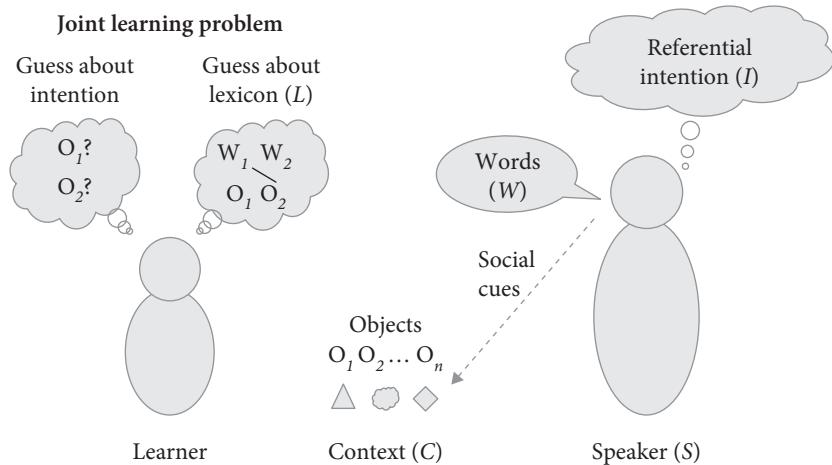
### *Basic cross-situational learning*

Consider a schematic description of the child's learning problem, shown in Figure 2. The child finds herself in a set of contexts  $C$ . Each of these contexts contains possible referents  $O_C = o_1 \dots o_n$ . In each of these contexts, a speaker has an intention  $I_C$ , unobserved to the child, that captures the idea that the speaker would like to convey (described in greater detail below – at this point, the speaker's intention is simply a placeholder). On the basis of this communicative intention, the speaker utters words  $W_C = W_1 \dots W_n$  and produces social gestures  $S_C$ . Having observed a set of these kinds of contexts, the learner's goal is to infer a set of correspondences between words and objects, which we denote  $L$  (the lexicon). This lexicon is assumed to be stable across contexts and individuals,<sup>3</sup> and can be modeled either

---

2. Of course, the idea of language as a rational action does not originate in the computational literature; for example, Grice (1975) writes that “... one of my avowed aims is to see talking as a special case or variety of purposive, indeed rational, behavior...” (p. 47). Clark (1988, 1990) then describes the linkage of this idea to the acquisition context.

3. Does the child know that her lexicon should be stable across time and identical across individuals? There are several such foundational assumptions that are necessary for all of the models described in this chapter. For example, we assume that words are linked to a particular level of description (objects or object concepts in the models we consider), and not to some others (e.g., motor actions, other sensory stimuli). The basic learning frameworks we describe could in principle be applied to a scenario with many more targets for words, no stability of word meaning across time, or vast individual differences in language use (whether due to bilingualism or



**Figure 2.** A schematic view of the intentional framework, linking learner and speaker via the two learning problems: guessing the speaker’s intended referent and guessing the lexicon of the language.

as a set of discrete links or continuous, probabilistic associations. The challenge for the child is that each individual context does not uniquely determine a set of lexical mappings. Which words go with which objects?

We can notate this problem of lexicon learning as a problem of Bayesian inference, that is, of inferring the most probable lexicon given the set of observed contexts:

$$P(L|C) \propto P(C|L)P(L) \quad (1)$$

The taxonomy in Figure 1 provides generative processes for four kinds of models that have been applied to this learning problem. The simplest approach to this problem is simply to estimate these probabilities, neglecting intentions or social cues. The model in the upper left corner, which assumes that words are generated via the observed objects in the context and the unobserved lexicon, can be written

$$P(L|C) \propto P(W|O, L)P(L). \quad (2)$$

If we represent the words in a context as the set  $W_C$  and the objects as  $O_C$ , we can expand this expression to

---

even simply random variation), but there is no guarantee that they would be sufficient. In practice, whether these assumptions are inborn or discovered, it is likely that they are necessary for learning to proceed.

$$P(L|C) \propto \prod_c P(W_c|O_c, L)P(L). \quad (3)$$

In other words, the probability of the lexicon under these models is the product across contexts of the probability of the words in the context, given the objects in the context and the lexicon. This “pure cross situational” approach is followed by a number of influential models (Fazly et al., 2010; Yu & Ballard, 2007).<sup>4</sup>

### *Differentiating associative and intentional models*

In this section, we compare the assumptions made by associative and intentional models, differentiating between the upper left and upper right panels of Figure 1. Following the approach above, the next step in defining a word learning model is to define the likelihood of a word being uttered, given the presence of some object and its lexical entry: the term  $P(W_C|O_C, L)$ . There are two important sub-problems that arise in defining this term. First, we must define the alignment between particular words and objects (assuming that there are multiples of each in each situation). Second, we need to define the probability of a word being used with a particular object (given that they are aligned). I will discuss only the first of these here, since the details of assigning probabilities are covered in several of the source articles for these models.<sup>5</sup>

It is with respect to the problem of aligning words and objects that the differences between associative and intentional models becomes clear. Associative models typically make minimal assumptions about alignment and assume that any

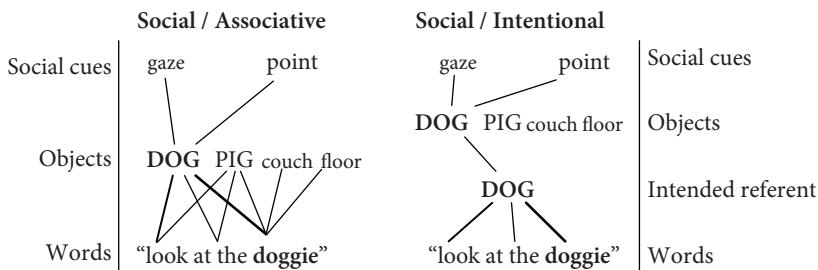
---

4. This modeling work has built directly on work on machine translation that attempted this optimization problem for aligned corpora (e.g. where words and objects are actually words in a target language and words in a source language; Brown, Pietra, Pietra, & Mercer, 1993).

5. The intentional assumption (the focus of this section) is orthogonal to the precise mechanics of how a model assigns probabilities to particular lexical mappings. Our initial work used a fairly discrete likelihood function to determine the probability that a particular word was used, given that the speaker had an intention to refer to some object:

$$P(w|I=o_c, L) = \begin{cases} \propto 1 & \text{if } L(w_c, o_c) = 1 \\ 0 & \text{otherwise} \end{cases} \quad (4)$$

where  $L(w_c, o_c) = 1$  indicated that  $w_c$  and  $o_c$  were linked in the lexicon. But it is equally possible to define a more clearly probabilistic function, using e.g. a multinomial distribution (which would then be conjugate to a dirichlet prior). More generally, it is likely very difficult to differentiate between a continuous lexical representation and a posterior distribution representing uncertainty over a discrete lexicon. Although our work on this topic is occasionally cited as providing a “hypothesis testing” (discrete) view of word learning, I see the discreteness of the lexicon in that particular model as an implementation choice rather than one that carries any particular theoretical weight.



**Figure 3.** A schematic of a single situation for associative and communicative models that both use social and prosodic information. Gaze and pointing cues signal that a dog toy is more salient than a pig toy. Prosodic focus on the word “doggie” raises its salience. As a consequence, the strongest association is between the object dog and the word “doggie” for both models. The communicative model includes a filtering step in which dog is assumed to be the correct referent. Salience is shown by type weight and size, while associative weights are shown by line weight.

word can be aligned with any object. In some sense, this is the basic tenet of an associative model: that all words and objects present in the context are associated with one another to some degree.

In an intentional model, in contrast, we assume that the relationship between words and objects is mediated by the speaker’s intention to refer to some set of objects. This mediation relationship is shown graphically by the intervening node between  $O$  and  $W$  in the generative process for models on the right side of Figure 1. The concept of the speaker’s intention to communicate mediates the relationship between the physical context and the words produced by that speaker. This notion of intention in our framework also corresponds to the notion of an agent’s goal, which plays a central role in work on social learning and rational action inference (Baker, Saxe, & Tenenbaum, 2009; Gergely, Bekkering, & Király, 2002; Shafto, et al., 2012). This intuition is shown graphically in Figure 3, which shows the mediating relationship that the speaker’s intention can play in learning from a single situation.

Formally, this mediation relationship results in a revision to Equation 2, where we notate this mediating intention  $I_C$ :

$$P(L|C) \propto P(W_C|O_C I_C L)P(L). \quad (5)$$

The addition of this mediating variable affects the process of finding the alignment between words and objects: While associative models assume that all words are linked to all objects, the intentional models assume that there is an extra step that removes some of these associations from consideration. In our work on cross-situational learning to date, the representation of the speaker’s intention has been quite basic, representing the speaker’s menu of possible intentions as the set of objects in the

context (making these *referential intentions*).<sup>6</sup> This restriction implicitly implemented a “whole object” assumption (Markman, 1991); but we reconsider this restriction in more recent work (described in the section on pragmatic inference, below).

Formally, in our initial model, we specified  $I_C$  as containing a subset of  $O_C$ , corresponding to the assumption that the speaker could talk about any subset of the objects in the context, including the empty set (Frank, Goodman, & Tenenbaum, 2009). Although this assumption allowed us to consider a wide variety of possibilities, if there were many objects present in a context it quickly became unwieldy, since it required considering the power set of objects in the context (which grows at  $2^n$  where  $n$  is the number of objects in  $c$ ). Hence, in more recent work we have begun using the simplifying assumption that  $I_C$  is a single object in  $O_C$  (Johnson, Demuth, & Frank, 2012), congruent with our empirical observation that most of the caregivers’ utterances in a corpus referred to at most one object.

Under an intentional assumption, regardless of how the intended referent is chosen, we can define the likelihood of a word as the product of two terms: the probability of the words given the intention, and the probability of a particular intention:

$$P(W_c | O_c, I_c, L) = \sum_{I_c \in O_c} P(W_c | I_c, L) P(I_c | O_c). \quad (6)$$

Thus, intentional models describes a two-step process of uttering a word: first decide which object to refer to, then decide the words to use to refer to it.

The key difference between associative and intentional models on our account is this two-step process, separating the choice of what to talk about from the choice of how to refer. Both types of models allow for information to “weight” the learner’s estimate of which objects are most salient, for example via social information. But in intentional models, this weighting influences the learner’s guess about which object(s) are being referred to. The definitional assumption of intentional models is that speakers have a discrete intention for each utterance, even if the learner is uncertain of what this intention is. In contrast, in associative models, all words are associated with all objects. Implicitly, this assumption is tantamount to assuming that all objects are being referred to and all words have referential status (just to greater or lesser degrees).

The intentional assumption – that there is a discrete choice of intended referent by a speaker – implies that one major part of word learning is in-the-moment

---

6. This construct has the potential to be far more flexible and powerful than the use to which we have so far put it. Provided that this distribution over intentions is limited by context, discourse, and other pragmatic factors, models could consider a much wider variety of possible interpretations – including interpretations not involving grounded reference to the current context.

interpretation (as posited in many pragmatic accounts of language learning, especially E. Clark, 2003 and Tomasello, 2003). If a learner knows what object is being talked about by the speaker (their intention, in our loose terminology), there is no need to compute associations between the words that are heard and the other objects that are present. In the language of causal models, the intention “screens off” the physical context from the words: knowing the speaker’s intended referent is enough for learning. In this respect, the intentional framework is deeply related to recent computational work by McMurray, Horst, and Samuelson (2012), who emphasized the role of learners’ interpretations of reference in the moment in longer-term word learning.

In Frank, Goodman, and Tenenbaum (2009), we reported results based on running associative and intentional models on a small, hand-annotated corpus of infant-directed speech. The intentional model outperformed other models in the lexicon it learned. This success was due at least in part to its ability to “filter out” spurious associations between function words and objects. Under the intentional model, evidence that a word was mapped to an object in one situation could help constrain hypotheses about which object was being referred to in another situation. This mutual constraint meant that irrelevant co-occurrences could be explained away, rather than providing noise (as in the associative models).

### *Adding social information*

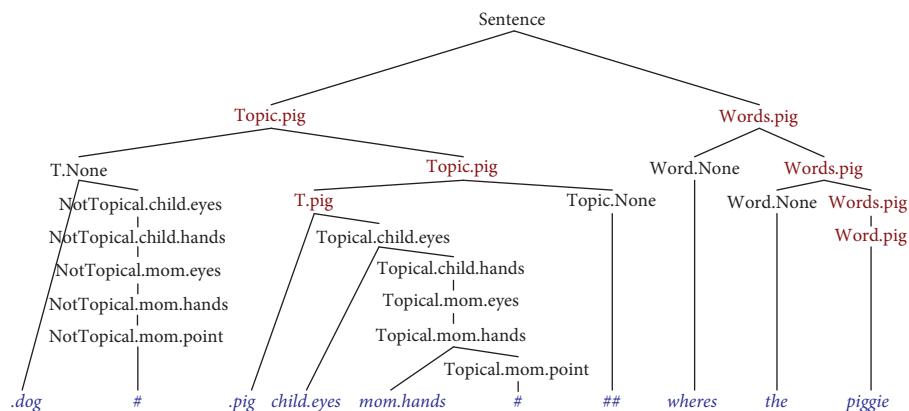
It is relatively rare for a speaker to talk about a physically present referent without giving some signal – at least at some point in the conversation – that the referent is indeed the one being talked about. Speakers gaze to conversational referents during language production both as an explicit social signal (H. Clark, 1996) and as a consequence of processes underlying language production (Griffin & Bock, 2000). In addition, speakers often signal reference by pointing. “Social cues” like eye-gaze or points to conversational referents are often cited as an important source of evidence for learning word-object mappings (Baldwin, 1993; Bloom, 2002; Hollich, Hirsh-Pasek, & Golinkoff, 2000; St. Augustine, 397/1963).

Social cues can be represented in an associative model as cues to which objects are most salient in a particular situation (Yu & Ballard, 2007). The presence of a point or gaze on a particular object endows it with some additional salience, which in turn strengthens its associations with co-occurring words. In the generative process for such a model, social cues are a reflection of the underlying salience of individual objects (Figure 1, lower left). If social cues are present, object salience is inferred to be higher. Figure 3, left side, shows a caricature of what the data for a single situation might look under such a model. The associative weights between words and objects are determined by the social cues and perceptual salience of the objects in the scene (as well as the prosodic salience of the words, which we do not

discuss here). The result is that the same associative computation is performed, but over a word/object set whose weights are no longer uniform.

In contrast, in an intentional model (Figure 1, lower right), social information informs the process of interpretation (deciding which object, if any, is being referred to). Social cues are generated by the speaker's intention to refer and hence are signals to that underlying intention (Figure 3, right). This interpretive, inferential use of social cues can be implemented in a number of ways. In our early work on this topic, we assumed that each cue could be the consequence of a relevant intentional action or could be the result of baseline looking without an underlying intention, and estimated these probabilities via a "noisy-or" model. This formulation allowed the model to learn that, for example, even though speakers' gaze was a frequent cue, its overall reliability was low (Frank, Goodman, & Tenenbaum, 2007).

Our more recent work embeds these social cue probabilities within a probabilistic, grammar-based formalism. In Johnson et al. (2012), we used an adaptor grammar, a probabilistic context free grammar that allows common structures to be reused efficiently (Johnson, Griffiths, & Goldwater, 2007). Consider the example shown in Figure 4. The observed input representation of the situation (shown in blue italics at the bottom of the tree) specifies that the words "where's the piggie" are observed along with dog and pig toys, and the pig toy is marked by two social cues, the child's eyes and the mom's hands.<sup>7</sup> The tree above it shows a possible parse of the situation. The extraneous dog toy is parsed as "non-topical" (e.g., not an intended



**Figure 4.** A parse tree for an entire situation, including a sentence along with its referential context and social cues. The sentence generates topical (referential) objects, the social cues that mark these objects, and the words that refer to them. Topic-specific words are marked in red and observed data are in blue. In this case, the referent is "pig" and referential words are propagated throughout the tree. Figure reprinted from Johnson et al. (2012).

7. We leave aside here the issue of whether the child's own eyes can be considered a "social cue" – this issue is discussed at length in Frank, Tenenbaum, and Fernald (2013).

referent, marked as *T.none*) and the pig toy and its accompanying social cues are generated on the basis of the topic. The words are also generated by the same topic, with several “non-topical” words (*Word.none*) followed by a word generated from the topical lexicon (marked as *Word.pig*).

This grammatical formalism, although distinct from our previous work in some of its details, still encodes the same two-stage computation we have associated with intentional models. The first decision is choosing the intention (topic, in the language of this model) underlying a sentence. This decision affects all other aspects of the sentence including the probabilities of both the social cues and the individual words, whose choice together constitutes the second decision: how to refer. The only difference is that these two “stages” are not broken out as separate computations but instead are instantiated as parts of the model’s grammatical representation. The intention or topic is represented through the topical productions in the parse tree, while the individual decisions about words in the sentence are represented through the choices of leaf nodes in the tree.

When we evaluated the grammatical model on a corpus tagged with social information, we found significant gains in the accuracy of both guessing the referents of utterances and the words associated with particular objects on the basis of adding the social cues. In addition, our analyses showed that the child’s own gaze on an object was the most predictive cue, suggesting that the corpus we used contained significant follow-in labeling by parents (replicating descriptive results from Frank et al., 2013).

To summarize: both associative and intentional models allow for the inclusion of social information, but associative models allow for social information to make particular referents more salient and bias the computation of associations between words and objects. In contrast, intentional models go beyond this interpretation of social cues as signals of salience and allow the social information to bias the computation of reference. Yet all of the models described here still treat only the mapping problem, implicitly equating referent identification with meaning learning. In this next section, we broaden the set of possible word meanings we consider, beginning the process of differentiating word meaning from reference.

### Adding pragmatic inference to intentional models

In the Quinian (1960) framing of the word learning problem, even if a single word is heard alongside a single object, there are still an infinite number of possible interpretations for the word. Quine distinguished between interpretations for which co-occurrence or pointing – the information sources relied on in the models above – failed to provide any traction for learners and those for which these cues might

in principle be informative. This first class includes, for example, the ambiguity between "rabbit" and "undetached rabbit parts"; Quine argues that these meanings may be impossible to distinguish. On the other hand, the second class contains many word meanings that may be empirically distinguishable but are likely to be confounded in any given context. To take a small set of the many possible modes of reference, a particular rabbit might be talked about as a "rabbit," but also as "white," (when the contrast is a brown rabbit), "animal" (when pointing out something in the bushes), or "small" (when the contrast is a larger rabbit).

Since basic-level object names are common in speech to children (Callanan, 1985), it should not be difficult to learn a word like "rabbit" from co-occurrence. The prospects for noticing the co-occurrence between animacy and "animal" seem somewhat lower; they are likely lower still for a color like "white," and close to nil for a gradable adjective like "small" (whose meaning changes from context to context). Information about the context of reference might provide a far more straightforward path to learning such terms (perhaps along with syntactic information, in the case of adjectives; Waxman & Booth, 2001). In this situation, a pragmatic word learner has an important advantage over any of the cross-situational learners described above: She can consider the context of use and the goal of the speaker in uttering a particular phrase, and crucially, she can consider why a term contrasting with the conventional descriptor is used (E. Clark, 1988).<sup>8</sup> This is the intuition that I will follow in this last section.

On a standard cross-situational view, pragmatic inferences belong to an entirely different class than the associative inference that leads the child to consider that "rabbit" = RABBIT. Our communicative/intentional framework provides a way to integrate these two inferences, however. Because long-term learning is mediated by in-the-moment interpretation, learners can use pragmatic computations to inform their guesses about what words refer and what objects (and even aspects of these objects) are being referred to. In this section I will show an example of our recent work modeling pragmatic inference in context, and then demonstrate that this kind of model of pragmatic inference can be integrated with the intentional word learning framework described above. There is much work yet to do, but I believe that this combined framework provides the beginnings of the tools necessary to extend the cross-situational paradigm beyond simple word-object mapping.

8. Critically, in the models described here, there is no requirement that words *must* be learned by one route or another. Much of the vocabulary that children acquire is no doubt learned ostensively or in socially-constrained situations that do not require inference. An important goal for future research is to quantify the contributions of different learning mechanisms to the growth of children's vocabulary (see e.g., Mitchell & McMurray, 2009 for an example of this kind of "macro-economic" analysis).

### *Modeling pragmatic disambiguation of reference*

Grice (1975) proposed a set of maxims for normative communication: Speakers should be truthful, relevant, clear, and informative. On Grice's account, listeners should interpret utterances as though these maxims are being followed, allowing them to go beyond the truth-functional meanings of the words in the sentences to derive richer meanings in context that he called "implicatures." In our work we have explored the idea that some of these implicatures can be captured in formal models (Frank & Goodman, 2012). We focus here on the maxim of informativeness and model listeners as doing statistical inferences about what a speaker's intended referent is, given a presumption of informativeness.

Our model assumes that in a context  $C$ , the listener is attempting to infer the speaker's intention  $I$  as one of a set of possible referents.<sup>9</sup> The listener considers two factors: first, the relative informativeness of the speaker's utterance with respect to each of the referents, and second, the "contextual salience" of the referents. Contextual salience here refers to the relative probability of reference, given the conversational context, the shared knowledge between communicators, and other factors that jointly determine that a particular referent will be the object of the speaker's expression. Although this quantity could perhaps be derived *a priori*, we have measured it empirically in our previous work (Frank & Goodman, 2012).

This two-part model can be notated as follows:

$$P(I|w, L, C) \propto P(w|I, L, C)P(I, C) \quad (7)$$

with the two terms on the right corresponding to the two factors being considered (informativeness and contextual salience). We pursue the idea that the informativeness of a word in context is inversely proportional to its specificity, so that

$$P(w|I, L, C) = \begin{cases} \propto \frac{1}{|w|} & \text{if } w(I) = 1 \\ 0 & \text{otherwise} \end{cases} \quad (8)$$

where  $|w|$  notates the number of objects in the context that can be referred to using  $w$  and  $w(I) = 1$  if word  $w$  is true of  $I$ .<sup>10</sup> This inverse proportionality corresponds to

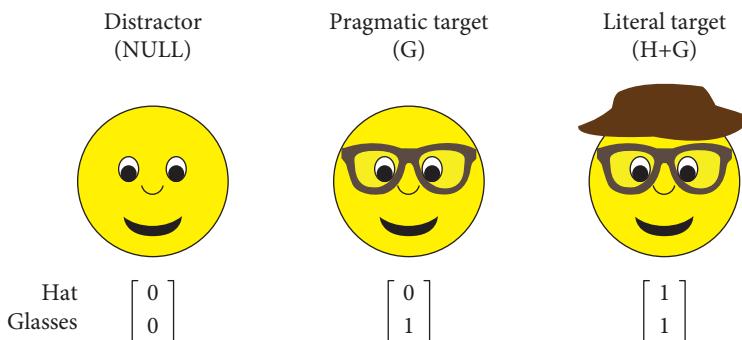
9. Note that in our other work we use  $R_{-S}$  to notate the speaker's intended referent, rather than  $I$  – we use  $I$  here for consistency with previous sections. As before, intentions are cashed out as referents, though again we believe that this framework could in principle be extended to consider intended *meanings*.

10. This set of definitions implicitly assumes that words are no longer names of objects. Instead, they are functions that can be applied to a context and that return true or false for each object in

the “size principle” of Tenenbaum & Griffiths (2001): A word is more informative if its use is “less coincidental” because it better picks out the intended referent from the context. Note here the parallel with Equation 4: In some sense, we are just replacing an uninformative speaker (where probability is proportional to 1) with an informative speaker (where probability is proportional to extension in context).

An example is useful in clarifying how this model setup naturally leads to Gricean pragmatic implicatures. Consider the situation pictured in Figure 5 (originally from Stiller, Goodman, & Frank, 2011). There are three possible referents shown, referred to below as NULL, G, and H + G. Each has different features, which lead to different possible expressions that can be used to refer to it. For the purpose of this game, we assume that this set is limited to the expressions “hat” and “glasses.” A speaker utters the word “glasses”; the job of the listener is to decide which face is being referred to. Many people share the intuition that the intended referent is G, the face without a hat.

This simple pragmatic inference has many of the elements of scalar implicature (the often-studied inference that “some” typically is strengthened pragmatically to mean “some but not all”), so it is a useful case study of how our model can be applied. In order to simplify the computation, we assume that the contextual salience of the three faces is even, and that the NULL referent is not considered. We can compute the strength of the pragmatic inference that “glasses” refers to G using Equation 7 and expanding the proportionality by normalizing over all possible referents:



**Figure 5.** An example stimulus from our pragmatic inference experiments. Participants would be asked to identify the referent of a phrase containing “glasses” as the descriptor; given this message, the middle face (who has glasses but no hat) is the pragmatic implicature target, whereas the right-hand face (who has a hat and hence has a better potential descriptor) is the literal target.

---

the context. This truth-functional model of word meaning can easily be used to capture predicates like “white” or “furry” and is in principle extensible to context-dependent adjectives like “small” (Schmidt, Goodman, Barner, & Tenenbaum, 2009).

$$\begin{aligned}
 P(G | "glasses", C) &= \frac{P("glasses" | G, C)}{\sum_{I \in C} P("glasses" | I, C)} \\
 &= \frac{P("glasses" | G, C)}{P("glasses" | G, C) + P("glasses" | H + G, C)}
 \end{aligned} \tag{9}$$

We can then expand this expression, using Equation 8 and notating the set of vocabulary items that can be used to describe e.g., item  $G$  as  $w \in G$ :

$$\begin{aligned}
 P(G | "glasses", C) &= \frac{\frac{1}{|“glasses”|}}{\frac{1}{\sum_{w' \in G} \frac{1}{|w'|}}} \\
 &= \frac{\frac{1}{|“glasses”|}}{\frac{1}{\sum_{w' \in G} \frac{1}{|w'|}} + \frac{1}{\sum_{w' \in H+G} \frac{1}{|w'|}}} \\
 &= \frac{\frac{1}{|“glasses”|}}{\frac{1}{|“glasses”|} + \frac{1}{\frac{|“glasses”|}{|“glasses”| + |“hat”|}}} = \frac{\frac{1/2}{1/2}}{\frac{1/2}{1/2} + \frac{1/2}{1/2 + 1}} = .75
 \end{aligned} \tag{10}$$

In other words, the probability of  $G$ , the face with glasses and no hat, given the descriptor “glasses,” is predicted to be .75 (and hence the probability of  $H + G$  is .25). This computation encodes the intuition that, had the speaker been talking about  $H + G$ , he or she would have chosen the more specific descriptor “hat.” This prediction of an implicature in favor of  $G$  corresponds well with the judgments of both adults and preschoolers (Stiller et al., 2011).

As illustrated above, our pragmatics model provides a framework for quantifying the Gricean maxim “be informative” from the perspective of both a speaker and a listener. It and its extensions can provide an account for a wide variety of pragmatic phenomena (Frank & Goodman, 2012; Goodman & Stuhlmüller, 2013).

The pragmatic implicature under this model can be computed for any situation in which the set of contextual and vocabulary alternatives is known, although the problem of identifying relevant alternative is still an open research challenge. In the next section, we illustrate how this pragmatic framework can be integrated with the intentional approach described above.

### *Using informativeness to learn words*

Our pragmatics model is deeply related to the intentional communication model described above. Recall that in Equation 6 of the cross-situational learning model, we defined the probability of a particular word being uttered, given some intention and context as a product of two terms: the probability of the word given the intention, and the probability of the intended referent given the objects in the context. This relation was stated as  $P(w|O, I, L) = P(w|I, L)P(I|O)$ . Intuitively, these two terms govern the probability of choosing a particular referent and choosing the proper referring expression. Note now that the pragmatic model described above uses the same breakdown of the process of inferring reference. The “contextual salience” term we described above maps directly onto the referent choice term  $P(I|O)$ , and the Gricean informativeness term in Equation 8 maps onto the term  $P(w|I, L)$ .

With this equivalence in hand, we can reverse the pragmatics model and derive a word-learning version (a version of this derivation is given in Frank, Goodman, Lai, & Tenenbaum, 2009 and Frank & Goodman, under review). In this version, we infer alternative meanings for a particular lexical item, given that a particular referent is known to be uttered. Consider the display in Figure 5. Imagine that a speaker pointed now to the literal target H + G, fixing the referent, but uttered a novel label, e.g., “fedora.” In this case, the referent is known, but the meaning of a novel element in the lexicon  $L$  is unknown. We notate the possibility that a particular word in  $L$  refers to a feature (e.g., having a hat) as  $w = f$ . Using the same formulation from the intentional model above, we can write

$$\begin{aligned} P(L|I, w, C) &\propto P(I|L, w, C)P(L) \\ &\propto P(I|w=f, w, C)P(w=f) \end{aligned} \tag{11}$$

and if we assume that there is no prior reason to prefer one meaning for  $w$  over another ( $P(w=f) \propto 1$ ) and substitute from Equation 8, then we have

$$P(L|I, w, C) \propto \frac{1}{|f|}. \tag{12}$$

In other words, all else being equal, a word is likely to have the meaning that would be informative in this context. So “fedora” would be more likely to refer to the hat (the feature that is most informative in this display) than the glasses. Thus, the pragmatic model described above can be used to capture inferences about the likely meanings of words in context.

In Frank, Goodman, Lai, and Tenenbaum (2009), we presented data that this relation in fact fit adult participants’ judgments about novel adjective meanings with high accuracy. We used schematic displays of shapes that reproduced the same kind of game as shown in Figure 5 and varied the number of shapes with each property (in Figure 5 this would be the number of faces with hats vs. the number with glasses). As the relative extensions changed, participants altered their guesses about novel adjective meanings, suggesting that (at least as a group) they were sensitive to the relative informativeness of different possible word meanings. Although this work is still ongoing, we now have some data from young children that they are able to make such judgments as well (Frank & Goodman, under review).

## Conclusions

Eve Clark’s (2003) book pushed me powerfully in the direction of considering communication as the grounding experience of language learning. The theoretical work that I have described here is my attempt to capture computationally some of the insights of Eve’s perspective, particularly the fundamental fact that children are not just passive absorbers of linguistic input but are participants in conversation. The consequence of this fact is that the task of understanding language in the moment – and ascertaining reference in particular – must interact with language learning.

The interaction between interpretation and learning leads to a class of models that I have referred to as communicative or intentional models. In the taxonomy described above, these models are dissociated from associative models not because of the information they include – both model classes can take advantage of social information – but because of the way they break down the learning task. While associative models use social information to bias associative learning, intentional models use social information (as well as pragmatic inference) to inform a guess about what speakers are trying to say. And of course, underlying this class of models is learners’ assumption that speakers are rational agents (Clark, 1988).

I have argued here that intentional models are a powerful framework for using social and pragmatic information in the service of learning the meanings of words. The evidence is strong that by the age of 18 months, children take advantage of this information and make inferences that go beyond the association of words and

objects (Baldwin, 1993; Hollich et al., 2000). But since recent evidence indicates the possibility of word knowledge in even younger children than had previously been suspected (Bergelson & Swoley, 2012), infants' assumptions during this early learning remain an important open question for both empirical and computational investigation.

Independent of whether infants begin life as probabilistic intentional learners, it seems likely that they converge to this position as their vocabulary expands to encompass terms that cannot be learned via contextual associations. For a learner who only need acquire basic level descriptors in a word of repeated exposures, the consequences of intentional learning are relatively modest. But for a learner of a language that contains a wide variety of complex, context-dependent predicates – in other words, a human child – it is essential to understand the contribution of the speaker's communicative intentions to the words they utter.

## References

- Akhtar, N., & Montague, L. (1999). Early lexical acquisition: The role of cross-situational learning. *First Language*, 19, 347–358. DOI: 10.1177/014272379901905703
- Baker, C., Saxe, R., & Tenenbaum, J. (2009). Action understanding as inverse planning. *Cognition*, 113, 329–349. DOI: 10.1016/j.cognition.2009.07.005
- Baldwin, D. (1993). Early referential understanding: Infants' ability to recognize referential acts for what they are. *Developmental Psychology*, 29, 832–843. DOI: 10.1037/0012-1649.29.5.832
- Baldwin, D. (1995). Understanding the link between joint attention and language. In C. Moore & P. Dunham (Eds.), *Joint attention: Its origins and role in development* (pp. 131–158). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Bergelson, E., & Swoley, D. (2012). At 6–9 months, human infants know the meanings of many common nouns. *Proceedings of the National Academy of Sciences*, 109, 3253–3258. DOI: 10.1073/pnas.1113380109
- Bloom, P. (2002). *How children learn the meanings of words*. Cambridge, MA: The MIT Press.
- Brown, P., Pietra, V., Pietra, S., & Mercer, R. (1993). The mathematics of statistical machine translation: Parameter estimation. *Computational Linguistics*, 19, 263–311.
- Callanan, M. (1985). How parents label objects for young children: The role of input in the acquisition of category hierarchies. *Child Development*, 56, 508–523. DOI: 10.2307/1129738
- Carpenter, M., Nagell, K., & Tomasello, M. (1998). Social cognition, joint attention, and communicative competence from 9 to 15 months of age. *Monographs of the Society for Research in Child Development*, 63(4/Serial no. 255), 1–174. DOI: 10.2307/1166214
- Clark, E.V. (1988). On the logic of contrast. *Journal of Child Language*, 15, 317–335. DOI: 10.1017/S0305000900012393
- Clark, E.V. (1990). On the pragmatics of contrast. *Journal of Child Language*, 17, 417–431. DOI: 10.1017/S0305000900013842
- Clark, E.V. (2003). *First language acquisition*. Cambridge: CUP.
- Clark, H.H. (1996). *Using language*. Cambridge: CUP. DOI: 10.1017/CBO9780511620539

- Csibra, G., & Gergely, G. (2009). Natural pedagogy. *Trends in Cognitive Sciences*, 13, 148–153. DOI: 10.1016/j.tics.2009.01.005
- Fazly, A., Alishahi, A., & Stevenson, S. (2010). A probabilistic computational model of cross-situational word learning. *Cognitive Science*, 34, 1017–1063. DOI: 10.1111/j.1551-6709.2010.01104.x
- Fisher, C., Hall, D., Rakowitz, S., & Gleitman, L. (1994). When it is better to receive than to give: Syntactic and conceptual constraints on vocabulary growth. *Lingua*, 92, 333–375. DOI: 10.1016/0024-3841(94)90346-8
- Frank, M.C., & Gibson, E. (2011). Overcoming memory limitations in rule learning. *Language Learning and Development*, 7, 130–148. DOI: 10.1080/15475441.2010.512522
- Frank, M.C., Goldwater, S., Griffiths, T.L., & Tenenbaum, J.B. (2010). Modeling human performance in statistical word segmentation. *Cognition*, 117, 107–125. DOI: 10.1016/j.cognition.2010.07.005
- Frank, M.C., & Goodman, N.D. (2012). Predicting pragmatic reasoning in language games. *Science*, 336, 998. DOI: 10.1126/science.1218633
- Frank, M.C., & Goodman, N.D. (Under review). Inferring word meanings by assuming that speakers are informative.
- Frank, M.C., Goodman, N.D., Lai, P., & Tenenbaum, J.B. (2009). Informative communication in word production and word learning. In N. Taatgen & H. van Rijn (Eds.), *Proceedings of the 31st Annual Conference of the Cognitive Science Society*, (pp. 1228–1233). Amsterdam: Cognitive Science Society.
- Frank, M.C., Goodman, N.D., & Tenenbaum, J.B. (2007). A Bayesian framework for cross-situational word learning. *Advances in Neural Information Processing Systems*, 20.
- Frank, M.C., Goodman, N.D., & Tenenbaum, J.B. (2009). Using speakers' referential intentions to model early cross-situational word learning. *Psychological Science*, 20, 578–585. DOI: 10.1111/j.1467-9280.2009.02335.x
- Frank, M.C., Tenenbaum, J.B., & Fernald, A. (2013). Social and discourse contributions to the determination of reference in cross-situational word learning. *Language Learning and Development*, 9, 1–24. DOI: 10.1080/15475441.2012.707101
- Gallistel, C. (1990). *The organization of learning*. Cambridge, MA: The MIT Press.
- Gallistel, C., Fairhurst, S., & Balsam, P. (2004). The learning curve: Implications of a quantitative analysis. *Proceedings of the National Academy of Sciences*, 101, 13124–13131. DOI: 10.1073/pnas.0404965101
- Geisler, W. (2003). Ideal observer analysis. In L.M. Chalupa & J.S. Werner (Eds.), *The visual neurosciences* (pp. 825–837). Cambridge, MA: The MIT Press.
- Gergely, G., Bekkering, H., & Király, I. (2002). Rational imitation in preverbal infants. *Nature*, 415, 755. DOI: 10.1038/415755a
- Gillette, J., Gleitman, H., Gleitman, L., & Lederer, A. (1999). Human simulations of vocabulary learning. *Cognition*, 73, 135–176. DOI: 10.1016/S0010-0277(99)00036-0
- Gleitman, L. (1990). The structural sources of verb meanings. *Language Acquisition*, 1, 3–55. DOI: 10.1207/s15327817la0101\_2
- Goodman, N.D., & Stuhlmüller, A. (2013). Knowledge and implicature: Modeling language understanding as social cognition. *Topics in Cognitive Science*, 5, 173–184. DOI: 10.1111/tops.12007
- Gopnik, A. (2012). Scientific thinking in young children. Theoretical advances, empirical research and policy implications. *Science*, 337, 1623–1627. DOI: 10.1126/science.1223416
- Grice, H. (1975). Logic and conversation. *Syntax and Semantics*, 3, 41–58.

- Griffin, Z., & Bock, K. (2000). What the eyes say about speaking. *Psychological Science*, 11, 274–279. DOI: 10.1111/1467-9280.00255
- Hollich, G., Hirsh-Pasek, K., & Golinkoff, R. (2000). Breaking the language barrier: An emergentist coalition model for the origins of word learning. *Monographs of the Society for Research in Child Development*, 65, 1–135. DOI: 10.1111/1540-5834.00091
- James, W. (1890). *The principles of psychology*, Vol. 1. New York, NY: Henry Holt and Company. DOI: 10.1037/11059-000
- Johnson, M., Demuth, K., & Frank, M. (2012). Exploiting social information in grounded language learning via grammatical reductions. In *Proceedings of the Association for Computational Linguistics* (pp. 883–891).
- Johnson, M., Griffiths, T., & Goldwater, S. (2007). Adaptor grammars: A framework for specifying compositional nonparametric Bayesian models. *Advances in Neural Information Processing Systems*, 19, 641–648.
- Li, P., Farkas, I., & MacWhinney, B. (2004). Early lexical development in a self-organizing neural network. *Neural Networks*, 17, 1345–1362. DOI: 10.1016/j.neunet.2004.07.004
- Locke, J. (1690/1964). *An essay concerning human understanding*. Cleveland, OH: Meridian Books.
- Markman, E.M. (1991). *Categorization and naming in children: Problems of induction*. Cambridge, MA: The MIT Press.
- Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information*. New York, NY: Henry Holt & Co.
- McMurray, B., Horst, J.S., & Samuelson, L.K. (2012). Word learning emerges from the interaction of online referent selection and slow associative learning. *Psychological Review*, 119, 831–877. DOI: 10.1037/a0029872
- Medina, T., Snedeker, J., Trueswell, J., & Gleitman, L. (2011). How words can and cannot be learned by observation. *Proceedings of the National Academy of Sciences*, 108, 9014–9019. DOI: 10.1073/pnas.1105040108
- Mitchell, C., & McMurray, B. (2009). On leveraged learning in lexical acquisition and its relationship to acceleration. *Cognitive Science*, 33, 1503–1523. DOI: 10.1111/j.1551-6709.2009.01071.x
- Onishi, K., & Baillargeon, R. (2005). Do 15-month-old infants understand false beliefs? *Science*, 308, 255–258. DOI: 10.1126/science.1107621
- Perfors, A., Tenenbaum, J., Griffiths, T., & Xu, F. (2011). A tutorial introduction to bayesian models of cognitive development. *Cognition*, 120, 302–321. DOI: 10.1016/j.cognition.2010.11.015
- Pinker, S. (1984). *Language learnability and language development*. Cambridge, MA: Harvard University Press.
- Quine, W. (1960). *Word and object*. Cambridge, MA: The MIT Press.
- Regier, T. (2005). The emergence of words: Attentional learning in form and meaning. *Cognitive Science*, 29, 819–865. DOI: 10.1207/s15516709cog0000\_31
- Sanborn, A., Griffiths, T., & Navarro, D. (2010). Rational approximations to rational models: alternative algorithms for category learning. *Psychological Review*, 117, 1144. DOI: 10.1037/a0020511
- Schmidt, L., Goodman, N.D., Barner, D., & Tenenbaum, J. (2009). How tall is tall? Compositionality, statistics, and gradable adjectives. In N. Taatgen & H. van Rijn (Eds.), *Proceedings of the 31st Annual Conference of the Cognitive Science Society*, (pp. 3151–3156).

- Schulz, L., Kushnir, T., & Gopnik, A. (2007). Learning from doing: Intervention and causal inference. In A. Gopnik & L. Schulz (Eds.), *Causal learning: Psychology, philosophy, and computation* (pp. 67–85). Oxford: OUP.
- Shafto, P., Goodman, N.D., & Frank, M. (2012). Learning from others the consequences of psychological reasoning for human learning. *Perspectives on Psychological Science*, 7, 341–351. DOI: 10.1177/1745691612448481
- Siskind, J. (1996). A computational study of cross-situational techniques for learning word-to-meaning mappings. *Cognition*, 61, 39–91. DOI: 10.1016/S0010-0277(96)00728-7
- Smith, L.B., & Yu, C. (2008). Infants rapidly learn word-referent mappings via cross-situational statistics. *Cognition*, 106, 1558–1568. DOI: 10.1016/j.cognition.2007.06.010
- St. Augustine. (397/1963). *The confessions of St. Augustine*. Oxford: Clarendon Press.
- Stiller, A., Goodman, N.D., & Frank, M. (2011). Ad-hoc scalar implicature in adults and children. In L. Carlson, C. Hoelscher, & T.F. Shipley (Eds.), *Proceedings of the 33rd Annual Meeting of the Cognitive Science Society*, (pp. 2134–2139).
- Tenenbaum, J., & Griffiths, T. (2001). Generalization, similarity, and Bayesian inference. *Behavioral and Brain Sciences*, 24, 629–640.
- Tomasello, M. (2003). *Constructing a language: A usage-based theory of language acquisition*. Cambridge, MA: Harvard University Press.
- 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, 126–156. DOI: 10.1016/j.cogpsych.2012.10.001
- Vouloumanos, A. (2008). Fine-grained sensitivity to statistical information in adult word learning. *Cognition*, 107, 729–742. DOI: 10.1016/j.cognition.2007.08.007
- Vouloumanos, A., Onishi, K., & Pogue, A. (2012). Twelve-month-old infants recognize that speech can communicate unobservable intentions. *Proceedings of the National Academy of Sciences*, 109, 12933–12937. DOI: 10.1073/pnas.1121057109
- Vouloumanos, A., & Werker, J. (2009). Infants' learning of novel words in a stochastic environment. *Developmental Psychology*, 45, 1611–1617. DOI: 10.1037/a0016134
- Waxman, S., & Booth, A. (2001). Seeing pink elephants: Fourteen-month-olds' interpretations of novel nouns and adjectives. *Cognitive Psychology*, 43, 217–242. DOI: 10.1006/cogp.2001.0764
- Yu, C., & Ballard, D. (2007). A unified model of early word learning: Integrating statistical and social cues. *Neurocomputing*, 70, 2149–2165. DOI: 10.1016/j.neucom.2006.01.034
- Yu, C., & Smith, L. (2012). Modeling cross-situational word-referent learning: Prior questions. *Psychological Review*, 119, 21. DOI: 10.1037/a0026182

# Word order as a structural cue and word reordering as an interactional process in early language acquisition

Aylin C. Küntay<sup>1,2</sup> and Duygu Özge<sup>1,3</sup>

<sup>1</sup>Department of Psychology, Koç University, Istanbul, <sup>2</sup>Educational and Learning Sciences, Utrecht University and <sup>3</sup>Department of Psychology, Harvard University

We present a critical review of the literature on how children exposed to flexible-word-order languages, especially Turkish, acquire word-order variation. We highlight two traditions of psycholinguistic research assuming different theoretical and methodological approaches, namely *language-as-product* and *language-as-action* views. While the former studies the underlying mechanisms of using word order as a structural cue in interpreting thematic roles, the latter focuses on how it is used to convey information structure embedded in communication and action. The review reveals that these seemingly independent views complement each other to better account for (i) how children come to interpret argument roles when the word order is not fixed and (ii) how they comprehend/use word-order variation as a pragmatic tool in communication.

## Introduction

Angel Chan recently interviewed Eve Clark about major developments in the field of child language in the last decade. Eve Clark (2012) responded, “we have now moved away from the formerly pervasive view of ‘language as product’ to a more realistic view of ‘language-as-process’” (IASCL, Child Language Bulletin, Volume 32). As always with Eve’s insights, this is an eye-opening and thought-provoking observation. Our chapter discusses two independently motivated approaches to studying children’s sensitivity and use of word order variation, allowing us to consider in juxtaposition evidence coming from *language-as-product* and *language-as-action* views.

Psycholinguistic research embodies two distinctly recognizable traditions contrastively dubbed as *language-as-product* and *language-as-action* (Clark, 1992;

Trueswell & Tannenhaus, 2005). The *language-as-product* approach has the goal of depicting syntactic representations underlying language comprehension and production, as part of a mission to discover how information-processing mechanisms underpin language knowledge (Fodor, 1983). On the other hand, the *language-as-action* approach locates linguistic behavior within a larger social communicative context, taking dialogue, situation, participants, and action as inseparable components of a communicative collaboration (Clark, 1992; 1996). These two distinct perspectives bring forth a bifurcation of preferred research methodologies (Tannenhaus & Trueswell, 2005). The *language-as-product* approach prefers experimental investigations of comprehension of typically isolated utterances displaying some single critical feature. In contrast, the *language-as-action* tradition typically focuses on naturalistic or context-embedded elicitations of linguistic production.

Studies of children's sensitivity to word order have mostly been conducted from the vantage point of the *language-as-product* tradition. A common developmental question is how early and with what degree of abstraction speakers or writers of natural languages utilize word order as a structural cue to meaning representations (Akhtar & Tomasello, 1997; Gertner, Fisher, & Eisengart, 2006). In these studies, sentences with different word orders (e.g., SOV vs. SVO) are presented to speakers of different ages both within and across languages to determine whether different meanings are assigned to differently ordered sentences. While the great majority of these studies conceive of the interpretation of word order as an end-sentence product constituted at the completion of a sentence, some investigate how interpretation is assigned dynamically as sentence elements unfold. The behavior tested is how the grammatical roles are assigned when the location of the arguments varies.

Here we first examine the crosslinguistic evidence about children's development of sensitivity to word order, gathered using the *language-as-product* approach. Next we turn to the far less studied effects of information structure in processes of word reordering during interactive communicative settings in caregiver-child dialogues. The bulk of our empirical examples come from Turkish, a free word order language that allows considerable ellipsis of nominal elements, with some evidence presented from other crosslinguistic research.

We propose that studies of word order variation taking place in naturalistic, context-motivated situations fruitfully complement experimental approaches that are purely concerned with the parsing of the underlying structure. Although the two lines of investigation are largely mutually independent, they need to be reconciled to give a full account of children's learning of word order. It is no wonder that studies should be concerned with how children eventually develop representations of word order as a guide to sentence meaning. However, it is the embodied joint interactions with others in actual situations that authentically generate the need

for word order alternations, and therefore provide a valuable set of cues for deriving abstract representational knowledge. Yet, studies about how children process and produce word order variation during conversational exchanges are surprisingly uncommon. It is paramount that we transport our experimental rigor to study the learner's challenges and opportunities in acquiring word order during the give-and-take that is typical of joint conversation.

### *Word order as a structural cue to participant roles in events*

Word order is often studied as a structural cue to sentence meaning, since the images conjured up when hearing *the bear chased the camper* and *the camper chased the bear* are so different – in a language like English, at least. The role of number and position of nouns is a conspicuous grammatical means of indicating *who* does *what* to *whom* in English (Naigles & Swensen, 2006; Yuan, Fisher, & Snedeker 2012). And there is evidence that English-learning infants display sensitivity to meaning differences between *Noun1-Verb-Noun2* and *Noun2-Verb-Noun1* as early as the middle of the second year of life (Fisher & Song, 2006; Golinkoff, Hirsh-Pasek, Cauley, & Gordon, 1987). Although the evidence is sparse for non-English-learning children, it is known that by 3 years of age, children use the most typical word order of their languages to derive meaning from sentences in languages as varied as Turkish, Mandarin, and Cantonese (Candan, Küntay, Yeh, Cheung, Wagner, & Naigles, 2012; Chan, Lieven, & Tomasello, 2009; Slobin & Bever, 1982). In these languages, unlike English, word order is not the only or the most reliable linguistic device to determine sentential meaning. Other features such as participant animacy, morphological markings on nouns, and prosody conspire with discourse properties to yield an interpretation of the participant roles in an event. Even in English, several studies have argued that animacy plays an important role in children's interpretation of word order (e.g., Cannizzaro, 2012; Chapman & Miller, 1975). Some cross-linguistic studies are presented below.

Slobin and Bever (1982) were the first to study cross-linguistically how young children (aged 24 to 52 months) learning two fixed word order languages (English and Italian) and two flexible word order languages (Turkish and Serbo-Croatian) encoded argument relations when arguments appeared in various orders. Children were asked to enact descriptions that had nouns differently sequenced in relation to the verb and to each other. Children acquiring all four languages used a *canonical sentence strategy*, employing the typical word order of their native language in assigning actor or undergoer roles to nouns mentioned in sentences describing transitive actions. Turkish learners interpreted the non-typical orders (NNV and NVN) correctly 70% of the time, paying attention to nominal inflection to determine the object NP; these are similar to results for Japanese, also a verb-final

language (Hakuta, 1982). Children acquiring Serbo-Croatian, another flexible word order language, tended to depend more on the typical order despite the presence of rich nominal case marking, acting more like Italian and English learners. The authors related this to the fact that Serbo-Croatian is not as regular as Turkish in terms of case morphology. First, both the subject and the object are overtly marked in Serbo-Croatian whereas subjects bear no overt case in Turkish, so that Turkish children need only to learn that accusative case marking applies consistently to the direct object in order to be able to assign correct participant roles, regardless of the order of the arguments. Besides, while Serbo-Croatian allows uninflected arguments (with appropriate stress) to appear in various orders, Turkish allows word order variation only when the object is overtly marked. Thus, while Turkish-speaking children might expect variation in a quite restricted domain (that the first NP would be more likely to be the subject if not marked overtly for the accusative case), Serbo-Croatians would have to consider a greater number of possibilities (e.g., case inflection or stress) in assigning participant roles. Overall, this early study of cross-linguistic word order comprehension suggested that there is no universally favored pattern for the acquisition of word order and that several factors act in competition to specify the most reliable cue of a language such as morphosyntax, lexical semantics, animacy or stress (Bates & MacWhinney, 1987). Yet, experimental research on such languages typically strips away all these co-occurring properties in order to study the effect of word order on sentence comprehension in isolation from all other linguistic, prosodic, and discourse-related cues.

In another act-out study with Turkish learners, Batman-Ratyosyan and Stromswold (1999) showed that 4-year-olds performed well on SOV and OVS orders with arguments lacking overt case marking, whereas 3-year-olds children performed significantly worse on OVS orders, showing an early preference for the typical order of SOV. Batman-Ratyosyan and Stromswold (2002) further investigated whether using pragmatically more felicitous contexts would improve Turkish children's comprehension of non-typical orders with no case marking cues, as shown in (1).

- (1) a. Context sentence (SV):

Bu oyun-da at-lar oyna-sın.  
This game-LOC horse-PL play-OPT  
“Horses play in this game.”

Trial sentence (OSV):

At bir fil it-sin.  
horse an elephant push-OPT  
“Let the horse push an elephant.”

## b. Context sentence (SV):

Bu oyun-da fil-ler oynasin.  
 this game-LOC elephant-PL play-OPT  
 “Elephants play in this game.”

## Trial sentence (OVS):

Bir fil it-sin at.  
 an elephant push-Opt horse  
 “Let the horse push an elephant.”

Children performed significantly better on SOV than on OVS sentences regardless of the contextual support. Interestingly, the youngest children at the age of two performed better in conditions without context, as though they did not benefit from the pragmatic boost provided by the context. This suggests that typical word order is initially acquired as a purely structural cue in assigning participant roles, and pragmatic cues do not improve young children’s interpretation of non-canonical orders. Importantly, however, language is always much more complicated, and possibly more informative, in contexts of natural communication than in controlled experiments. Specifically, in the examples in (1), from Batman-Ratyosyan and Stromswold (2002), the context provided in (1a) – *Horses play in this game* – does not satisfy the uniqueness presupposition<sup>1</sup> of the definite NP *the horse* in the trial sentence *Let the horse push an elephant*. It is highly likely that the child expects to hear about horses in the upcoming sentence, but if the speaker is not going to talk about all of the horses (i.e., as a group) mentioned in the context, s/he needs to identify which particular horse will be picked up in the upcoming sentence. In other words, it would be more appropriate to follow this context with a uniquely identifiable NP, either with a demonstrative pronoun (e.g., *that horse*) or a partitive construction (e.g., *one of the horses*), as illustrated in (2).

## (2) Context:

Bu oyun-da at-lar oyna-sin.  
 this game-LOC horse-PL play-OPT  
 “Horses play in this game.”

## Trial:

Şu at /at-lar-dan bir-i bir fil it-sin.  
 that horse /horse-PL-ABL one-ACC an elephant push-OPT  
 “Let that horse/one of the horses push an elephant.”

1. The context provided in (1a) (i.e., *horses play in this game*) does not establish a specific or unique horse in the context but rather several horses. Therefore, in the trial sentence it is not contextually appropriate to use a definite NP *at (the horse)*, which presupposes the existence of a uniquely identifiable horse referent.

Similarly, in (1b), given that the context mentions a set of elephants, the object in the trial sentence should either be preceded by a demonstrative or contain an accusative marked NP, which would receive a partitive interpretation, as in (3).

(3) Context:

Bu oyun-da fil-ler oyna-sın.  
this game-LOC elephant-PL play-OPT  
“Elephants play in this game.”

Trial:

Şu fil-i /fil-ler-den bir-i-ni it-sın at.  
that elephant-ACC /elephant-PL-ABL one-Poss3SG-ACC push-OPT horse  
“Let the horse push that elephant/one of the elephants.”

The findings in Batman-Ratyosyan and Stromswold (2002) contradict certain comprehension studies with children acquiring Japanese, another verb-final language, suggesting that non-typical orders are interpreted correctly when provided with appropriate discourse contexts around three years of age (Otsu, 1994). Sano (2004) showed that not only topicalizing discourse but also the definite marker *sono* are required for successful interpretation of the non-canonical OSV sentences in Japanese. Thus, the fact that Turkish-speaking children did not significantly benefit from contextual information in OVS and SOV sentences in Batman-Ratyosyan and Stromswold’s (2002) study may actually have resulted from the omission of other necessary cues such as specificity or partitivity markers, demonstrative pronouns, or more appropriate contexts, as illustrated above.

As exemplified here, most of the studies on the acquisition of word order variation in verb-final languages are based on offline comprehension studies that require children to listen to a particular utterance and perform an action (e.g., act out the sentence using toys, or pointing to the correct picture). Although capturing general patterns of acquisition, these studies do not provide detailed information regarding the underlying processing mechanisms involved during the interpretation of non-canonical structures.

In a recent study, Özge, Marinis and Zeyrek (2013) used a self-paced listening paradigm to investigate the processing of word order variation in Turkish-speaking children (5–8 years old) and adults. The aim was to investigate whether object-initial sentences would pose any processing challenge in verb-medial orders (i.e., OVS and SVO). In this paradigm, pre-recorded spoken sentences are presented in a phrase-by-phrase fashion and participants are asked to press a button after each phrase to transition to the next phrase. The duration between successive button-presses reveals the listening times devoted to each phrase (e.g., *naughty gorilla*-Nom vs. *naughty gorilla*-Acc), with longer listening times associated with a

greater amount of processing effort (Ferreira, Anes, & Horine, 1996). Listening times devoted to each phrase were compared for simple transitive sentences in SVO and OVS orders, as in (4).

- | (4)    Phrase 1        | Phrase 2  | Phrase 3       | Phrase 4        |
|------------------------|---|----------------|-----------------|
| a. SOV: Haylaz goril   | hızlıca it-ti   | güçlü aslan-ı  | dün sabah.      |
|                        | naughty gorilla-NOM   | fast push-PAST | strong lion-ACC |
|                        | "The naughty gorilla pushed the strong lion yesterday morning." |                |                 |
| b. OVS: Haylaz goril-i | hızlıca it-ti   | güçlü aslan    | dün sabah.      |
|                        | naughty gorilla-ACC   | fast push-PAST | strong lion-NOM |
|                        | "The strong lion pushed the naughty gorilla yesterday morning." |                |                 |

The results showed that during the listening of the first phrase (i.e., NP-Nom in (4a) vs. NP-Acc in (4b)), both children and adults needed longer times to integrate the sentence-initial nominative (unmarked) subject NP compared to the sentence-initial accusative-marked object NP. For the listening time of the second phrase (i.e., verb) participants showed no significant difference between the two structures. During the third phrase (i.e., NP-Acc in (4a) and NP-Nom in (4b)) longer times were needed to integrate the nominative-marked NP (i.e., subject) compared to the accusative-marked NP (i.e., object). Faster processing of sentence-initial accusative object compared to the nominative in the first phrase was attributed to both the frequency of subject-less sentences (Demiral, Schlesewsky, & Bornkessel-Schlesewsky, 2008), and the ambiguity of the nominative (unmarked) case since it could both mark an Agent in a declarative sentence (e.g., *maymun muzu yedi*; *the monkey ate the banana*) or a Patient in a passive sentence (e.g., *maymun mağarada bulundu*; *the monkey was found in the cave*). Slower processing of the subject argument in the third phrase in the OVS order compared to the object in SVO was attributed to the misinterpretation of the OV as a subject-drop sentence. Overall there were no significant differences between the structures in terms of the total listening times. This study demonstrated that offline findings that are based on end-sentence measures might not fully reflect the underlying processes involved in word order processing and that children acquiring head-final languages, just like adults, incrementally integrate local morphosyntactic cues based on their frequency and reliability.

In sum, experimental studies following the *language-as-product* tradition have not always been successful in depicting the full story on how word order is processed on the fly. Offline studies that gather end-sentence behavior from participants may not yield correct generalizations about what cues are used at what point of the comprehension. Although insightful in terms of moment-by-moment processes involved,

real-time studies still lack crucial information as to how pragmatic and information structural aspects are incorporated during the interpretation of argument structure in free word order languages. In what follows, we consider studies that emphasize the integration of pragmatic cues in more naturalistic contexts.

### *Word order flexibility as an indicator of information structure in discourse*

Word order variation in authentic human communication is used across stretches of discourse as a means of managing the flow of information during interactive conversations and extended discourse, such as narratives. Lambrecht (1994) defines information structure as:

(The) component of sentence grammar in which propositions as conceptual representations of states of affairs are paired with lexico-grammatical structures in accordance with the mental states of interlocutors who use and interpret these structures as units of information in given discourse contexts. (Lambrecht, 1994: 5)

In languages with a relatively free word order, the position and ordering of sentence constituents (i.e., word order) is a major marker of information structure, along with other devices such as prosody, case-marking, or choice of construction type. Alteration of the typical word order is determined by discourse-pragmatic factors such as what the addressee has been previously attending to or what has been mentioned before in the ongoing discourse context, often interacting with prosodic cues such as sentential stress. There is controversy, though, about how early children start to monitor the addressee's attention and previously mentioned discourse topics to employ word ordering strategies, such as positioning *given* elements before *new* elements or saying prominent elements before less prominent or backgrounded ones (Gundel, 1988).

Although the effect of information structure on the use of referring terms has been studied (reviewed by Graf & Davies, in press; Küntay, Nakamura, & Ateş Şen, 2014), there is little work on whether or how awareness of information structure is reflected in children's comprehension and production of word order. Most of the existing evidence comes from elicited picture-based narratives, monologically delivered by speakers of different ages, indicating that use of word order as an information-structuring device appears late. Hickmann, Hendriks, Roland, and Liang (1996) find that Mandarin Chinese learners as old as seven years of age do not consistently use the obligatory newness marking through a word order device that involves postverbal placement of just-introduced referents. Word order is also not used by children as efficiently as by adults in Finnish, a language where both postverbal placement of new nouns and NP-level indefiniteness markers are optional devices for indicating information structure (Dasinger, 1995). These studies

indicate that either as an obligatory or an optional device, variation of basic word order for indicating information structure is not as accessible to young storytellers as local article-like devices that function purely at the level of the noun phrase. The planning of sentence positioning of nouns calls for a more sophisticated global discourse processing strategy than just localized editing of information status of referents through article-like elements (Hickmann et al., 1996). Besides, word orders that diverge from the one that is typical in a given language might, in fact, be specialized constructions that are not yet as accessible or practiced enough to be called on by young speakers in structured narrative tasks.

Observations of naturalistic conversations paint a different picture. For example, Russian children around age two have been reported to consider semantic or pragmatic constraints on object placement, consistently locating specific objects preverbally (Avrutin & Brun, 2001) and locating topical objects sentence initially (Dyakonova, 2004). Ukrainian two-year-olds also refrain from word order changes in indefinite-nonspecific-nonpartitive contexts while altering word order in specific-partitive contexts, reflecting the discourse-semantic features of their language (Mykhaylyk, 2009). The same is true for three-year-old Serbo-Croatian children (Ilic & Deen, 2004).

Evidence in young Turkish-speaking children's corpus data is consistent with the patterns observed in these languages. Ekmekçi (1986) showed that children effectively alter word order in their conversations with regard to information structure. As young as 1;7 years of age, children reported consistently located modifiers to the immediate left of the modified item (e.g., *soğuk su*; *cold water*). They were able to use word order appropriately to make a distinction between the predicative (e.g., *su soğuk*; *water is cold*) and attributive function (e.g., *soğuk su*; *cold water*) of adjectives in line with their communicative aims. In addition, unmarked, indefinite and non-referential direct objects were appropriately placed to the immediate left of the verb (e.g., *su getir*; *water bring*), and direct objects were located in the postverbal position only when the NP was overtly marked for accusative case. Earlier data from Ekmekçi (1979) showed that children appropriately locate the constituents containing new discourse information sentence initially. Similarly, Küntay (2002) demonstrated that children use unmarked object arguments in the preverbal position only when they wanted to introduce an indefinite object NP into their discourse, following similar discourse-pragmatic principles they are exposed to in adult speech. These studies showed that young preschool children acquiring Turkish use word order flexibly in their conversations, reflecting considerable knowledge regarding encoding of information structure in their language.

Sensitivity to information structure is manifested not only in using specialized non-basic word order constructions, but also in knowing when to omit or mention

referents in relation to what was mentioned in prior discourse. Indeed, the patterns of lexical addition and deletion effectively change the order of linguistic elements in an utterance. There is some convincing experimental evidence that, in more interactive situations, children show sensitivity to prior discourse in constructing the argument structure of upcoming utterances (Graf & Davies, *in press*). Children as young as 2;5 are able to determine the referent of a backward-referring pronoun by selecting the character that was mentioned first and in subject position in prior discourse (Song & Fisher, 2007). German-speaking 2-year-olds are more likely to use pronouns and null forms to answer questions such as *where's the doll?*, showing sensitivity to the notion that the full noun had already been introduced in the question of the previous speaker (Wittek & Tomasello, 2005). Clark and Bernicot (2008) found that 2-year-olds repeat newly introduced topics from the previous utterance to indicate that it is now mutually shared information. In Turkish, children drop arguments that have been previously established as discourse topics (Demir, So, Özyürek & Goldin-Meadow, 2011), and they place first-person pronouns after the verb to make object contrast (e.g., *cay iç-er-im ben* 'tea drink-Aor-1sg I' 'I would drink TEA') or to make an assertive announcement of action in the face of opposition (*bu aç-a-yim ben* 'this open-Opt-1sg I' 'Let me open THIS') (Slobin & Talay, 1986).

Given the nature of child-directed input, it is not surprising that Turkish preschoolers use word order flexibly to indicate different nuances of information structure. This has been observed in a corpus of naturalistic child-caregiver interaction in Turkish (the KULLDD corpus, Ural, Yüret, Ketrez, Koçbaş, & Küntay, 2009) showing that both caregivers and children produce pragmatically appropriate utterances in terms of word order variation (Küntay & Slobin, 1996; 2001; 2002). These studies show that alteration of word order across successive utterances is a commonly encountered discourse pattern used by Turkish caregivers interacting with their young children, as demonstrated in speech addressed to children at 12 to 18 months of age (Ateş Şen, Demir, & Küntay, *in preparation*). In what we called *variation sets*, caregivers repeat the same communicative content but alter features of the linguistic construction across successive utterances. These reformulations serve different discourse-pragmatic goals, displaying a general intent of increasing the likelihood of comprehension and compliance on the part of the child. Many of these reformulations of the same communicative intention across successive utterances in Turkish involve changing the position of sentence elements. Analysis reveals that as high as 40% of the variation sets that retain the same lexical items involve a change in word order (Küntay & Slobin, 1996), as illustrated in (5).

- (5) Ver el-ler-in-i  
 Give hand-PL-Poss2SG-ACC  
 “Give me your hands.”
- El-ler-in-i ver-ir-mi-sin?  
 hand-PL-Poss2SG-ACC give-Aor-QuestPart-2SG  
 “Could you give me your hands?”
- El-ler-in-i ver.  
 hand-Pl-Poss2SG-ACC give  
 “Give me your hands.

Across successive utterances in (5), the noun *ellerini* retains the same morphology, while the verb goes from a simple imperative to an indirect imperative, and back to a simple imperative, in addition to changing position. Verbs were found to be repeated across more than one utterance in the vast majority of variation sets (about 80%) – while they change their morphological form in nearly half (43%) of the cases where they are repeated from one utterance to the next. In other words, although Turkish is commonly referred to as a verb-final language in typological frameworks, the position of the verb changes frequently in conversational discourse addressed to children. Even in sequences in which word order does not change, patterns of addition and deletion of lexical items often function to change the utterance position of a verb. Naturalistic corpus studies complement the aforementioned online comprehension findings by Özge et al. (2013), showing fast integration of the verb in OV and SV structures regardless of the argument type preceding the verb.

Caregiver-child interactions are full of examples of reordering of sentence elements in successive utterances of topically continuous discourse from young ages on. In the following example (6), the mother is trying to get Azra, a 15-month old girl, to tell her dad about the outing they did together earlier. The mother in this case is prompting a retelling of a shared event for a third member of the family who has not participated in the event. To make sure that her directive gets through, the order of main clause (*tell your dad*) and the subordinate clauses (*how we went*) displayed in the first utterance is reversed in the second utterance.

- (6) Mother: Anlat-sana baba-ya nasil git-ti-k?  
 Tell-Opt dad-Dat how go-Past-1PL  
 “Tell your dad how we went.”
- Mother: Nasil gi-ti-k, baba-ya anlat-sana tatlı-m.  
 How go-Past-1PL dad-Dat tell-Opt sweet-Poss.1SG  
 “How did we go? Tell your dad honey.”

Example (6) from a Turkish-speaking mother to her preverbal 12-month old daughter shows positioning of a word in different locations across successive utterances in a flexible word order language. We call these discourse patterns *referential sets*, to describe successive utterances that maintain reference to a given referent across a stretch of discourse (Ateş Şen, Demir, & Küntay, in preparation). In the excerpt in (7), the mother refers to a gift-wrap using the word *paket*, after having unwrapped a gift given by the child's grandmother. Within 14 seconds of the first reference, there are five more mentions of the same referent across extended discourse and interaction about this object: two by means of nouns, the other three involving noun elision.

- (7) a. Mother: (the mother is holding up the gift wrap; the child looks at the gift wrap)
- Sen bu paket-i de sev-er-sin.  
you this wrap-ACC FocPart love-Aor-2SG  
“You like this gift wrap as well.”
- b. Mother: (the mother is folding the gift wrap)
- Şimdi kaldır-ma-ya-lım paket-i.  
now hold-Neg-Opt1PL wrap-ACC  
“Let’s not put this gift wrap away now.”
- c. Mother: (the mother is holding up the gift wrap)
- Paket-in kendisi de çok önemli di mi?  
wrap-GEN RefPron FocPart very important isn’t it?  
“The gift wrap itself is very important, isn’t it?”
- d. Mother: (the mother is still holding up the wrap, child is looking at and touching the wrap)
- Bakar mı-sın?  
look QuestPart-2SG  
“Could (you) look at this?” (argument drop)
- e. Mother:
- Çok süslü  
very smart-looking  
“Very smart-looking.” (argument drop)
- f. Mother:
- Sen-in gibi. Sen de süslü-sün  
You-GEN like you FocPart smart-looking-2SG  
“Like you, you are smart-looking as well.” (argument drop)

In this referential set, *paket*, the word for gift-wrap, is placed in the focal preverbal position in the first utterance, then moved to the postverbal slot reserved for previously attended referents once the child's attention is secured. In the third utterance, the referent moves back to utterance initial, preverbal position, only to be represented by a null form in the fourth utterance, when the child is both looking at and touching the object in question. The timing of word ordering appears contingent on the securing of attention of the child on the part of the mother. These kinds of discourse stretches where referential intentions are nuanced across consecutive utterances are very common in Turkish child-directed speech. Such referential sets play a major role in introducing new words such as *paket gift-wrap*, displaying language-specific morphosyntactic properties such as case-marking (*paket-i*; *wrap-Acc* vs. *paket-in*; *wrap-Gen* in utterances (7a), (7b) and (7c)), argument omission (in utterance (7d), (7e), and (7f)) in addition to pragmatically motivated word order variation. Word ordering appears to be governed by interactive cues during this exchange, as demonstrated by the repositioning of the discourse-old noun after the verb or dropping of the noun once the referent is placed in the common ground between the mother and the child (as evidenced by the gaze direction and touch of the child). Thus, word ordering serves as an interactional device that manages the status of continuously changing given and new information, updating the common ground of the interlocutors regarding the referent.

The examples of re-ordering in (5) to (7) are common in Turkish child-directed speech in multi-utterance requestive discourse that calls for action or verbalization from the child. These examples demonstrate that, when the children are so young that their verbal answers to questions are not yet anticipated, caregivers often restate their questions and prompts, typically together with reordering. When children start to make explicit contributions that are temporally and semantically contingent on previous turns (Casillas, 2014), reordering becomes more sensitive to the specific content of the turn supplied by the child.

Studies of how general characteristics of child-directed speech affect children's linguistic complexity are abundant (Snow, 1995), but what is lacking is extensive turn-by-turn analysis focusing on children's contributions to the local dialogic process. We do not know, for example, how young children start to use variation of word order across successive utterances as aptly as their caregivers to contribute to the flow of information in conversations. In a recent study modeled after Kamide, Scheepers & Altmann (2003), we have some evidence that children vary word order and use contrastive focus appropriately when correcting an experimenter's wrong descriptions of a pictured event (Özge, Küntay, & Snedeker, in progress). Four-year-olds were presented with a visual context comprised of three objects on the screen of an eye-tracker (e.g., a rabbit, a fox, and a carrot), at the same time listening to a pre-recorded utterance about this context (e.g., *The rabbit*

*is going to eat that carrot vs. The fox is going to eat that rabbit)* and their eye movements on the visual scene were recorded in relation to the different parts of the utterance they heard. Immediately after this scene-tracking phase, the children were presented with an animation serving as an offline comprehension task and asked to report whether the animation correctly depicted the event in the utterance they had heard. During this comprehension phase of the study, children produced utterances where they flexibly manipulated the word order and prosody in line with the requirements of the information structure, as illustrated in (8).

- (8) Visual context presents the following three referents: a rabbit, a fox, and a carrot. Utterance accompanying the visual context:

Tavşan birazdan yi-yecek şurada-ki havuc-u.  
rabbit shortly eat-Fut that-Rel carrot-ACC  
“The rabbit will shortly eat the carrot over there”

Animation following the utterance depicts the following: The fox is eating the rabbit. Child’s response judging whether the animation depicts the utterance correctly:

Yanlış! TİLKİ yedi tavşan-ı.  
Wrong! FOX-Nom ate rabbit-ACC  
“Wrong! FOX ate the rabbit.”

Note that the animation here does not depict the utterance correctly, since the utterance created an expectation that the rabbit would eat the carrot, whereas in the animation it was the fox, not the rabbit, who performed the act of eating and the rabbit was actually the Patient. In providing a reason why the animation is wrong, the child locates the new Topic *the fox* sentence-initially, to the immediate left of the verb, backgrounding the old Topic *the rabbit* to post-verbal position. In addition, the new Topic receives a contrastive focus because the child corrects what was said in the spoken utterance in relation to what was depicted in the animation, saying that the expectation that the rabbit would eat something was wrong, since in fact the fox was eating the rabbit. That is, with this reordering of constituents and contrastive prosody on the first constituent, the child is not only varying the word order flexibly but also constructing an emphatic denial fully in line with the information structural elements presented in discourse. Such naturalistic displays of word order variation could also be studied with younger children producing unscripted interactive dialogues with caregivers. This requires a rigorous system of tracking moment-to-moment changes that might be associated with word order alternations in discourse, for example, changes in the attentional states, actions, and real-time comprehension of interlocutors. We predict that children would be more apt to use word ordering in these interactionally and multimodally supported conversations than in their elicited monologic narratives.

## Final remarks

Although word order is certainly a structural cue to participant roles in sentences, it is simultaneously used as a tool to package information in interactional contexts. As children are learning to derive *who did what to whom* from the ordering cues they hear in sentences, they also need to acquire knowledge about word order as an indication of information structure. How early do children develop awareness of the two functions of word order? We discussed a set of cross-linguistic studies showing that children's sensitivity to word order as a structural cue to participant roles appears as early as three years of age. We also showed that using word order to structure information in monologic narratives does not appear until relatively late, although children manage interactional flow of information within a turn-taking structure from early on. It could be that elicited narratives strip away the motivation found in dialogic interaction, where children can use word ordering to re-instate a point, refuse an offer, or correct a statement in relation to what was said by an interlocutor. Studies of children's contributions in dialogically constructed narratives are needed to determine whether interactive contexts lead to earlier deployment of their skills in using word order as an indication of information structure.

Could sensitivity to word order be acquired devoid of its pragmatic functions? Data coming from research with newborns indicate that infants are ready to perceive the sequential position of linguistic units in addition to their relation to one another in artificially constructed stimuli (Gervain, Berent, Werker, 2012). However, having the perceptual apparatus to detect sequencing relations in speech might not be adequate to comprehend the pragmatic motivations for changing word order in actual discourse. In real languages, change in the order of linguistic units serves interactional information management functions, rendering word order a cue to information structure and pragmatic meaning from the beginning. What is still unknown is how infants develop the ability to pair the statistically associated sequences with information content and how social and pragmatic interpretations that take place at the moment of actual interactions lead to appropriate comprehension and production of word ordering. After all, it is attentional and social cues that lead to a change in sequential position of words in the course of actual, realtime conversational discourse. Thus, the sensitivity to linguistic sequence markers that could be used as structural cues to participant roles in events might fall short in explaining how word ordering is acquired as an interactional process. An ability to pay attention to relative sequence markers in input is probably necessary but not sufficient to learn how patterns of word re-ordering are sensitive to discourse-pragmatic principles in the course of language use in interaction. When eliciting interpretations of word order from children, future work should employ discourse-sensitive stimuli to overcome some of the limitations of

typical experimental designs. In addition, systematic studies of continuous conversations between young children and caregivers should be conducted, taking into account nonverbal attentional and social cues that appear to play a role in the ways word order is altered across successive interactive turns. Recently, with head-mounted mini cameras, children's eye gaze during speech-based interactions with caregivers was tracked, providing a more reliable first-person view of the child, rather than the not-so-reliable third-person coding of nonverbal behavior recorded with tripod-mounted cameras (Yurovsky, Smith, & Yu, 2013). Use of eye movements promises an advance in bridging the *language-as-product* and *language-as-action* traditions for studies of word ordering and other linguistic devices (Tanenhaus & Trueswell, 2005). Even more auspicious would be to be able to seamlessly track other nonlinguistic movements such as actions and gestures, as they are integrated with eye-gaze events and speech during regular interactions. After all, how sensitivity to pragmatic nuances of word order is acquired cannot be realistically studied without paying attention to nonlinguistic action accompanying linguistic interactions.

## References

- Akhtar N., & Tomasello, M. (1997). Young children's productivity with word order and verb morphology. *Developmental Psychology, 33*, 952–965. DOI: 10.1037/0012-1649.33.6.952
- Ateş Şen B., Demir, Ö.D., & Küntay, A.C. (In preparation). Communicative actions and words in referential sets.
- Avrutin, S., & Dina B. (2001). The expression of specificity in a language without determiners: Evidence from child Russian. In A.H.-J. Do, L. Domínguez, & A. Johansen (Eds.), *Proceedings of 25th Boston University Conference on Language Development* (pp. 70–81). Somerville, MA: Cascadilla Press.
- Bates, E., & MacWhinney, B. (1987). Competition, variation, and language learning. In B. MacWhinney (Ed.) *Mechanisms of language acquisition* (pp.157–194), Hillsdale, NJ: Lawrence Erlbaum Associates.
- Batman-Ratyosyan, N., & Stromswold, K. (1999). What Turkish acquisition tells us about underlying word order and scrambling. *Proceedings of the 23rd Annual University of Pennsylvania Linguistics Conference, 6*, 37–52.
- Batman-Ratyosyan, N., & Stromwold, K. (2002). Morphosyntax is easy, discourse pragmatics is hard. In A.H.-J. Do, L. Domínguez, & A. Johansen (Eds.), *Proceedings of 25th Boston University Conference on Language Development* (pp. 793–804). Somerville, MA: Cascadilla Press.
- Candan, A., Küntay, A.C., Yeh, Y., Cheung, H., Wagner, L., & Naigles, L.R. (2012). Language and age effects in children's processing of word order. *Cognitive Development, 27*, 205–221. DOI: 10.1016/j.cogdev.2011.12.001
- Cannizzaro, C.L. (2012). *Early word order and animacy*. Unpublished doctoral dissertation. University of Groningen.

- Casillas, M. (In press). Turn-taking. In D. Matthews (Ed.), *Pragmatic Development in First Language*. Amsterdam: John Benjamins.
- 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. DOI: 10.1515/COGL.2009.015
- Chapman, R.S., & Miller, J.F. (1975). Word order in early two and three word utterances: Does production precede comprehension? *Journal of Speech and Hearing Research*, 18, 346–354.
- Clark, E.V. (2012). An interview with our new President, Eve Clark / Interviewer: Angel Chan. *IASCL Child Language Bulletin*, 32.
- Clark, E.V., & Bernicot, J. (2008). Repetition as ratification: How parents and children place information in common ground. *Journal of Child Language*, 35, 349–372. DOI: 10.1017/S0305000907008537
- Clark, H.H. (1992). *Arenas of language use*. Chicago, IL: Chicago University Press.
- Clark, H.H. (1996). *Using language*. Cambridge: CUP. DOI: 10.1017/CBO9780511620539
- Dasinger, L. (1995). *The development of discourse competence in Finnish children: The expression of definiteness*. Unpublished doctoral dissertation. University of California, Berkeley.
- Demir, Ö.E., So, W., Özürek, A., & Goldin-Meadow, S. (2011). Turkish- and English-speaking children display sensitivity to perceptual context in the referring expressions they produce in speech and gesture. *Language and Cognitive Processes*, 27, 844–867. DOI: 10.1080/01690965.2011.589273
- Demiral, S.B., Schlesewsky, M., & Bornkessel-Schlesewsky I. (2008). On the universality of language comprehension strategies: Evidence from Turkish. *Cognition*, 106, 484–500. DOI: 10.1016/j.cognition.2007.01.008
- Dyakonova, M. (2004). Information structure development: Evidence from acquisition of word order in Russian and English. *Nordlyd: Tromsø Working Papers*, 32, 88–109.
- Ekmekçi, F. (1979). *The Effects of Simulation / Games on Foreign Language Learning*. Unpublished MA thesis. The University of Texas at Austin.
- Ekmekçi, F. (1986). Significance of word order in the acquisition of Turkish. In D.I. Slobin & K. Zimmer (Eds.), *Studies in Turkish linguistics* (pp. 253–264). Amsterdam: John Benjamins.
- Ferreira, F., Anes, M.D., & Horine, M.D. (1996). Exploring the use of prosody during language comprehension using the auditory moving window technique. *Journal of Psycholinguistic Research*, 25, 273–290. DOI: 10.1007/BF01708574
- Fisher, C., & Song, H. (2006). Who's the subject? Sentence structures as analogs of verb meaning. In K. Hirsh-Pasek & Golinkoff, R.M. (Eds.), *Action meets word: How children learn the meanings of verbs* (pp. 392–428). Oxford: OUP. DOI: 10.1093/acprof:oso/9780195170009.0.03.0016
- Fodor, J.A. (1983). *Modularity of mind*. Cambridge, MA: The MIT Press.
- 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. DOI: 10.1111/j.1467-9280.2006.01767.x
- Gervain, J., Berent, I., & Werker, J.F. (2012). Binding at birth: The newborn brain detects identity relations and sequential position in speech. *Journal of Cognitive Neuroscience*, 24, 564–574. DOI: 10.1162/jocn\_a\_00157
- Golinkoff, R.M., Hirsh-Pasek, K., Cauley, K.M., & Gordon, L. (1987). The eyes have it: Lexical and syntactic comprehension in a new paradigm. *Journal of Child Language*, 14, 23–45. DOI: 10.1017/S030500090001271X

- Graf, E., & Davies, C. (In press). The development of reference. In D. Matthews (Ed.), *Pragmatic Development in First Language*. Amsterdam: John Benjamins.
- Gundel, J.K. (1988). Universals of topic-comment structure. In M. Hammond, E.A. Moravcsik, & J.R. Wirth (Eds.), *Studies in syntactic typology* (pp. 209–239). Amsterdam: John Benjamins.
- Hakuta, K. (1982). Interaction between particles and word order in the comprehension and production of simple sentences in Japanese children. *Developmental Psychology, 18*, 62–76. DOI: 10.1037/0012-1649.18.1.62
- Hickmann, M., Hendriks, H., Roland, F., & Liang, J. (1996). The marking of new information in children's narratives: A comparison of English, French, German and Mandarin Chinese. *Journal of Child Language, 23*, 591–619. DOI: 10.1017/S0305000900008965
- Ilic, T., & Deen, K.U. (2004). Object raising and cliticization in Serbo-Croatian child language. In the A. van Kampen & S. Baauw (Eds.), *Proceedings of GALA 2003* (pp. 235–243). Utrecht: LOT.
- Kamide, Y., Scheepers, C., & Altmann, G. (2003). Integration of syntactic and semantic information in predictive processing: Cross-linguistic evidence from German and English. *Journal of Psycholinguistic Research, 32*, 37–55. DOI: 10.1023/A:1021933015362
- Küntay, A.C. (2002). Development of the expression of indefiniteness: Presenting new referents in Turkish picture-series stories. *Discourse Processes, 33*, 77–101. DOI: 10.1207/S15326950DP3301\_04
- Küntay, A., Nakamura, K., & Ateş Şen, B. (2014). Crosslinguistic and crosscultural approaches to pragmatic development. In D. Matthews (Ed.) *Pragmatic Development in First Language*. Amsterdam: John Benjamins.
- Küntay, A., & Slobin, D.I. (1996). Listening to a Turkish mother: Some puzzles for acquisition. In D. Slobin, A. Gerhardt, A. Kyrtzis, & J. Guo (Eds.) *Social interaction, social context, and language: Essays in honor of Susan Ervin-Tripp* (pp. 265–287). Mahwah, NJ: Lawrence Erlbaum Associate.
- Küntay, A.C., & Slobin, D.I. (2001). Discourse behavior of lexical categories in Turkish child-directed speech: Nouns vs. verbs. In M. Almgren, A. Barreña, M. Ezeizabarrena, I. Idiazaabal, & B. MacWhinney (Eds.), *Research on child language acquisition: Proceedings for the 8th Conference of the International Association for the Study of Child Language* (pp. 928–946). Somerville, MA: Cascadilla Press.
- Küntay, A.C., & Slobin, D.I. (2002). Putting interaction back into child language: Examples from Turkish. *Psychology of Language and Communication, 6*, 5–14.
- Lambrecht, K. (1994). *Information structure and sentence form: Topic, focus, and the mental representation of discourse referents*. Cambridge: CUP. DOI: 10.1017/CBO9780511620607
- Mykhaylyk, R. (2009). Developmental patterns in flexible word order acquisition. In J. Crawford, K. Otaki, & M. Takahashi (Eds.), *Proceedings of the 3rd Conference on Generative Approaches to Language Acquisition North America (GALANA 2008)* (pp. 165–174). Somerville, MA: Cascadilla Proceedings Project.
- Naigles, L.R., & Swensen, L.D. (2006). Syntactic supports to word learning. In M. Shatz & E. Hoff (Eds.), *Blackwell handbook of language development* (pp. 212–231). Malden, MA: Wiley-Blackwell.
- Otsu, Y. 1994. Early acquisition of scrambling in Japanese. In T. Hoekstra & B.D. Schwartz (Eds.), *Language acquisition studies in generative grammar* (pp. 253–264). Amsterdam: John Benjamins.
- Özge, D., Küntay, A.C. & Snedeker, J. (In preparation). Predictive processing of morphosyntax in Turkish-speaking children.

- Özge D., Marinis, T., & Zeyrek D. (2013). Object-first orders in Turkish do not pose a challenge during processing. In U. Özge (Ed.), *Proceedings of the 8th Workshop on Altaic Formal Languages* (pp. 269–280). Cambridge, MA: MITWPL.
- Sano, T. (2004). Scope relations of QP's and scrambling in the acquisition of Japanese. In A. van Kampen & S. Baauw (Eds.), *The Proceedings of GALA 2003* (pp. 421–431). Utrecht: LOT.
- Slobin, D.I., & Bever, T.G. (1982). Children use canonical sentence schemas: A crosslinguistic study of word order and inflections. *Cognition*, 12, 229–265. DOI: 10.1016/0010-0277(82)90033-6
- Slobin, D.I., & Talay, A. (1986). Development of pragmatic uses of subject pronouns in Turkish child language. In A.A. Aksu-Koç & E. Erguvanlı Taylan (Eds.), *Proceedings of the Turkish Linguistics Conference* (pp. 207–228). Istanbul: Boğaziçi University Press.
- Snow, C.E. (1995). Issues in the study of input: Fine-tuning, universality, individual and developmental differences, and necessary causes. In P. Fletcher & B. MacWhinney (Eds.), *Handbook of child language* (pp. 180–193). Oxford: Blackwell.
- Song, H., & Fisher, C. (2007). Discourse prominence effects on 2.5-year-old children's interpretation of pronouns. *Lingua*, 117, 1959–1987. DOI: 10.1016/j.lingua.2006.11.011
- Tannenhaus, M.K., & Trueswell, J.C. (2005). Eye movements as a tool for bridging the language-as-product and language-as-action traditions. In J.C. Trueswell & M.K. Tanenhaus (Eds.), *Approaches to studying world-situated language use: Bridging the language-as-product and language-as-action traditions* (pp. 3–37). Cambridge, MA: The MIT Press.
- Trueswell J. C., & Tanenhaus, M.K. (2005). *Approaches to studying world-situated language use: Bridging the language-as-product and language-as-action traditions*. Cambridge, MA: The MIT Press.
- Ural, A.E., Yüret, D., Ketrez, F.N., Koçbaş, D., & Küntay, A. (2009). Morphological cues vs. number of nominals in learning verb types in Turkish: The syntactic bootstrapping mechanism revisited. *Language and Cognitive Processes*, 24, 1393–1405. DOI: 10.1080/01690960902775525
- Yuan, S., Fisher, C., & Snedeker, J. (2012). Counting the nouns: Simple structural cues to verb meaning. *Child Development*, 83, 1382–1399. DOI: 10.1111/j.1467-8624.2012.01783.x
- Yurovsky, D., Smith, L.B., & Yu, C. (2013). Statistical word learning at scale: The baby's view is better. *Developmental Science*, 16(6), 959–966. DOI: 10.1111/desc.12011
- Wittek, A., & Tomasello, M. (2005). Young children's sensitivity to listener knowledge and perceptual context in choosing referring expressions. *Applied Psycholinguistics*, 26, 541–558. DOI: 10.1017/S0142716405050290



# **The discourse basis of the Korean copula construction in acquisition**

Patricia M. Clancy  
University of California, Santa Barbara

This chapter examines the discourse foundations of the Korean copula construction as it is acquired by two young children in interaction with their mothers. The findings suggest that the construction develops through gradual generalization of an initially limited set of context-specific structural schemas, and that it is profoundly sensitive to discourse factors throughout the course of development. From the outset, the realization and form of the construction's components are shaped by the functions it serves (e.g., labeling vs. describing), the discourse-pragmatic properties of referents (e.g., given vs. new information), the interactional roles of participants (e.g., questioner vs. respondent), and the ideologies (e.g., child as conversational partner, mother as tutor and playmate) that motivate particular types of caregiver-child activities and talk.

## **Preface**

Inspired by Eve's monumental body of work on lexical acquisition and her focus on the role of interaction, I attempt in this chapter to explore the ways in which discourse provides the basis for the forms and functions of the copula construction, the primary vehicle for introducing vocabulary. Eve's research on new words in interaction – how caregivers introduce them and children respond – provides an especially rich foundation for this endeavor.

## **Introduction**

### *A discourse-based approach to constructions*

Research on the acquisition of constructions during the past 15 years has given new form to the long-standing debate between formal, generativist approaches to the acquisition of grammar and functional, usage-based approaches. Like earlier

functionalist approaches (e.g., Bates & MacWhinney, 1979), the definition of constructions as form-meaning pairs (Goldberg, 1995) has focused attention on the form-function relations in caregiver speech as the basis for children's grammars. Longitudinal research showing gradual, piecemeal acquisition of grammatical constructions (e.g., Lieven, Pine, & Baldwin, 1997; Tomasello, 2000) has challenged the predictions of formal theory for rapid, across-the-board generalization of grammatical forms based on innate linguistic knowledge. The accumulated evidence favors a view of the child as gradually building up mental representations of grammatical constructions based on experience with adult grammar, i.e., an acquisition theory that is functional and constructionist rather than formal and nativist.

To date, research on the acquisition of constructions has focused primarily on the sentence level, with little attention to the discourse foundations of constructions. A discourse-functional approach, in contrast, is grounded in the assumption that there is an inextricable relationship between discourse and grammar. Since actual uses of grammar never occur apart from specific discourse contexts, children always experience grammar and discourse simultaneously. Discourse context, therefore, is not merely an envelope from which grammatical constructions can be extracted and analyzed independently; rather, the form of constructions is sensitive to the discourse motivations for their use.

A discourse-functional approach highlights the interactive, culturally situated nature of grammatical constructions. In everyday discourse, the form of grammatical constructions is regularly associated with activities that the child either observes or participates in directly. Since grammatical constructions tend to be associated with particular activities, participating in these activities and related talk affords a discourse-based incentive for children to master grammatical constructions, as well as a pathway to achieving that mastery. Children are also sensitive to the sociocultural context of grammatical constructions, including the social identities of interlocutors and their roles in the interaction (Ochs & Schieffelin, 1995: 73-75). This sensitivity is especially relevant for constructions that figure prominently in cultural routines with clear interactional roles for participants, such as labeling routines with the copula construction.

In this study, I take a discourse-functional approach to the Korean copula construction in caregiver-child interaction, focusing on the ways in which the form of the construction arises from the ongoing activities and accompanying conversation, as well as on children's sensitivity to the discourse motivations of grammatical forms during the process of acquisition.

### *Prior research on English*

Eve Clark's research on the acquisition of new words in English paves the way for a discourse-based approach to the copula construction, which is the primary

syntactic vehicle for introducing, negotiating, and displaying new vocabulary in caregiver-child interaction. Three areas of Clark's research are particularly relevant to the present study: (1) the discourse-pragmatic factors of physical co-presence and linguistic prior mention, (2) the limited set of syntactic frames used in presenting new words, and (3) the sequential organization of caregiver-child talk about objects and their properties.

In the caregiver-child interactions in which new words are introduced, the primary activities are looking at picture books and joint play with blocks (Clark & Wong, 2002). The physical co-presence of the objects to be labeled or described allows participants to indicate the intended referent non-verbally, e.g., by pointing, touching, and tapping to identify the object and its parts and properties (Clark & Estigarribia, 2011). Linguistic co-presence – prior mention in the discourse – links the information in one utterance, such as an object label, to the information in subsequent utterances, such as definitions, set membership, whole/part relations, and properties of objects including color, number, and function (Clark, 1997, Clark, 2006; Clark & Grossman, 1998).

As Clark's research has shown, the organization of talk about objects and their properties takes place in a consistent sequence. The caregiver first establishes joint attention with the child through eye gaze, gesture, or a verbal summons (e.g., *Look!*), and produces a label only after the child gazes at the intended referent (Estigarribia & Clark, 2007). Once a new word is introduced, the child usually responds by repeating it, producing an agreement token (e.g., *yeah*), or making further comments about the referent (Clark, 2007; Clark & Bernicot, 2008). If the child labels an object incorrectly, the caregiver's next turn is likely to be an explicit correction, a clarification question, or simply an utterance using the correct term (Chouinard & Clark, 2003; Clark & Wong, 2002). When introducing new words, caregivers use a limited set of syntactic frames, e.g., *That's \_\_\_, This is \_\_\_, What's this/that?* (Clark & Wong, 2002); these simple structures reflect the discourse context in which new words are introduced: seeking and providing labels for co-present objects.

As this brief overview indicates, Clark's research provides ample evidence for the claim that "children discover word meanings in the course of interaction" (Clark, 2010: 267). In some cultures, the relevant interactions may not involve the kind of labeling routines analyzed in Clark's work. However, in cultures where labeling is a common practice with young children, such routines serve as a key site for acquisition of the copula construction.

Recent research on the English copula construction has provided evidence for a gradual progression from rote-learned chunks to a more general, abstract construction. Wilson (2003), for example, has found that children begin producing the copula in item-specific pronoun + copula frames, e.g., *It's \_\_\_ and That's \_\_\_,*

before generalizing to a more abstract construction with an analyzed copula, and to lexical as well as pronominal subjects: NP<sub>SUBJ</sub> COP. Wilson's results have been replicated and expanded by Pine et al. (2008), who found a significantly higher rate of pronominal vs. lexical subjects with the copula, as well as higher frequencies of certain pronouns (*it, he*) than others (*I*). Overall, the English findings support a constructionist account, in which early lexically specific schemas are gradually generalized.

The discourse bases for the initial limited forms of the English copula construction have not been explicitly addressed. We can speculate, however, that their form reflects the properties of the discourse contexts in which co-present or depicted inanimate (*it*) and animate (*he*) referents are labeled. The referents are presumably encoded with pronouns because they are given information, i.e., information that can be assumed to be currently activated in the mind of the listener (Chafe, 1976), on the basis of their physical co-presence and/or recent mention in the discourse.

### *The Korean copula construction*

The Korean copula construction differs in certain respects from its English counterpart. Since it accommodates neither adjectives (which are inflected like verbs) nor locative complements (which take the verb *issta* 'exist' rather than the copula), the Korean copula construction always has an NP complement. As schematized in (1), the full form of the construction includes four grammatical components: the subject, morphological marking on the subject, the predicate nominal, and the copula.

- (1) NP<sub>SUBJ</sub>-MARKER NP<sub>PRED</sub>-COPULA

Morphological markers attach phonologically to the preceding noun (Kim, 1997), and the copula forms a lexical word with the head noun of its complement (Oh 1991, cited in Kim, Sells, & Wescoat 2008). Depending on the discourse context, any of the four components of the copula construction may be left unexpressed; there is thus a direct relation between the form of the construction and the discourse factors motivating realization of its parts.

### *Research questions*

Prior research on the Korean transitive construction (Clancy, 2004) has shown that it has strong discourse-functional correlates. In this study, the discourse foundations of the Korean copula construction will be investigated, focusing on the following four questions: (1) How does discourse function impact the form of the

copula construction – specifically, the realization of its four component parts? (2) How do the grammatical forms and discourse functions of the construction change over time? (3) Is there evidence for a transition from limited structural schemas to more abstract, general patterns in development, as documented for English? and (4) What ideologies and related activities provide the cultural matrix for use and acquisition of the Korean copula construction? Discourse function is taken to include the discourse-pragmatic properties of referents (e.g., their status as given vs. new information), the interactional roles of participants and associated speech acts (e.g., questioner/question vs. respondent/answer), and the function of the construction with respect to its semantic content (e.g., labeling vs. describing the subject referent). Such properties constitute an important part of the discourse foundations of the copula construction, potentially impacting the course of its acquisition.

## Methodology

The data for this study are taken from 90-minute audio-recordings of two Korean girls, Wenceng and Hyenswu, and their interlocutors. The children, who were 1;8 and 1;10 years old, respectively, at the beginning of the study, were recorded by one or two Korean graduate student research assistants in bimonthly sessions for just over a year. During the recording sessions, each child interacted with her mother, the research assistants, and occasionally other family members, as they played with toys, assembled puzzles and Lego structures, drew pictures, read from storybooks, looked at photo albums, ate snacks, and talked about their ongoing activities.

The children's fathers were graduate students at a private university on the east coast, where they were living with their families until completion of their doctoral degrees. The children spoke Korean with their parents, but acquired some passive comprehension of English during the year. Wenceng was an only child, but Hyenswu was exposed to English by her five-year-old sister, who was attending kindergarten; by the end of the year she could produce a number of fixed phrases (e.g., *no, don't touch*). As temporary residents in the United States, the families maintained a Korean lifestyle culturally: speaking Korean at home, eating Korean food, and socializing with Korean friends.

For this study, transcripts of one session per month for the children have been analyzed, as well as 350 utterances from each of the mothers, comprising their first 150 utterances from following periods: the early months of the study (1st–4th months), the mid (5th–9th) months, and the later (10th–13th) months. All affirmative, unembedded copula constructions in these data are included in the following analyses.

**Table 1.** Participants and copula constructions produced.

Participant	Age	Number of Copula Constructions	Percentage of Copulas among Verb Tokens
Hyenswu (H)	1;10 – 2;10	470	14%
H's mother	Adult	228	19%
Wenceng (W)	1;8 – 2;8	839	22%
W's mother	Adult	186	15%

Table 1 summarizes the participants and data for this study, including the number and proportional frequency of copula constructions out of all verb tokens. An overt predicate nominal was treated as criterial for identifying instances of the copula construction, with two exceptions: (1) questions consisting solely of the subject, e.g., *i-ke-nun?* this-thing-TOP ‘What about this one?’ when the preceding discourse had established the copula as the intended verb, and (2) clarification questions such as *i-ke?* this-thing ‘(Do you mean) this one?’ following questions with the copula.

The copula is the most frequent verb for the children and their mothers. The copula construction is proportionally more frequent in the early months, decreasing by the later months of the study to 7% of verb tokens for the mothers, 10% for Hyenswu, and 12% for Wenceng.

## Results

### *Developmental changes in form*

The children’s copula constructions show gradual development from a few structurally limited schemas to the full four-component form of the construction. Table 2 presents the sequence in which the most common forms of the construction developed. Each form is treated as having “emerged” at the age of its first use, and as “established” when the child uses it appropriately at least five times in a single session. (The term “established” is used rather than “acquired,” since certain forms may already have been acquired in other constructions.) When a component meets the “established” criterion during the first session in which it appears, an age of emergence is not given. (PN = predicate nominal, COP = copula, SUBJ = subject, MRK = morphological marker)

With production data from just one session per month, Table 2 can give only a rough measure of development; the children most likely could understand more than they could produce. Furthermore, since Wenceng was already using different

**Table 2.** Sequence of development of the copula construction.

Form of construction	Hyenswu		Wenceng	
	Emerged	Established	Emerged	Established
PN		1;10		1;8
PN + COP	1;10	2;0		1;8
SUBJ + PN	1;11	2;0		1;8
SUBJ + PN + COP	2;1	2;2	1;8	1;10
SUBJ + MRK + PN	2;2	2;5	1;8	2;6
SUBJ + MRK + PN + COP	2;2	2;5	1;8	2;3

forms of the construction from the outset, we must rely primarily on Hyenswu's data to extrapolate an approximate sequence of acquisition. The earliest form of the construction is the predicate nominal alone, usually single-word labels responding to *mwe-ya* what-COP.IE 'What is (it)?' questions. These simple labels comprise 81% of Hyenswu's copula constructions in the first month of the study and 76% in the second month.

The earliest grammatical complexity appears at 2;0 years of age, when Hyenswu begins attaching the inflection *-ya* to labels in the schema N-*ya* COP.IE,<sup>1</sup> e.g., *baby-ya* babydoll-COP.IE '(It)'s a babydoll'. In the same month, the demonstrative pronoun *i-ke* this-thing 'this' appears in subject position, creating the frame: *ike* N 'This (is a) N'. Both subject<sup>2</sup> and copula do not appear in the same utterance, suggesting a limit on the complexity of the construction at this point. The predicate nominal is sometimes elaborated as the possessive N + N construction, as in *hyenswu ppang* Hyenswu bread '(It)'s Hyenswu's (my) bread' or *ike appa-kke*<sup>3</sup> this-thing daddy-thing 'This (is) daddy's'. The final component of the construction, morphological marking on overt subjects, appears at 2;2 years of age. By this point,

1. Most abbreviations are from Lee (1991):

ADD	additive	COMM	committal	EXCL	exclamation	NOM	nominative
ANT	anterior	CONN	connective	IE	informal ending	TOP	topic
ATTRIB	attributive	COP	copula	INTERR	interrogative	VOC	vocative
CIRCUM	circumstantial	DECL	declarative	NEG	negative		

2. In referring to the children's speech, the term 'subject' is used to denote the single argument of the copula construction, with no implication that they have acquired an adult-like subject category.

3. Korean *ke(s)* is a lexical noun meaning 'thing', which can serve as the indeterminate head of prenominal adjectives and relative clauses (Kim, 1987). As the head of a possessive NP, the initial consonant is tensed (written *-kke*), and is analyzable as a possessive suffix on the noun (Choi & Gopnik, 1995).

Hyenswu can produce four-component constructions, and the items appearing in each component slot are becoming increasingly diverse.

Overall, then, Hyenswu's data suggest that the Korean copula construction follows the expected trajectory from a few lexically specific, structurally simple schemas to a more abstract, generalized construction. Denser longitudinal data from additional children is necessary, however, to obtain a less compressed developmental sequence, and to confirm the generality of the order of acquisition in Hyenswu's data.

The increasing complexity of the copula construction can be measured more precisely in terms of the number of components per construction that the children produce. Figures 1 and 2 display the percentage of constructions produced by each child with from one to four components in the early, mid, and later months of the study, with their mothers' average number of components per construction for comparison.

Both children start out with at least 70% of their copula constructions consisting of a single component (the predicate nominal) and about 24% having two components (the predicate nominal plus an overt subject or copula). Hyenswu

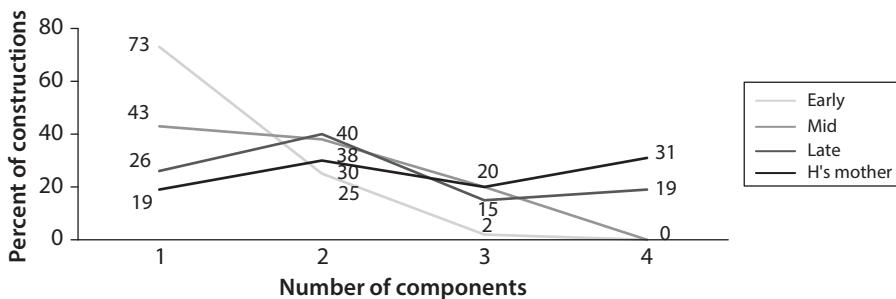


Figure 1. Hyenswu: Increasing complexity of the copula construction.

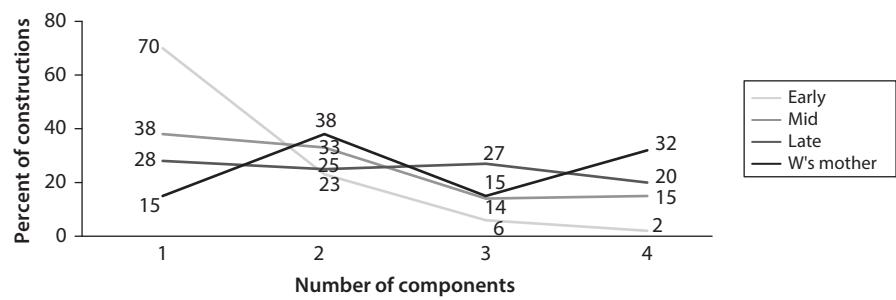


Figure 2. Wenceng: Increasing complexity of the copula construction.

does not begin using forms with three and four components until later than Wenceng, but by the later months of the study both children are using the full form of the construction only 12% less frequently than their mothers.

The children's slowly increasing percentage of more complex forms of the construction, as well as their gradual approximation to their mothers' usage, is consistent with an acquisition model in which children build up the form of constructions gradually based on their experience with caregivers' speech, as prior research on constructions has postulated.

### *Developmental changes in interactional roles*

In Clark's research on English, labeling interactions were found to have a predictable organization with distinct roles for caregiver and child. This kind of routine, illustrated in Example (2), is the earliest and most frequent use of the copula construction in the Korean data.

- (2) Hyenswu (1;11) and her mother are looking at a picture book.

- 1 H's mother: *hyenswu-ya i-ke-n mwe-ya i-ke?*  
hyenswu-voc this-thing-TOP what-COP.IE this-thing  
'Hyenswu, what's this?'
- 2 Hyenswu: *wuywu.*  
milk  
'Milk.'
- 3 H's mother: *wuywu. yo-ke-nun mwe-ya?*  
milk. this.DIM-thing-TOP what-COP.IE  
'Milk. What's this?'
- 4 Hyenswu: *banana-ta.*  
banana-COP.DECL  
'It's a banana.'
- 5 H's mother: *banana-ta. ung. i-ke-nun?*  
banana-COP.DECL yes this-thing-TOP  
'It's a banana. Yes. What about this?'
- 6 Hyenswu: *kkem.*  
gum  
'Gum.'
- 7 H's mother: *ppang. ppang-i-ci kkem-i-ni? ppang.*  
bread bread-COP-COMM gum-COP-INTERR bread  
'Bread. Is it bread or gum? Bread.'

As shown in this example, labeling routines are usually organized as repetitions of the following sequence: (1) the questioner asks the respondent to label a present item, (2) the respondent answers with a label, and (3) the questioner (optionally) provides an assessment, e.g., using an agreement token, repeating the label to acknowledge/accept it, questioning the label to confirm/challenge it, or supplying a different word to correct the child's label. In (2), this question-answer-assessment format is repeated three times.

When interactional roles are clearly distinguished, as in the labeling routine, each role may be associated with different linguistic choices. As Example (2) demonstrates, the copula construction takes different grammatical forms for questioner vs. respondent. The questioner usually uses a demonstrative pronoun subject and a copula when introducing the subject referent to be labeled (lines 1 and 3), while the respondent usually produces only the label (lines 2 and 6), but sometimes adds a copula (line 4). This pattern of using a fuller form in questions than in answers is a very general one in these caregiver-child interactions.

In the labeling routine, the mothers usually enact the role of questioner, while the children take the role of respondent, resulting in an asymmetrical distribution of questions and answers. Figure 3 displays the children's much higher percentage of answers during the year compared with their mothers. Since questions and statements use fuller forms of the copula construction than answers, the complexity of the children's constructions increases as they spend less time in the role of respondent.

Conversely, Figure 4 shows the mothers' much higher percentage of questions compared with the children.

There are certain individual differences; for example, Hyenswu does not reach Wenceng's proportion of questions until mid-year. But as Figures 3-4 show, there is no overlap between the percentages of questions vs. answers produced by the children and their mothers; throughout the year, they tend to assume different interactional roles when using the copula construction.

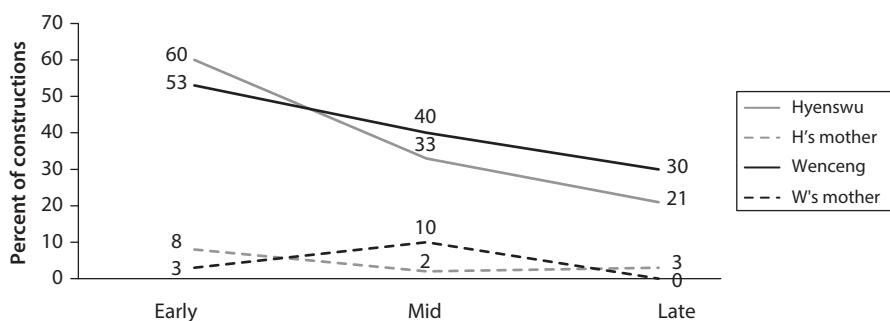


Figure 3. Percent of copula constructions used as answers.

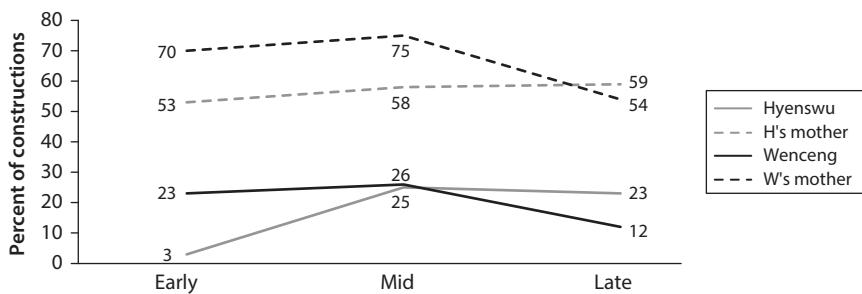


Figure 4. Percent of copula constructions used as questions.

#### *Developmental changes in function*

During the year, the function as well as the form of the copula construction undergoes change. Labeling, i.e., providing a lexical category or classification for the subject referent, is initially the most common function. As the labeling routine becomes less common, the copula construction is increasingly used to describe referents, i.e., to provide information about their properties. The earliest and most frequent type of description encodes possession, as in (3).

- (3) Wenceng (1;9) wants to drink juice from a bottle, but her mother has insisted that she use a cup.

Res. Asst.: *i-ke nwukwu khep-i-ya?*  
 this-thing who cup-COP-IE  
 'Whose cup is this?'

Wenceng: *wencengi khep.*  
 wencengi cup  
 'Wenceng's (my) cup.'

Other frequent types of description indicate the color, size, and similarity (*ttok kathun ke* 'exact same thing') of objects.

Using the copula construction to identify referents, i.e., to specify a unique individual, is comparatively rare in these interactions. Naming the individuals in picture books or photo albums and assigning roles to participants and props in pretend play are the two most common activities for identifications. Depending on the context, the subject referent is encoded with a demonstrative pronoun, as in labels and descriptions; this is typical of picture-based identifications such as (4).

- (4) Hyenswu (2;5) is looking at a pictures in a photo album with the research assistant.

Hyenswu: *i-key appa-ya?*  
 this-thing.NOM daddy-COP.IE  
 'Is this daddy?'

Compared with labels and descriptions, however, identifications are more likely to have a lexical noun (Example 5) or personal pronoun (Example 6) as subject.

- (5) Hyenswu (2;8) has her doll, Kunceng, and is playing with her father.  
 Hyenswu: *ne-ka kuncengi appa-ya.*  
 2p-NOM kuncengi daddy-COP.IE  
 ‘You’re Kunceng’s daddy.’
- (6) Hyenswu (2;4) is putting her doll to bed, but doesn’t have pajamas for her.  
 Hyenswu: *camos-i ppansu.*  
 pajama-NOM panties  
 ‘Her pajamas are panties.’

The following figures display the distribution of the three most frequent semantic functions of the copula construction during the year – labeling, describing, and identifying referents – for the children (Figure 5) and their mothers (Figure 6).

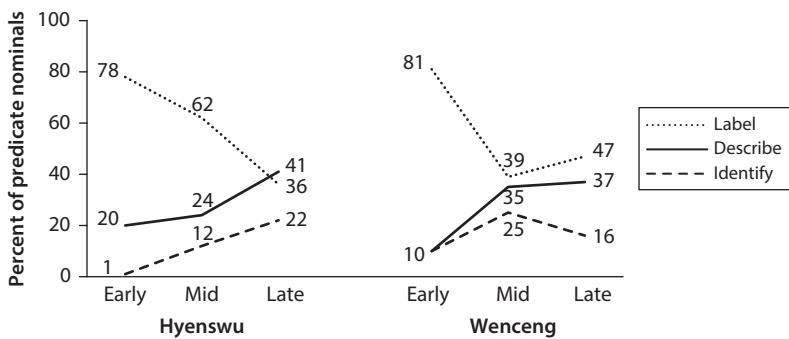


Figure 5. Children: Percentage of predicate nominals used to label, describe, and identify subject referents.

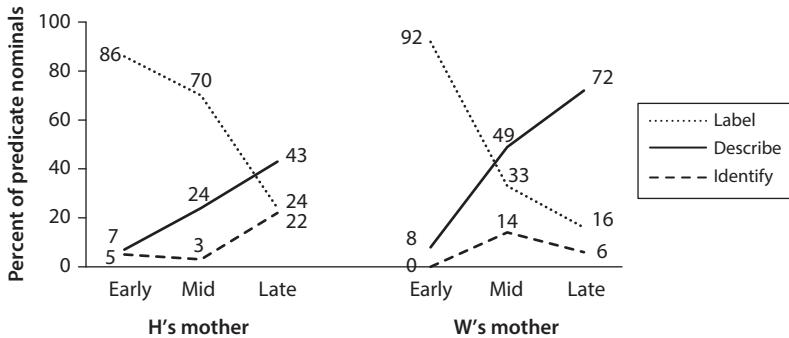


Figure 6. Mothers: Percentage of predicate nominals used to label, describe, and identify subject referents.

For all speakers, the early preponderance of labeling gives way to an increasing use of description; identifications are comparatively rare. (One unusually long sequence of pretend play is responsible for the high frequency of identifications in the later data for Hyenswu and her mother.) Overall, the developmental progression with the copula construction is from heavily scripted labeling routines to freer conversation in which objects are described, and the people and objects in pictures and pretend play are identified.

As these developmental changes in the functions and speech acts encoded with the copula construction clarify, the children are not merely acquiring grammatical forms; they are learning culturally valued ways of interacting with their caregivers, in which particular forms of the copula construction are mastered along with the concomitant interactional roles and communicative functions. The following sections take a closer look at the four components of the construction: subject, morphological marker, predicate nominal, and copula. As we shall see, each component is sensitive to discourse-functional factors that motivate its form and use during the acquisition process.

### *Subject forms*

The subject of the copula construction takes various grammatical forms: ellipsis, demonstrative pronouns, lexical nouns, and personal pronouns. Table 3 gives the proportional frequency of these forms produced by each speaker in the course of the study. It is important to note that pervasive ellipsis is characteristic of ordinary adult-adult speech in Korean, and not merely of child-directed speech. Similarly, personal pronouns, which are highly sensitive to the social context, are rare in adult speech compared with ellipsis.

By far the two most frequent subject forms are ellipsis and demonstrative pronouns, which exhibit a strong asymmetry between the mothers and children; the children have a much higher percentage of elliptical subjects than their mothers,

Table 3. Distribution of subject forms.

FORMS	Hyenswu	Wenceng	H's mother	W's mother
Ellipsis	312 66%	579 69%	110 48%	84 47%
Demonstrative	141	238	100	75
Pronoun	30%	29%	44%	42%
Noun	12 3%	16 2%	14 6%	15 8%
Personal Pronoun	5 1%	2 0.2%	3 1%	1 1%

**Table 4.** Order of development of subject forms.

Form of Subject	Hyenswu		Wenceng	
	Emerged	Established	Emerged	Established
Ellipsis		1;10		1;8
Demonstrative pronoun	1;11	2;0		1;8
Noun	1;11	2;7	1;8	2;6
Personal pronoun	2;7	2;8	2;5	

who use approximately equal proportions of ellipsis and demonstrative pronouns. Nominal subjects, including both lexical nouns and proper names, are rare, but more common in the mothers' speech; personal pronouns are extremely rare for all four speakers.

The order in which subject forms are acquired mirrors these frequencies, as shown in Table 4. (Since nouns and pronouns are very rare, the criterion for "established" status in Table 4 is three correct uses in a single recording.)

Hyenswu's developmental sequence suggests that the earliest uses of the copula construction lack an overt subject. This is consistent with research showing that Korean children initially use verbs without subjects (Choi, 1998; Kim, 1997). Children acquiring languages like English must progress beyond this usage, but subjectless sentences are fully grammatical in adult Korean, depending on the discourse context. Given the high frequency of ellipsis in adult speech, the task of the Korean child is best conceptualized as learning when to produce an overt subject, rather than learning when to use ellipsis.

In these interactions, ellipsis is pragmatically appropriate most of the time, since the subject referent is almost always present in view and can be established as the focus of attention non-verbally; contextual notes taken during the recording sessions indicate numerous instances of pointing at subject referents. In addition, the overwhelming majority of elliptical subject referents have just been mentioned in the immediately prior clause (Hyenswu: 95%, H's mother: 84%, Wenceng: 78%, W's mother: 81%). Thus there are strong motivations in the verbal and nonverbal discourse context for subject ellipsis.

The frequency of demonstrative pronoun subjects – usually *ike* 'this' – also has clear a discourse-pragmatic basis. In contrast with elliptical subjects, the referents of demonstrative pronouns are more likely to encode new information and hence to warrant explicit mention. For Hyenswu and her mother, 52% of demonstrative pronoun subjects have new referents; for Wenceng and her mother, 38%. The referents of demonstrative pronouns are also almost invariably inanimate, and these three properties – co-presence, newness, and inanimacy – combine to make demonstrative pronouns the most common form of overt subject.

How can the asymmetrical distribution of ellipsis and demonstrative pronouns in the speech of the children vs. their mothers be accounted for? At least two types of reasons can be proposed: developmental (in learning when to use overt subjects, the children initially use ellipsis in contexts where their mothers use an overt form) and social-interactive (the organization of caregiver-child conversation affords more opportunities for the children to use ellipsis). To evaluate a developmental explanation, cases in which a new referent is introduced into discourse are especially relevant, since this is a context that elicits overt mention in adult speech (Clancy, 1997). Example (7) shows a typical case in which the child mentions a new referent with an elliptical subject.

- (7) Wenceng (1;9) has been drawing a picture.

Wenceng: *say-ya.*  
*bird-COP.IE*  
 'It's a bird.'

Since Wenceng's drawing is present and in full view of the addressees, it is not pragmatically inappropriate to use ellipsis rather than an explicit form such as *ike* 'this' to mention the new referent. Nevertheless, the mothers generally use demonstrative pronoun subjects when introducing new referents, even when they are in full view. Table 5 shows the children's percentage of elliptical vs. demonstrative pronoun subjects over time, with their mothers' frequencies for comparison.

Compared with their mothers, the children start out with a much higher rate of ellipsis for new referents and a correspondingly lower rate of demonstrative pronouns, but exhibit a strong shift toward their mothers' pattern by mid-year. Thus development proceeds from an early reliance on non-verbal context – the addressee's shared attention to the intended referent – to a more adult-like use of demonstrative pronouns for new referents.

Table 5. Ellipsis vs. demonstrative pronouns for new subject referents.

	Hyenswu		Wenceng	
	Ellipsis	Dem. Pro.	Ellipsis	Dem. Pro.
Early	10	2	49	25
	83%	17%	65%	33%
Mid	22	40	23	37
	34%	62%	37%	60%
Late	14	31	26	59
	32%	66%	28%	63%
Mother	17	52	5	29
	24%	72%	14%	81%

The interactional basis for the children's comparatively high rate of subject ellipsis lies in their role as respondent in question-answer exchanges with the copula construction. Since questions typically seek information about new referents, both mothers and children use a higher percentage of demonstrative pronoun subjects in questions than in answers, as shown by the following ratios: Hyenswu: 62:8, Wenceng: 53:8, H's mother: 65:8, W's mother: 59:0. Conversely, there is a lower percentage of ellipsis in questions vs. answers: Hyenswu: 32:90, Wenceng: 45:92, H's mother: 25:83, W's mother: 30:100. Since the mothers usually take the role of asking questions about new referents, while the children respond with information about the just-mentioned referent, the children are much more likely to use elliptical subjects than their mothers. Thus there is a clear interactional motivation for the asymmetrical distribution of subject forms, in addition to the developmental shift from ellipsis to overt mention for new referents.

As shown in Table 3, nouns and personal pronouns are rare as subjects of the copula construction. When labeling objects, lexical subjects would be anomalous pragmatically, since providing a category for the subject referent is the function of the predicate nominal. Nominal reference is more frequent for the mothers because they are more likely to talk about human referents, using lexical nouns or proper names, as in (8).

- (8) Hyenswu (2;3) has stuffed her mouth with apple slices.

H's mother: *hyenswu twayci-ci?*  
*hyenswu pig-COP.COMM*  
 'Is Hyenswu (are you) a pig?'

Personal pronouns are extremely rare for both mothers and children for various reasons. Korean lacks the equivalent of English third person pronouns, and more than 90% of subject referents in the copula construction are third person. First and second person reference is typically lexical; Korean parents refer to themselves with kin terms such as *emma* 'mommy' and to their children by name, as in (8), while the children use kin terms for their parents. For self-reference, the children start out using their own names, but by about mid-year they shift to the first person pronoun *na*, as in (9).

- (9) Wenceng (2;7) is playing with a naked doll and the research assistant comments that girls should wear clothes.

Wenceng: *na yeca-ya.*  
 1p girl-COP.IE  
 'I'm a girl'

Identifications with lexical or pronominal subjects do sometimes occur, as in (5) and (6), but except for occasional comments like (8) and (9), the children and their mothers rarely apply labels or descriptions to themselves or one another.

In sum, the frequency and distribution of subject forms in the copula construction reflect their discourse functions. New, co-present subjects are usually introduced with demonstrative pronouns, often in questions, while just-mentioned referents are typically left unexpressed in answers. As the children increasingly take on the role of questioner in interaction, they rely less on ellipsis and start using the referential form appropriate for that role – a demonstrative pronoun. The use of nouns and personal pronouns as subjects is limited by the fact that participants in face-to-face interaction rarely need to label, identify, or describe one another, except in contexts that are irrealis (e.g., pretend play) or displaced in time (e.g., photo albums). It is clear that in acquiring particular subject forms, the children are also learning the discourse factors motivating their use.

### *Morphological markers*

When the subject of the copula construction is expressed overtly, a morphological marker is sometimes attached. Use of a marker is not obligatory in spoken Korean; rather, it has a strong basis in the discourse context. Opportunities for using subject markers in these data are comparatively rare, since only about 50% of the mothers' subjects and 30% of the children's subjects are overt; when overt subjects are used, they are often left unmarked. The developmental task for Korean children, then, is to learn which discourse contexts warrant use of morphological marking.

The most common subject markers in the copula construction are the topic -(n)un, nominative -i/ka, and additive -to 'also, even' particles. Their use in this construction differs rather dramatically from overall usage. Figure 7 gives the percentage of topic, nominative, and additive particles on overt subjects for verbs other than the copula.

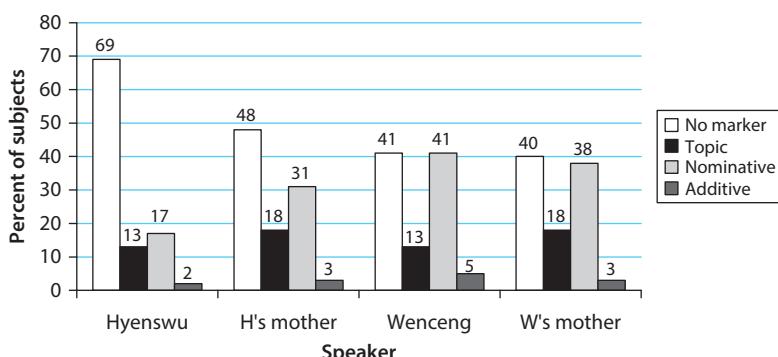


Figure 7. Distribution of morphological markers on overt subjects of verbs other than the copula.

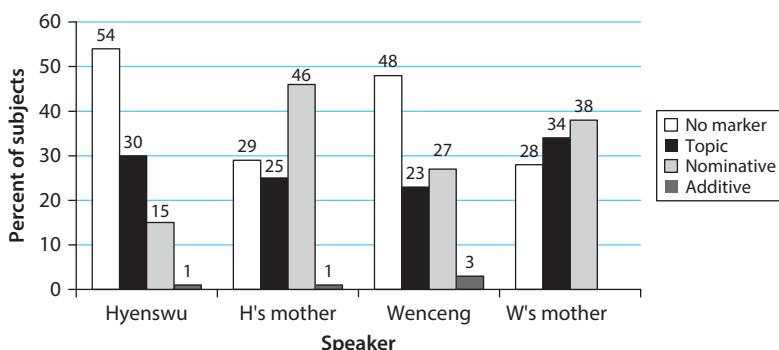


Figure 8. Distribution of morphological markers on overt subjects of the copula.

In general, approximately 40–50% of overt subjects are left unmarked. Nominative marking is more frequent than topic marking, and the additive particle is quite rare. (Hyenswu's high rate of unmarked subjects reflects her comparatively late acquisition of the relevant morphemes, especially the nominative particle.)

In the copula construction, there is a different distribution of subject markers, shown in Figure 8. Although the picture is complicated, in general there is a higher rate of topic markers and a lower percentage of unmarked subjects; Wenceng is exceptional in having a somewhat higher rate of unmarked subjects.

The prominence of topic marking in the copula construction is reflected in the order of acquisition for subject markers, which is presented in Table 6. Topic marking is acquired first, followed by nominative marking; the additive marker emerges last and remains rare.

Why is the copula construction so conducive to morphological marking in general and to topic marking in particular? Since markers can only occur when a subject is overt, they are necessarily associated with discourse contexts in which it is appropriate to use an overt subject, such as when the subject referent is contrastive (Clancy, 1997). One important function of the topic marker is to mark contrast (Lee, 2003); this is the earliest function for the topic particle that Korean

Table 6. Order of development for topic, nominative, and additive markers.

Subject marker	Hyenswu		Wenceng	
	Emerged	Established	Emerged	Established
Topic	2;2	2;5		1;8
Nominative	2;3	2;7	1;10	2;1
Additive	2;5		2;2	

children acquire (Kim, 1997). In the labeling routine, contrastive topic-marking has the following format: A-TOP is an X; B-TOP is a Y. This use of topic marking highlights the relationship between successive predicates, which are both labels but differ in semantic content, as well as the relationship between successive subject referents, which are both members of the same set of present items, such as the people in a photo album, but contrast in that they refer to different entities. The contrastive function of the topic marker is illustrated in (11):

- (11) Hyenswu (2;5) is looking at pictures in a photo album.  
*i-ke-n appa-ya! i-ke-n oppa-ya!*  
 this-thing-TOP daddy-COP.IE this-thing-TOP uncle-COP.IE  
 ‘This one is daddy! This one is uncle!’

Like topic marking, nominative marking often has a contrastive function, although of a subtly different type. This function of the nominative particle, described as exclusive or exhaustive focus (Lee, 2003), highlights the subject referent as one among various candidates to which a particular predicate applies, as in the English cleft construction ‘It is X who/that \_\_\_\_’. Example (12) illustrates this use of the nominative particle. (Nominative marking with *-ke* ‘thing’ surfaces as a final raised vowel [e], written as *-ey*.)

- (12) Hyenswu (2;5) is looking at a photo album.
- |               |   |
|---------------|---|
| 1 → Hyenswu:  | <i>i-key appa-ya?</i>                       |
|               | this-thing.NOM daddy-COP.IE                 |
|               | ‘Is this one daddy?’                        |
| 2 Res. Asst.: | <i>an-i-ntey.</i>                           |
|               | NEG-COP-CIRCUM                              |
|               | ‘It’s not.’                                 |
| 3 Hyenswu:    | <i>appa an-i-ntey?</i>                      |
|               | daddy NEG-COP-CIRCUM                        |
|               | ‘It’s not daddy?’                           |
| 4 Res. Asst.: | <i>ung.</i>                                 |
|               | ‘Yes.’ (It’s not daddy.)                    |
| 5 → Hyenswu:  | <i>i-key appa-ya, i-key.</i>                |
|               | this-thing.NOM daddy-COP.IE, this-thing.NOM |
|               | ‘This one is daddy, this one.’              |
| 6 Res. Asst.: | <i>ung.</i>                                 |
|               | ‘Yes.’                                      |

In line 1, Hyenswu uses the nominative particle as she (incorrectly) identifies one potential referent as her father; following correction, she chooses the appropriate

subject referent in line 5, again with nominative marking. Although nominative marking is not more frequent than usual in the copula construction, opportunities for the exhaustive focus function of the nominative particle are common, e.g., when the accuracy of children's labels and descriptions is at issue, as in (12).

Thus the discourse contexts in which the copula construction is used, as well as the organization of conversational sequences featuring this construction, are well-suited both to the contrastive function of the topic marker and to the exhaustive focus function of the nominative marker. Further research is necessary to shed light on other functions of morphological markers in the copula construction, but the discourse motivations for contrastive subject marking in this construction are apparently accessible to Korean children at an early age.

### *Predicate nominal*

The predicate nominal is the heart of the copula construction. It is the first component of the construction to be acquired, and is omitted only in those rare cases when the discourse context renders it fully predictable, e.g., in the question *ike-nun?* this-TOP 'What about this one?' in a series of label-seeking questions (e.g., Example 2, line 5).

While the subject of the copula construction remains a simple NP, the predicate nominal develops from a single-noun label to increasingly complex noun phrases. This development is functionally motivated, reflecting the shift from labeling to describing in the use of the copula construction (Figures 5–6). The function of description is typically accomplished simply by modifying the head of the predicate nominal.

Table 7 summarizes the developmental changes that occur in the grammatical form of the predicate nominal as the functions of the copula construction expand.

Table 7. Order of development of predicate nominal forms.

Predicate Nominal Type	Hyenswu		Wenceng	
	Emerged	Established	Emerged	Established
Noun		1;10		1;8
<i>mwe</i> 'what'	1;10	2;6		1;8
N- <i>kke</i> 'thing' (possession)	1;10	1;11	1;9	1;9
N + N (possession)	2;1	2;6	1;8	1;9
N + N	2;3	2;8	1;8	1;10
ADJ- <i>ke</i>	2;3		1;10	2;7
ADJ + N			1;9	1;10
Rel. clause- <i>ke</i>	2;5	2;8	1;11	2;3
Rel. clause + N			1;11	2;1

In the first recording session, both children can already produce one-word predicate nominals in the labeling routine, as illustrated in Example (1), as well as ‘What is it?’ questions with *mwe* ‘what’. Next, the descriptive function emerges, initially with the possessive construction, either in [Possessor-*kke* ‘thing’] format, as in Examples (13) and (14) below, or in N + N [Possessor + Possessed] format with a specific lexical head, e.g., *wencengi khep* ‘Wenceng’s cup’ in Example (3). Perhaps because she has an older sister, Hyenswu’s possessives often occur in contexts where ownership is being challenged, as in (13). For both children, possessives are also common when items such as food or playthings are being distributed, as in (14).

- (13) Hyenswu (2;0) and her sister are arguing over a slice of apple.

- Sister: *i-ke hyenswu-kke an-i-ci.*  
           this-thing hyenswu-thing NEG-COP-COMM  
           ‘This isn’t Hyenswu’s (yours).’
- Hyenswu: *hyenswu-kke-ya.*  
           hyenswu-thing-COP.IE  
           ‘It’s Hyenswu’s (mine).’

- (14) Wenceng (1;9) is handing out crayons.

- Wenceng: (giving the Research Assistant a crayon):  
*yoke acwumma-kke.*  
           this.DIM-thing aunt-thing  
           ‘This (is) auntie’s.’

Prenominal adjectives emerge next, at first mainly in set phrases, such as *ttok kath-un ke/N* ‘exact same thing/N’ or with color adjectives, as in (15).

- (15) Wenceng (1;9) is drawing with crayons.

- Wenceng: *say-ya. ppala-n say-ya.*  
           bird-COP:IE blue-ATTR bird-COP.IE  
           ‘It’s a bird. It’s a blue bird.’

Relative clauses are the last form of predicate nominal to be acquired. The earliest relative clauses are used to characterize the function of the subject referent, as in (16).

- (16) Wenceng (1;11) is looking at a newspaper ad for a vacuum.

- Wenceng: *emma i-ke pwu-pwung ha-nun-ke-ya.*  
           mommy this-thing vroom-vroom do-ATTR-thing-COP.IE  
           ‘Mommy, this is the thing that goes vroom-vroom.’

After such generic characterizations, relative clauses describing the subject referent in terms of specific events appear, as in (17).

- (17) Wenceng (2;6) is playing with a purse and trying to close it.

Wenceng: *i-ke samchon-i sa-cwu-n-ke-ya.*  
 this-thing uncle-NOM buy-give-ATTR-thing-COPIE  
 ‘This is the thing uncle bought for me.’

Although descriptions are sometimes used in question-answer-evaluation routines, as in a series of questions about the color of different objects, they are less likely than labels to be prompted; descriptions typically appear as spontaneously produced statements. And while they do occur in the context of looking at pictures in books, as in (16), descriptions are generally used in talking about co-present objects that are being acted upon, as in Examples (13-15) and (17). In the function of describing referents, the predicate nominal of the copula construction affords a high-frequency, structurally expandable site for the acquisition of complex NPs.

### *Copula*

The Korean copula *-i-*, which attaches to the head of the predicate nominal, has a grammatical rather than semantic function; serving as a carrier for inflections, the copula enables predicate nominals to participate in the system of verbal suffixes. Following a vowel-final noun, the copula elides completely and verbal suffixes attach directly to the predicate nominal. Although the copula lacks semantic content and has only minimal phonetic form, since Korean is an SOV language, it usually occurs sentence-finally – a salient position for language learners (Slobin, 1985).

Acquisition of the copula poses a number of challenges. First of all, it is not always used in adult speech; Hyenswu’s mother, for example, uses a copula in 69% of her copula constructions, Wenceng’s mother in 82% of hers. Moreover, use of the copula is unevenly distributed across different types of predicate nominal head, as shown in Table 8. Both mothers and children use the highest percentage of copulas following wh-words, a somewhat lower rate following the generic head noun *-ke* ‘thing, one’, and the lowest frequency following lexical head nouns.

Table 8. Percentage of predicate nominal head types followed by a copula.

Predicate Nominal	Hyenswu	Wenceng	H’s mother	W’s mother
Lexical	111 36%	218 34%	68 54%	77 77%
Generic head noun <i>-ke</i> ‘thing’	69 67%	64 78%	15 94%	16 94%
Wh-word	39 93%	96 100%	74 99%	54 100%

From the perspective of grammaticalization, this distribution makes sense. The copula is less likely to form a grammaticalized bond with lexical heads than with the generic head noun *-ke*, since no single lexical head is as frequent as *-ke*. As a head noun, *-ke* occurs with various types of descriptive predicate nominals, including possessives (e.g., *hyenswu-kke* ‘Hyenswu’s’), prenominal adjectives (e.g., *khun-ke* ‘big-one’), and relative clauses (e.g., *acwumma-ka sacwun-ke* ‘the one auntie bought for me’). The copula is most likely to form a grammaticalized bond with wh-words such as *mwe* ‘what’ and *nwukwu* ‘who’, which are short, high-frequency, and lack the diversity and complexity of the descriptive material preceding *-ke*. With the most frequent wh-word, *mwe* ‘what’, there is also a need to avoid ambiguity; as in English, *mwe?* ‘What?’ without a copula indicates that the speaker has not heard/understood the prior utterance.

As Table 8 shows, the children’s use of copulas, although less frequent than their mothers’, roughly approximates the same distribution across types of predicate nominal head. The children’s production of copulas with wh-words is at (Wenceng) or close (Hyenswu) to adult frequency, and like their mothers, they use copulas more frequently with the indeterminate head *-ke* than with lexical head nouns. Mastery of this distribution does not depend solely on differentiating between different types of predicate nominal head; in fact, there are strong functional correlates to support its acquisition. For example, wh-words with copulas generally appear in the questioner’s turn and lexical predicate nominals without copulas in the respondent’s turn, while descriptions with *-ke* tend to occur in spontaneously produced statements.

This association between use of the copula and speech acts is shown in Table 9. The copula is most frequent in questions, least common in answers, and of intermediate frequency in statements. (Answers are too few in the mothers’ speech to yield reliable percentages, as indicated by parentheses.) The distribution of copulas in Table 9 mirrors the distribution of overt subjects. The mothers use both overt subjects and copulas in their questions, while the children, in their role as respondents, tend to produce answers lacking both overt subjects and copulas, as in Example (1), excerpted below as (18).

Table 9. Percentage of copulas in answers, statements, and questions.

Speech Acts	Hyenswu	Wenceng	H's mother	W's mother
Answers	157	360	11	7
	17%	23%	9%	43%
Statements	219	293	88	45
	55%	52%	51%	62%
Questions	90	172	124	115
	77%	80%	86%	90%

- (18) Hyenswu (1;11) and her mother are looking at a picture book.

H's mother: *hyenswu-ya i-ke-n mwe-ya i-ke?*  
                   hyenswu-voc this-thing-TOP what-COP.IE this-thing  
                   ‘Hyenswu, what’s this?’

Hyenswu: *wuywu.*  
                   milk  
                   ‘Milk.’

In acquiring their mothers’ distribution of copulas across the different types of predicate nominal head, the children are simultaneously mastering the different interactive functions associated with that distribution.

The organization of conversational sequences is another discourse-functional factor that potentially facilitates acquisition of the copula construction. Beginning with the labeling routine, the copula construction is often used in dense clusters of cohesive conversational sequences, with a progression from question-answer sequences to more flexible patterns in which, for example, labels are followed by various types of descriptions.

Table 10 presents the frequency of three types of conversational sequence in which the adult interlocutor models use of an overt copula: (1) immediately prior to the child’s use of a copula, (2) immediately following the child’s copula, and (3) in an expansion of a child’s copula construction that lacked a copula. The first two rows in the table are based on the children’s first 50 uses of the copula. The top row gives the frequency with which an adult copula occurs in the turn immediately prior to the child’s copula, while the second row gives the frequency with which the child’s use of a copula is reinforced by an adult copula in the immediately following turn. The third row is based on a different data set – the children’s first 50 uses of the copula construction without a copula – and gives the frequency with which these constructions are followed by an adult expansion with a copula in the immediately following turn.

Table 10. Conversational sequences with adult modeling of the copula.

Type of conversational sequence	Hyenswu	Wenceng
adult copula → child copula	23 46%	13 26%
child copula → adult copula	21 42%	14 28%
no child copula → adult copula	22 40%	21 44%

As the first two rows of Table 10 show, Hyenswu, who was not yet producing the copula spontaneously at the start of the study, receives a higher rate of preceding and following adult support than Wenceng. In (18), for example, the adult's preceding question provides a model for one of Hyenswu's earliest uses of the copula.

- (19) The adults are trying to get Hyenswu (2;0) to label a toy cake.

Res. Asst.: *ppang-i-ci?*  
cake-COP-COMM  
'Is it a cake?'

Hyenswu: *mwul-i-ci, mwul.*  
water-COP-COMM water  
'It's water, water.'

Acknowledgements that repeat the child's utterance provide immediate reinforcement for the children's use of a copula, as when Hyenswu's mother acknowledges her production of the label *panana-ta* banana-COP.IE 'It's a banana' by repeating *panana-ta* banana-COP.IE 'It's a banana' (Example 2, lines 4-5). For both children, production of a copula construction without the copula is followed at least 40% of the time by an adult expansion with a copula, as in (19).

- (20) Hyenswu (2;3) finishes singing a song, and the adults applaud.

1 Hyenswu: (picking up a purse)  
*cwumeni.*  
pocket  
'Pocket.'

2 → H's mother: *cwumeni-ya? cikap-i-ta.*  
pocket-COP.IE purse-COP-DECL  
'Is it a pocket? It's a purse.'

Thus the sequence of conversational turns in these interactions provides the children with ample context-sensitive models for using overt copulas appropriately. Such close juxtapositions of incomplete child forms with full adult forms have long been cited as a type of indirect negative evidence supporting the acquisition of grammar (e.g., Sokolov & Snow, 1994). A discourse-functional approach highlights the fact that such evidence arises naturally in conversational sequences that elicit, acknowledge, and correct the child's speech.

## Discussion and conclusions

The findings of this study provide a variety of evidence for the discourse foundations of the Korean copula construction in early caregiver-child interaction. Each

component of the construction is sensitive to the discourse context; from the outset, what the children are acquiring is a set of context-specific versions of the copula construction that serve particular discourse functions.

With respect to the first research question of this study – how discourse function impacts the form of the copula construction – the results clarify the relationship between grammatical form and discourse function for each component of the construction. The form of the subject depends on the physical and linguistic co-presence of its referent, as Clark and Estigarribia (2011) have found for English. Elliptical subjects are preferred for previously mentioned (i.e., given) referents, while demonstrative pronouns – usually *ike* ‘this’ – highlight new, co-present referents. The distribution of these subject forms in discourse is organized to a large extent by the participants’ interactional roles. In their predominant role as questioner, the mothers introduce a higher rate of new referents and therefore use a higher proportion of demonstrative pronoun subjects. In their typical role as respondent, the children use a higher percentage of elliptical subjects because most of their subject referents have just been mentioned in the preceding question and therefore constitute given information.

Discourse context also has a major impact on the forms of the other three components of the copula construction. Morphological markers on overt subjects – as well as the selection of topic vs. nominative markers – are sensitive to whether and how the subject referent is being contrasted with other referents in the discourse. The form of the predicate nominal varies in complexity depending on whether the copula construction is being used to describe, rather than to label or identify, the subject referent. Production of an overt copula varies with the type of head noun in the predicate nominal (wh-word, generic head noun *-ke*, or lexical noun); in turn, the distribution of these types is associated with different functions for the construction (describing vs. labeling or identifying) and different types of speech acts (question, answer, statement).

When constructions are situated in discourse, the organization of conversation itself can be recognized as a potential facilitator of the acquisition process. As Clark’s research has shown, labeling interactions create a highly predictable structure for grammatical forms and speaker roles. In these Korean caregiver-child interactions, cohesive, repetitive sequences of turns juxtapose grammatical similarities and differences in form across turns, while associating each form of the copula construction with particular speech acts and interactional roles. Contrastive topic marking of the subject is probably facilitated by conversational sequences in which referents and their labels are contrasted in successive turns. Similarly, sequences of turns in which the child’s choice of subject referent for a particular predicate nominal is assessed or corrected provide an ideal discourse context for acquiring the exhaustive focus function of nominative marking. Repetition and

alternation of forms across multiple turns may facilitate acquisition of the skewed distribution of the copula, in which different types of predicate nominal head are associated with particular functions, such as describing vs. labeling, along with the related interactional roles and speech acts.

Findings with respect to the second research question – how the grammatical form and discourse functions of the copula construction change over time – also provide support for a discourse-based view of constructions. Use of overt subjects and copulas increases over time, as the children appropriate the role of questioner – a role that calls for introducing new referents with overt subject forms and grammaticalized wh-form + copula sequences. The core component of the copula construction, the predicate nominal, is structurally elaborated over time, reflecting a shift in function from labeling referents with a simple lexical noun to describing them with increasingly complex NPs. The shift from labeling to description also involves a shift in speech acts and interactional roles, as the question-answer format characteristic of labeling gives way to more varied conversational sequences in which the child spontaneously provides descriptive information in declarative statements.

The results of this study also suggest a positive answer to the third research question: whether there is a developmental transition from lexically specific structural formulas to more abstract, general constructions, as documented for English. The high frequency of particular goals in these Korean caregiver-child interactions (e.g., to engage the child in conversation and elicit talk), as well as the frequency of particular activities (e.g., reading books and playing with toys), initially limits use of the copula construction to talk about referents with specific semantic and discourse-pragmatic properties (e.g., labeling of co-present inanimate objects). The specificity of these discourse contexts leads to use of a very limited set of forms at first, both by the mothers (*ike N-ya* ‘This is a N’, *ike-nun mwe-ya?* ‘What’s this?’) and by the children (*ike N* ‘this (is a) N’, *mwe-ya* ‘What is it?’, *N-ya* ‘(It’s a) N’). These simple schemas are essentially the Korean equivalents of those found by Clark and Wong (2002), Pine et al. (2008), and Wilson (2003) in English. While denser longitudinal data is required to clarify details of this developmental transition, there is preliminary support for a constructionist view of acquisition, in which a small set of structural schemas is built up incrementally into a more general, abstract construction. At the discourse level, this evolution mirrors the transition from rigidly structured interactional routines to more loosely organized conversational sequences (Peters & Boggs, 1986).

The discourse foundations of the copula construction go deeper than its immediate contexts of use. Caregivers who ask young children questions that test and display their knowledge of labels, properties of objects, and identities of people, are enacting culture-specific practices with strong ideological roots. For example,

the belief that mothers play a tutorial role in their children's language development and that it is appropriate for children to display their knowledge to social superiors are well-documented ideologies for middle-class caregivers of European and Asian ancestry (Clancy, 1989; Heath, 1983; Ochs & Schieffelin, 1984). These ideologies are foregrounded when caregivers' skills as tutors and the accuracy of their children's knowledge are being observed by researchers.

Cultural ideologies also motivate the activities regarded as appropriate for caregivers and young children. Joint book reading, the single most common interactional setting for the copula construction in these caregiver-child interactions, as in their American counterparts (Clark & Wong, 2002), is rooted in the high cultural value placed on literacy. In these Korean interactions, not only are children regarded as appropriate conversational partners for adults (Ochs & Schieffelin, 1984), but caregivers see themselves as appropriate playmates for children, participating in activities such as assembling puzzles and Lego constructions, pretend play, and drawing. Many of the activities in Korean and American caregiver-child interactions reflect a culture in which material objects are valued and made available to young children.

Cultural associations between types of interaction and grammatical constructions is a fundamental aspect of the discourse basis of constructions, and provides the cultural context in which certain types of interaction facilitate or even provoke the use of particular grammatical forms and constructions (Ervin-Tripp, 2012). Thus there is a clear answer to the fourth research question concerning the cultural ideologies and related activities that provide the cultural matrix for use and acquisition of the Korean copula construction. In these data, the relevant ideologies include the importance of literacy, the mother as tutor, the child as an appropriate conversational partner for adults, adults as appropriate playmates of young children, and hospitality to guests (e.g., sharing food). These ideologies motivate the typical activities during the recording sessions, and a consistent set of discourse and grammatical patterns arising from these activities.

At the most abstract theoretical level, the findings of this study pose a challenge to the notion of an autonomous grammar that can be explained and acquired without reference to discourse. Each component of the Korean copula construction is sensitive to its discourse context in ways that affect the frequency and distribution of forms and the order in which they are acquired by children. Drawing on Eve Clark's foundational work on new words in caregiver-child interaction, this study has attempted to situate the acquisition of the Korean copula construction in its natural discourse context. The findings, I believe, argue for an expansion of research on constructions to include their discourse foundations – a research agenda that can yield a fuller understanding of what constructions are and how they are acquired.

## References

- Bates, E., & MacWhinney, B. (1979). A functionalist approach to the acquisition of grammar. In E. Ochs & B. Schieffelin (Eds.), *Developmental pragmatics* (pp. 167–209). New York, NY: Academic Press.
- Chafe, W. (1976). Givenness, contrastiveness, subjects, topics, and point of view. In C. Li (Ed.), *Subject and topic* (pp. 25–55). New York, NY: Academic Press.
- Choi, S. (1998). Verbs in early lexical and syntactic development in Korean. *Linguistics*, 36(4), 755–780. DOI: 10.1515/ling.1998.36.4.755
- Choi, S., & Gopnik, A. (1995). Early acquisition of verbs in Korean: A cross-linguistic study. *Journal of Child Language*, 22, 497–529. DOI: 10.1017/S0305000900009934
- Chouinard, M.M., & Clark, E.V. (2003). Adult reformulations of child errors as negative evidence. *Journal of Child Language*, 30, 637–669. DOI: 10.1017/S0305000903005701
- Clancy, P. (1989). A case study in language socialization: Korean wh-questions. *Discourse Processes*, 12, 169–191. DOI: 10.1080/01638538909544725
- Clancy, P. (1997). Discourse motivations for referential choice in Korean acquisition. In H. Sohn & J. Haig (Eds.), *Japanese/Korean linguistics VI* (pp. 639–659). Stanford, CA: CSLI.
- Clancy, P. (2004). The discourse basis of constructions: Some evidence from Korean. In E. Clark (Ed.), *Proceedings of the 32nd Stanford Child Language Research Forum* (pp. 20–29). Stanford, CA: CSLI.
- Clark, E.V. (1997). Conceptual perspective and lexical choice in acquisition. *Cognition*, 64, 1–37. DOI: 10.1016/S0010-0277(97)00010-3
- Clark, E.V. (2006). Color, reference, and expertise in language acquisition. *Journal of Experimental Child Psychology*, 94, 339–343. DOI: 10.1016/j.jecp.2006.03.002
- Clark, E.V. (2007). Young children's uptake of new words in conversation. *Language in Society*, 36, 157–182. DOI: 10.1017/S0047404507070091
- Clark, E.V. (2010). Adult offer, word-class, and child uptake in early lexical acquisition. *First Language*, 30, 250–269. DOI: 10.1177/0142723710370537
- Clark, E.V., & Bernicot, J. (2008). Repetition as ratification: How parents and children place information in common ground. *Journal of Child Language*, 35, 349–371. DOI: 10.1017/S0305000907008537
- Clark, E.V., & Estigarribia, B. (2011). Using speech and gesture to introduce new words to young children. *Gestalt*, 11, 1–23. DOI: 10.1075/gest.11.1.01cla
- Clark, E.V., & Grossman, J. (1998). Pragmatic directions and children's word learning. *Journal of Child Language*, 25, 1–18. DOI: 10.1017/S0305000997003309
- Clark, E.V., & Wong, A. (2002). Pragmatic directions about language use: Offers of words and relations. *Language in Society*, 31, 181–212.
- Ervin-Tripp, S. (2012). Pragmatics as a facilitator in child syntax development. In M. Meeuwis & J.-O. Östman (Eds.), *Pragmaticizing understanding: Studies for Jef Verschueren*. Amsterdam: John Benjamins.
- Estigarribia, B., & Clark, E.V. (2007). Getting and maintaining attention in talk to young children. *Journal of Child Language*, 34, 799–814. DOI: 10.1017/S0305000907008161
- Goldberg, A. (1995). *Constructions: A construction grammar approach to argument structure*. Chicago, IL: Chicago University Press.
- Heath, S. (1983). *Ways with words: Language, life and work in communities and classrooms*. Cambridge: CUP.

- Kim, J., Sells, P., & Wescoat, M. (2008). Korean copular constructions: A lexical sharing approach. In M. Endo Hudson, P. Sells, S.-A. Jun, P. Clancy, S. Iwasaki, & S.-O. Sohn (Eds.), *Japanese/Korean linguistics, Vol. 13* (pp. 179–192). Stanford, CA: CSLI.
- Kim, Y. (1987). *The acquisition of relative clauses in English and Korean: Development in spontaneous production*. Unpublished doctoral dissertation. Harvard University.
- Kim, Y. (1997). The acquisition of Korean. In D. Slobin (Ed.), *The crosslinguistic study of language acquisition, Vol. 4* (pp. 335–443). Mahwah, NJ: Lawrence Erlbaum Associates.
- Lee, C. (2003). Contrastive topic or contrastive focus. In W. McClure (Ed.), *Japanese/Korean linguistics, Vol. 12* (pp. 1–13). Stanford, CA: CSLI/University of Chicago Press.
- Lee, H.S. (1991). *Tense, aspect, and modality: A discourse-pragmatic analysis of verbal affixes in Korean from a typological perspective*. Unpublished doctoral dissertation. University of California, Los Angeles.
- Lieven, E., Pine, J. & Baldwin, G. (1997). Lexically-based learning and early grammatical development. *Journal of Child Language*, 24, 187–219. DOI: 10.1017/S0305000996002930
- Ochs, E., & Schieffelin, B. (1984). Language acquisition and socialization: Three developmental stories and their implications. In R. Schweder & R. LeVine (Eds.), *Culture theory: Essays on mind, self, and emotion* (pp. 276–313). Cambridge: CUP.
- Ochs, E., & Schieffelin, B. (1995). The impact of language socialization on grammatical development. In P. Fletcher & B. MacWhinney (Eds.), *The handbook of child language* (pp. 73–94). Cambridge, MA: Blackwell.
- Oh, M. (1991). The Korean copula and palatalization. *Language Research*, 27, 701–724.
- Peters, A., & Boggs, S. (1986). Interactional routines as cultural influences upon language acquisition. In B. Schieffelin & E. Ochs (Eds.), *Language socialization across cultures* (pp. 80–95). Cambridge: CUP.
- Pine, J., Conti-Ramsden, G., Joseph, K., Lieven, E., & Serratrice, L. (2008). Tense over time: Testing the Agreement/Tense Omission Model as an account of the pattern of tense-marking provision in early child English. *Journal of Child Language*, 35, 55–75. DOI: 10.1017/S0305000907008252
- Slobin, D. (1985). Crosslinguistic evidence for the language-making capacity. In D. Slobin (Ed.), *The crosslinguistic study of language acquisition, Vol. 2: Theoretical issues* (pp. 1157–1256). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Sokolov, J., & Snow, C. (1994). The changing role of negative evidence in theories of language acquisition. In C. Galloway & B. Richards (Eds.), *Input and interaction in language acquisition* (pp. 38–55). Cambridge: CUP. DOI: 10.1017/CBO9780511620690.004
- Tomasello, M. (2000). The item-based nature of children's early syntactic development. *Trends in Cognitive Sciences*, 4, 156–163. DOI: 10.1016/S1364-6613(00)01462-5
- Wilson, S. (2003). Lexically specific constructions in the acquisition of inflection in English. *Journal of Child Language* 30, 75–115. DOI: 10.1017/S0305000902005512

# **Emergent clause-combining in adult-child interactional contexts\***

Ruth A. Berman and Lyle Lustigman  
Tel Aviv University

The study analyzes early clause-combining (CC) complex syntax of three Hebrew-acquiring girls in self-initiated autonomous constructions compared to three types of interlocutor-supported contexts – by means of Affirmations (encouraging comments), Scaffolding (structural prompts), and Co-Construction (adult-child clause-combining adjacency pairs). Two of the girls show a marked rise in self-initiated use of explicitly-marked canonical CC constructions along with a concurrent increase in reliance on interlocutor-supported contexts. The third child, who progressed more slowly in producing such autonomous CC constructions, also received relatively little supportive input overall. These findings for an increase in canonical CC usage together with a rise in reliance on supportive contexts suggest that interactive contexts play a role in advancing the consolidation of early complex syntax.

## **A personal note from Ruth Berman**

Writing this chapter was both a pleasure and a trial, since my prior research has not involved issues of interaction, which is at the center of much of Eve Clark's current work on acquisition. So thank you, Eve and the editors of this splendid volume, for the opportunity to expand our horizons in this direction.

---

\* The study is supported by Grant No. 190/10 of the Israel Science Foundation to Ruth Berman and Bracha Nir for the study of "Clause Combining in Hebrew from Early Childhood across Adolescence: Usage-Based Analyses". The authors are indebted to Dr. Nir for her invaluable input, to Netanel Haim for his assistance in coding and CLAN applications, to Enav Kedar for help with editing, and to the editors of this volume and two anonymous reviewers for their excellent feedback on an earlier draft. The authors are responsible for any inadequacies that remain.

## 1. Introduction

This study considers the emergence of clause-combining in early child Hebrew, in the domain termed “acquisition of complex sentences” – for example, by Bowerman (1979) and Limber (1973), more recently by Diessel (2004) or Lust, Foley, and Dye (2009) – a topic reviewed by Clark under the heading of “combining clauses – more complex constructions” (2009: 229–253).<sup>1</sup> We, too, take for our basic unit of analysis the “clause”, to avoid attributing the abstract construct of “sentence” to early child utterances – possibly to oral language in general (Halliday, 1989: 66–67). Focus is on early “clause-combining” (henceforth CC) constructions in children’s speech, from around age two years, once they go beyond the level of the simple clause (Bowerman & Perdue, 1990; Vasilyeva, Waterfall, & Huttenlocher, 2008), and prior to the consolidation of complex syntax around age 4 to 6 years (De Villiers, 1995; Karmiloff-Smith, 1986; Reilly, 1986) and in school-age text construction (Berman & Nir, 2009; Berman & Slobin, 1994:538–554; Silva, Sánchez, & Borzone, 2010).

Several interrelated assumptions underlie the study. First, following Bowerman’s (1979: 286) observation that “The first complex sentences appear after simple sentences about four words long become common”, we suggest that clause-combining emerges only after simple-clause structure is established. Second, the ability to combine two or more predication within a single syntactic envelope is indicative of concurrent linguistic and cognitive development by more condensed and cohesive packaging of information (Berman, 2009; Karmiloff-Smith, 1986; Hickmann, 2003). Third, taking into account Clark’s observation that “Even in two-word combinations, children may actually be combining two distinct propositions or proto-clauses” (2009: 230), children’s initial clause-combining is expected to be marked ambiguously or not at all, and to show an age-related increase in appropriately explicit marking (Berman, 1996). Finally, parental and other contextual support will play a role in children’s early clause-combining, as in other grammatical domains like verb inflection (Clark & de Marneffe, 2012) and stringing of successive single-word utterances (Scollon, 1976; Veneziano, 1999). Our study thus takes into account the nature of *supportive contexts* provided by adult input, on the assumption that with age, children will show increasing ability to produce canonical CC constructions that are independently self-initiated CC and explicitly marked by appropriate lexico-grammatical connectives.

Prior research on the acquisition of complex syntax in Hebrew has generally not taken into account the nature of input-based contextual support. Experimental studies on relative clause acquisition are reviewed in Arnon (2011), who also

---

1. In linguistics, the domain is variously referred to as “complex sentences” (Bybee & Noonan, 2001), “clause-union” (Givón, 2009), or “clause-combining” (Haiman & Thompson, 1988).

investigated the RC use of two siblings aged 2 to 6. Other studies of Hebrew preschoolers' clause-combining include: coordinating constructions (Berman, 1996); distribution of subordinate clauses in conversational and narrative contexts (Berman, 1997); and levels of connectivity as indicators of stages in grammatical development (Berman & Lustigman, 2012). Two studies are particularly relevant to the *emergence* of CC. Dromi and Berman's (1986) cross-sectional survey of interactions between adult investigators and over 100 children aged 1 to 5 years revealed increased use from age 2;0 up of: (a) clauses with lexical verbs versus verbless copular constructions; (b) clauses with "expanded VPs", in which modal and aspectual verbs take an infinitival complement; and (c) "polyclause" utterances – around 8% of all utterances between ages 2 to 3 years, and over one-third at ages 4 and 5, compared with over 90% single-clause utterances below age two. Armon-Lotem (2004) concludes from her analysis of four children between ages 1;4 to 3;0 (two of them participants in the present study) that initially, complex sentences are "preconjunctionals" – utterances with more than one clause, but without the complementizer *še-* 'that' – and that children begin to mark subordinate clauses with *še-* at around age 2;0, at first without, and later with, a preceding main clause. Results of these studies combine with findings from English (e.g., Clark, 2009; Diessel, 2004) in identifying age two years as a critical period for the emergence of clause-combining abilities.

In Hebrew, a language in which early grammatical acquisition is devoted largely to word-level morphology (Berman, 1986; Lustigman, 2012; Ravid, 1997), clause-combining is relatively straightforward. Coordination and subordination rely largely on two general-purpose conjunctions, *ve-* 'and' and *še-* 'that' respectively; word-order inversion is not involved; and same-subject ellipsis is permitted but not obligatory in both conjoined and complement clauses. These different options are illustrated in (1) and (2) below from stories based on a picture-book told by mothers to their three-year-old daughters (from Berman, 1990).

(1) Excerpts from mother's story to Noa, aged 3;0<sup>2</sup>

- a. *ha-yéled nivhal me'od mimena ve-hu nafal al ha-ricpa*  
'The-boy got a big fright from it and he fell on the ground'
- b. *ve-ba-láyla ba-láyla, še-af exad lo sam lev,*  
'And at-night at-night, that = when nobody was looking,  
*ha-cfardéa kama ve-0 yač'a mi-tox ha-cincénet*  
'the frog got up and-0 went out of the jar'...

2. The following conventions are used in transcribing Hebrew data: (a) Broad phonemic transcription represents current Hebrew pronunciation; (b) word-stress is final unless marked otherwise by an accent aigu; (c) hyphens indicate grammatical morphemes attached prefixally to the following word in Hebrew orthography, including the prepositions meaning 'to', 'in', the definite article, and the conjunctions 'and', 'that', 'when' (Ravid, 2012).

(2) Excerpts from mother's story to Sharon, aged 3;2

- a. *hu lakax ota me-ha-bica ve-0 hevi ota ha-bayta*  
'He took it from the swamp and-0 brought it home'
- b. *ve ha-yéled omer la-kélev "shhh", še-0 lo yafria*  
'And the boy says to the dog 'shhh', that-0 shouldn't disturb'

The study reported below, while limited in scope and exploratory in approach, aims to make an original contribution to research on the acquisition of complex syntax from two intersecting perspectives. First, early clause-combining has not been the subject of detailed study in Hebrew nor, second, has the role of supportive contexts provided by parental input been widely considered in research on the emergence of complex syntax in other languages.<sup>3</sup>

## 2. The study

The corpus and coding categories used in the study are described below.

### 2.1 Data-Base

The study is based on weekly audio-recordings of three typically developing girls, between ages 1;4 to 3;6, in interaction with their university-educated native Hebrew-speaking parents.<sup>4</sup> Smadar, the youngest of three daughters, was recorded one hour weekly between ages 1;4 to 1;7 and once or twice a month between ages 1;8 to 2;4 with a hiatus between age 1;9 to 1;10 due to illness; Lior, a firstborn child with a baby brother, was recorded one hour weekly from 1;5 to 2;9 and then once or twice a month until age 3;6; Hagar, an only child, was recorded one hour weekly from age 1;7 to 3;3.

Recordings were transcribed in broad phonemic script by Hebrew-speaking linguistics majors, with digitalized versions consulted when necessary.<sup>5</sup> Coding was conducted by the two authors working together, with disagreements resolved by a research team directed by Dr. Bracha Nir. Interpretation of children's utterances

---

3. An exception is the study of Huttenlocher, Vasilyeva, Cymerman, & Levine (2002), who – like other researchers they cite in this connection – consider the role of parental or teacher input in complex syntax of older children, from age 4 years.

4. The grammatical development of these children has previously been studied in relation to: subjectless constructions (Berman, 1990); use of the connective 'and' (Berman, 1996); argument ellipsis (Uziel-Karl & Berman, 2000); and verb inflections (Lustigman, 2013).

5. Thanks to Brian MacWhinney and Aviad Albert for making digitalization possible.

took into account situational non-linguistic context, preceding adult utterances, and the child's subsequent utterances (Bloom, 1970). The mothers who did the recording were explicitly instructed to provide "rich semantic context" (Brown, 1973), so that they consistently offered expansions and/or clarifications of children's utterances which were not clearly interpretable phonetically or contextually. Transcribers also consulted with mothers to resolve remaining difficulties.

The data-base selected for analysis began with the first occurrences of children's clause-combining (see Section 2.2 for criteria) – hence excluding approximately six months of previous recordings – and continued to the last available recording for each child. Table 1 gives the breakdown of transcripts analyzed, where number of utterances equals number of textlines, as specified by intonation following CHAT conventions (MacWhinney, 2000). Textlines thus include not only clauses but single words, non-clausal phrases, and multi-clause strings.

As evident from Table 1, the materials recorded for the three children differed in size and duration. Smadar's sample was far smaller than of the other two children, yet she provided particularly rich data for CC (see Table 2): *Inter alia*, Smadar's first CC utterances were identified as early as age 1;8, but from age 2;0 for Lior and Hagar. Importantly, all three *mothers* used complex syntax extensively in addressing their daughters well before then. For example, when the girls were 18 to 19 months old: Smadar's mother tells her daughter, playing with blocks: *at mex-apéset xatixa še-tiheye lo po, aval at oméret "sim lo po"* 'You-are looking (for) a piece that will fit it here, but you say "put it here"'; Lior's mother comments that: *ha-yom ani lo mevina ma at oméret* 'Today I (do) not understand what you are-saying'; and Hagar's mother responds to her question *ma ze?* 'what's that?' by *ze bul še-madbikim al mixtav* 'It's (a) stamp that (people) stick on (a) letter'. The figures in the last line of Table 1 show that the three children received different proportions of input utterances: Smadar talked more than her mother, Lior produced about 20% fewer utterances than her mother, and Hagar about 33% less than her parents.

**Table 1.** Breakdown of transcripts analyzed for CC for three children, by age of initial CC occurrence.

Coded Corpus \ Child	Lior	Hagar	Smadar
Age of first CC	2;0	2;0	1;8
No. of transcripts	66	58	17
Child utterances	12,496	8,152	3,895
Adult utterances	15,829	11,860	3,408

Note: There were gaps in recordings or transcriptions between 3;1 to 3;5 for Lior and between 3;0 to 3;2 for Hagar.

## 2.2 Categories of analysis

The basic unit of analysis was the *clause*, defined as “a unified predicate” expressing a single activity, event, or state (Berman & Slobin, 1994: 660-662), and specified for present purposes as (1) a string containing at least one verb or (2) a verbless present tense copular construction. The process of Clause Combining was defined as applying to all and only clauses that are syntactically and/or thematically *contingent on* another (typically although not necessarily adjacent) predication, where contingency means that a given clause is dependent on another for grammatical well-formedness and/or for semantic interpretation.

The analysis included all instances of each child’s CC, except for rote-learned strings lifted from storybooks, television programs, songs, games, and other routines. Relevant instances were divided into those supported by caretaker-child interaction (Section 2.2.1) and child-initiated CC constructions – marked by juvenile or non-explicit means (2.2.2) or fully marked (2.2.3). Inter-clausal syntactic categories were coded on the basis of prior work on complex syntax in Hebrew (Berman & Lustigman, 2012; Berman & Nir, 2009, 2012).<sup>6</sup> Coding of input-supported and non-canonic CC was based on innovative criteria devised specifically for this study.

### 2.2.1 *Supportive contexts*

We identified three types of *supportive contexts*, ranked by amount of grammatical structuring afforded by caretaker input – Affirmation, Scaffolding, and Co-construction, as illustrated in (3) to (5). Input utterance are bolded, clause boundaries marked by square brackets], explanatory comments in curly brackets {...}.

- i. *Affirmation*: encouraging, but quite minimal input comments (like English ‘really,’ ‘is that so?’), and partial reiterations, as in (3).

(3) Lior, aged 2;5.0<sup>7</sup>

CHI: *ima, ani orid et ze.*

‘Mommy, I’ll take it down’ {her doll}

MOT: *toridi!*

‘You-take (it) down!’

CHI: *še-yiheye lax tinok*

‘(So) that-you’ll have (a) baby (doll)’

6. For example, so-called procedural “discourse markers” that lack a conventional syntactic function and/or semantic content linking two clauses were **not** counted as indicating CC contingencies (Nir & Berman, 2010b).

7. Children’s utterances are presented in the target forms, rather than replicating juvenile or partial articulations. Examples are translated freely into English, except where morpheme-by-morpheme glossing is needed to interpret CC constructions.

- ii. *Scaffolding* provides more directly structural supportive input, in the form of prompts intended to instruct the child as well as to keep the conversation moving. Thus, in (4), the adult query ‘and then?’ scaffolds the child’s next clause as contingent on what she (the child) said previously on the topic of the pacifier.

(4) Smadar 1;11:13

CHI: *ne’elam ha-moceč šel ha-dod!*

‘Disappeared the-pacifier of the-uncle = the man’s pacifier disappeared’

MOT: *ve-az?*

‘And then?’

CHI: *ve-az hem lo mac’u et ha-moceč shel ha-yeled*

‘And-then they didn’t find the pacifier of the (little) boy’

- iii. *Interlocutor-Child Co-Constructions* are cases where the child’s output clause is *structurally* contingent on input from the adult – often in the form of a question or other type of adjacency pair which *together* serve a clause-combining function, as in (5).

(5) Hagar, aged 2;8:7

FAT: *ze me’od lo yafe. az lama ciyart?*

‘It’s very not pretty. So why (did) you-draw (it)?’

CHI: *ki ani roca rak lasim et ze.*

‘Because I want just to-put it (there)?’

All three – affirmation, scaffolding, and interlocutor co-construction – also play “supportive” roles in adult interchanges, too, in moving a conversation along.

### 2.2.2 Autonomous, non-explicitly marked CC

Four types of child-initiated “autonomous” CC strings were identified as not marked by explicit connectives: Two juvenile (omission or ambiguous general-purpose use of a conjunction) and two grammatical but unmarked (imperative directives and juxtaposed clauses).

- i. *Omission* of an obligatory connective – generally of the multi-functional subordinator *še-* ‘that’, required in complements as in (6) and relatives (7), and also in most adverbial clauses. The English gloss in (6) sounds acceptable, but the complement clause *hitkavanti* ‘I meant’ is ungrammatical since it lacks the obligatory complementizer *še-* ‘that’, while (7) lacks a required temporal or conditional connective.

(6) Smadar, aged 1;11.18

CHI: *hine, hine beyt-sefer!*

‘here, here is (the) school!’

MOT: *naxon.*

‘right.’

CHI: *ze ma] 0 hitkavanti.]*

‘that’s what] 0 I-meant’]

- (7) Lior, aged 2;7.25

CHI: *0 ani yihye gdola] ani lo ohevvet mic-anavim*

‘0 I will-be grown-up] I (do) not like grape-juice.]

- ii. *General Purpose Markings* are cases where še- is used so opaquely that its relation to the antecedent clause is multiply ambiguous between a Coordinate, Relative, or Adverbial Clause (if the latter, Time, Reason, or Purpose) – as in (8), where the child wants a towel to dry her wet hair.

- (8) Lior, aged 2;3.0

CHI: *gam ze bišvili*

‘That (is) also for me’

MOT: *ma amart, xamuda?*

‘What (did) you-say, honey?’

CHI: *še-lo yihye li kar, še-lo na'im li ba-rosh.*

‘(so) that I shouldn’t be cold, **that** my head isn’t comfy’

Two types of unmarked clause-combining – also found in adult conversation – were distinguished: Imperative-mood directives (iii) and thematically related juxtaposed clauses (iv).

- iii. “*Directive introducers*” (called “serial-verb” type constructions by Vasilyeva, et al, 2008) function pragmatically as directives to action, mainly with an imperative verb meaning ‘come’ (masculine *bo*, feminine *bői*) occasionally also *lex(i)* ‘go’, followed by a verb in irrealis mood (Infinitive or Future). These function as hortatives, analogously to English ‘let’s’, in both adult input and child output. For example, Hagar’s father says to her [aged 2;1:8] *boi nasim ba-makom* ‘come-IMP will-1STPL [= let’s] put (the blocks) back in place’.

- iv. *Juxtaposed (asyndetic) clauses* are cases where two or more adjacent clauses are closely related thematically and/or pragmatically, even though without an overt connector between them – as in adult conversational interaction, too (Mann & Thompson, 1986; Nir, 2008). Juxtaposed CC is illustrated in (9) and (10).

- (9) Hagar 2;11;11

CHI: *ani roca múzika.] ani adlik múzika.]*

‘I want music. I’ll turn-on music {= the radio}

- (10) Lior, aged 2;1.14

CHI: *aba halax la-rofe,] koev lo ha-gav aba.*]  
 ‘Daddy went to the doctor,] his back hurts, Daddy.]’

In (9) and (10), the second clause, while lexically unmarked as such, expresses a clear thematic contingency of cause or motivation with respect to the preceding clause, hence constituting an instance of unmarked CC.

### **2.2.3 Autonomous explicitly marked CC constructions**

Children’s CC constructions were coded as canonical if and only if they met four conditions: they were (a) self-initiated, (b) marked by coordinating or subordinating conjunctions, (c) occurred within a single turn, and (d) were thematically related. Such occurrences were coded by accepted syntactic categories such as Coordinate, Complement, Adverbial, and Relative Clauses in Hebrew and other languages (Berman & Nir, 2009; Nir, 2008; Nir & Berman, 2010a). Sequential stringing of clauses, like those bolded in (11), was added as a sub-category of coordination, as a common CC device in young children’s narration (Berman, 1990; Berman & Neeman, 1994).

- (11) Lior, aged 2;7.9

CHI: *téxef savta tavo,] ve-, ve-axarkax, ve-axarkax telxi la-šuk]*  
*ve-al tiša’ari otanu po.*]  
 ‘soon Granny will-come] and afterwards, and afterwards you’ll-go to-the-market] and don’t stay (*sic*) us here.]’

## **3. Results**

Total occurrence of CC output for each child and the proportion of these supported by the three types of adult input specified in Section 2.2.1 are presented in Table 2.

Table 2 shows that, despite the disparity in sample size of the three girls (Table 1), in the period analyzed for their “emergent clause combining”, a similar

**Table 2.** Total instances and proportion of input-supported CC, by child.

CC Breakdown \ Child	Lior	Hagar	Smadar
Total CC instances	444	424	429
Supported CC instances	203 (45%)	46 (10.8%)	125 (29%)

number of over 400 instances fell into the class of child output that we identified as analyzable for CC – defined as strings of two or more clauses linked together in a relation of contingency. These breakdowns formed the basis for analysis *across the data-base*, since concern was with the nature and development of early CC observed for each child, rather than the proportion of CC occurrences out of her total speech output. Importantly, Table 2 shows that the overall proportion of supportive contexts identified for each child differs considerably, from 45% of Lior's and 29% of Smadar's CC output, down to only 10.8% for Hagar.

Below, findings are presented first for children's self-initiated or autonomous CC (Section 3.1) followed by analysis of different kinds of supportive contexts (3.2).

### 3.1 From juvenile to canonic CC

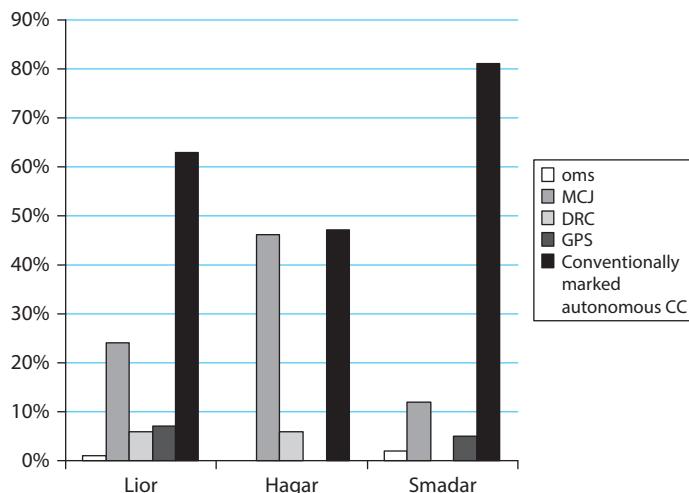
In order to trace the transition from juvenile to canonically grammatical marking of inter-clause relations, we first analyzed the children's self-initiated CC outputs for cases where these were not marked by explicit lexico-syntactic connectives. Table 3 shows the number of contingent clauses occurring in autonomous (i.e., non-supported) CC constructions and the proportion of these that were not explicitly marked due to: omission of a required conjunction (OMS), ambiguous use of the general-purpose subordinating conjunction *še-* 'that' (GPS), imperative-mood directives (DRC), and unmarked thematically-related juxtaposed clauses (MCJ) – as defined in the preceding section.

Table 3 shows that, although Hagar produced nearly twice as many autonomous, self-initiated CC strings as the other two children, she uses far more juvenile or non-explicit marking of CC – over half the time, compared with 19% to 37.3% for Lior and Smadar. Figure 1 shows the breakdown of these four types of non-explicitly marked contingent clauses, compared with canonic CC with contingent clauses marked by appropriate, explicit lexico-grammatical connectives.

Figure 1 shows that the bulk of the children's self-initiated CC constructions are grammatically marked: Juxtaposed strings account for most non-marked instances, omissions are relatively rare, while Smadar and Lior but not Hagar make

**Table 3.** Proportion of non-explicitly marked (omitted, ambiguous, directive, juxtaposed) clauses participating in CC out of total autonomous contingent clauses, by child.

Child	Lior	Hagar	Smadar
Total autonomous contingent clauses	370	623	372
Juvenile (OMS, GPS) and non-marked (DRC, MCJ) contingent clauses	138 (37.3%)	330 (52.9%)	71 (19.08%)

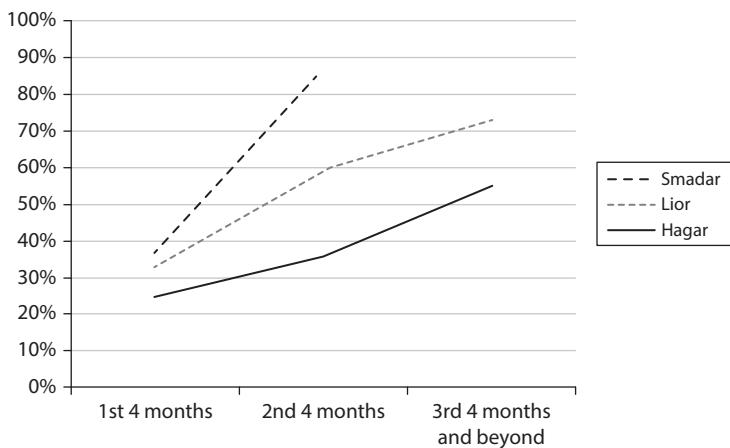


**Figure 1.** Breakdown in percentages of juvenile (OMS = omissions and GPS = general purpose subordinator), syntactically unmarked (DRC = directive imperatives and MCJ = juxtaposed) and explicitly marked contingent clauses produced autonomously, by child.

some use of the Hebrew-specific juvenile device of a vague general-purpose subordinating *še-*. Hagar further differs from the other two girls in more use of the imperative directive 'come!' followed by an irrealis form verb.

Developmentally, each child revealed an age-related increase in the proportion of autonomously produced canonical contingent clauses. Figure 2 compares the three girls across two four-month intervals in this respect, beginning with the earliest CC strings for each child (age 1;8 for Smadar and 2;0 for Lior and Hagar) and proceeding to the end of the documented period, lasting eight months for Smadar (2;4) and over a longer stretch for Lior (to 3;6) and Hagar (to 3;3).

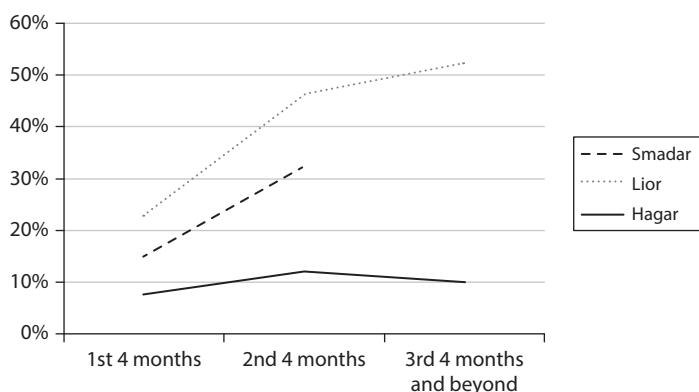
Figure 2 shows different developmental trajectories in proportion of explicitly marked canonic contingent clauses out of total CCs for each child: Lior shows a consistently gradual development, from one-third to 60% at age 2;4 and up to over 70% from one 4-month interval to the next; Smadar starts at 40% and reaches a ceiling of 87% by the end of her second 4-month interval; Hagar starts at a low of under 30%, arriving at 55% explicitly marked contingent clauses by her third interval. That is, Smadar progresses very rapidly, Lior advances more slowly but consistently, while Hagar fails to achieve near ceiling more than a year after her first CC productions.



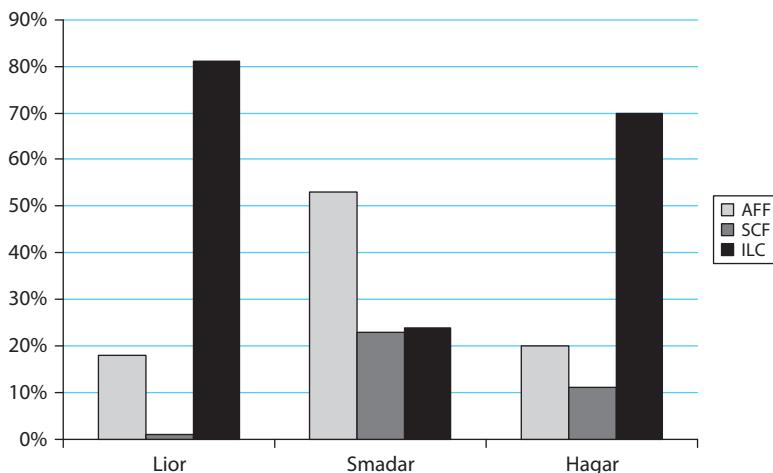
**Figure 2.** Increase in proportion of explicitly marked self-initiated contingent clauses by 4-month intervals, by child – from age 1;8 to 2;4 for Smadar, 2;0-3;6 for Lior, and 2;0 to 3;3 for Hagar.

### 3.2 Supportive contexts in early clause-combining

Three kinds of supportive contexts were identified (Section 2.2.1): affirmation-based, scaffolded, and interlocutor-triggered. The total amount of contextual support differed considerably from child to child, when calculated out of total CC instances, from 45% for Lior to 29% for Smadar, and only 10.8% for Hagar (see Table 2). Figure 3 shows the relative proportion out of total CC instances of the three types of interlocutor supports combined.



**Figure 3.** Proportion of context-supported CC out of total CC instances by 4-month intervals, by child – from age 1;8 to 2;4 for Smadar, 2;0-3;6 for Lior, and 2;0 to 3;3 for Hagar.



**Figure 4.** Breakdown (in percentages) of three types of CC supportive contexts (aff = affirmation, scf = scaffolding, and ilc = interlocutor co-construction), by child.

A marked rise in proportion of supported CC is apparent for both Smadar and Lior during their second four-month interval: from under one-fifth ( $12/80 = 15\%$ ) to one-third ( $113/349 = 32.4\%$ ) for Smadar and from  $23\% 18/78$  to  $46.4\% 51/110$  increasing to  $52.3\% 134/256$  in the third interval for Lior. In contrast, the proportion of supported CC for Hagar starts low at  $7.7\% (=5/65)$  and rises to only  $12\% (=14/118)$  in her second four-month interval, then down again to  $10\% (27/241)$ . Not only did the number of supportive contexts for CC differ for each child, but so did the type of support each received, as shown in Figure 4.

Recall that although all three children produced a similar number of overall around 400 CC strings in the period analyzed, the number of these that were supported hence non-autonomous, differed considerably by child (Table 2) Lior's mother provided the most supportive contexts overall, and 80% of these (over 200) were structurally-based interlocutor triggered inputs; Smadar's mother provided relatively more affirmations, with a similar number (around 20% each) of the remaining 125 instances of her supportive contexts in the form of scaffoldings and structurally based input. The few supports provided to Hagar's CC output (less than 50 in all) were mainly structurally based with little in the way of affirmations or scaffolding.

As for CC *functions*, all three girls from early on produced each of the different types of major CC constructions (coordinating, adverbial, complement, and relative clauses), they did so more or less concurrently, and in both supported and autonomous contexts. Moreover, coordinating and adverbial (mainly reason)

clauses were overall most frequent, again for all three children, and again, in both supported and autonomous contexts of CC production.

#### 4. Discussion

Two developmental trends emerged as *shared* in the children's clause-combining abilities, despite disparities in total size and duration of their samples. First, as expected from studies on early complex syntax in other languages, they all demonstrated progress in appropriate lexico-syntactic marking of their autonomously produced, self-initiated CC constructions. Initially, clause-combining involved little explicitly marked inter-clausal connectivity, rising markedly in the second 4-month intervals for Smadar and Lior, and in the third for Hagar (Figure 2). This shift from non-marking to explicit use of appropriate lexico-grammatical connectives is in line with Clark's (2009: 230) observation that, "as children's utterances become more complex, their references to two or more events within a single utterance become more clearly discernible, even when they do not yet produce adult constructions and omit all connectives for linking one clause to another".

Second, all three girls relied, although in different proportions, on the four different types of unconventionally marked CC constructions that we identified – juvenile omissions and vague use of the Hebrew subordinating 'that' along with acceptable, asyndetic use of imperative directives and juxtaposed main clauses (Table 3, Figure 1) Such devices might in fact function as *precursors* to the canonical lexico-syntactic marking of clause-combining, particularly since the first two are rare if not outright impossible in adult speech, and all four usages decrease with age for the three children (Figure 2). As such, they are in line with the more general tendency of young children to initially rely on non-marked, even if adult-like, options in acquisition of grammar. This is seen, for example, in the early preference of Hebrew-speaking toddlers for inflectionally neutral non-finite stem or participial forms of verbs (Lustigman, 2012) and singular masculine forms of nouns (Dromi & Berman, 1982). Besides, our impression is that adults' use of the two conventional devices (directives marking hortative mood and unmarked juxtaposition of thematically related clauses) in our data-base as well as in adult-adult conversational interchanges constitutes a smaller proportion of CC constructions than in the speech output of young children – an issue that requires further analysis of a larger and more representative data-base.

As against the shared trends revealed by our analysis, individual patterns emerged for each child along several dimensions. For example, each showed a different developmental trajectory in the relative proportion of explicitly marked

autonomous CC outputs (Figure 2), so that although, on the surface, Hagar and Lior appear quite alike, both starting CC production at around age 24 months compared with the more precocious Smadar. In fact Lior and Smadar show quite similar developmental trajectories from two points of view. First, their route to a relatively large proportion of explicitly marked autonomous CC constructions was shorter than that of Hagar who, moreover, showed a far flatter developmental curve in this respect. Second, Smadar and Lior across-the-board show far greater reliance in their CC productions on supportive contexts provided by adult caretakers than Hagar (Figure 3). Overall, Lior's CC was interactively supported around half the time, Smadar's in around one-third of the cases, less so, and Hagar's almost not at all. Moreover, this quantitative disparity in amount of CC supportive contexts is reflected qualitatively in the *types* of non-canonic clause-combining devices favored by each child (Figure 1). For example, Smadar and Lior, but not Hagar, made some use of a vague, general purpose subordinating *še-*, whereas Hagar relied relatively more on the imperative directive meaning 'come!' plus an unrealis-form verb. In fact, Hagar showed far more avoidance of lexico-syntactic marking of inter-clausal relations, even though in terms of sheer amount of CC, she stringed clauses together no less than the other two girls. This is consistent with the finding that in the CC constructions she produced in the few supportive contexts available to her (Table 2), Hagar used mainly reason adverbials introduced by *ki* 'cos = because' in response to parental *why* questions. She rarely if ever produced constructions introduced by subordinating *še-*, again, in contrast to the two other girls.

These differences can in part be attributed to "variations in individual experiences ... as one reason why children can follow different developmental paths as they acquire a first language" (Arnon & Clark, 2011:1). Thus, beyond the marked quantitative differences in amount and type of supportive contexts provided by the adults (Figure 4), clear discrepancies were detected in *parental styles* of interaction with their toddlers: Hagar's mother (and occasionally her father) were generally very talkative and communicatively invasive, giving her little or no chance to practice clause-combining in interaction. (Much of her CC activity took place when she was on her own, playing with dolls or looking at picturebooks). In contrast, Lior's mother was a very cooperative interlocutor-caretaker (providing the highest proportion of adult-child CC co-construction) and Smadar's mother was very encouraging (giving relatively far more affirmatory and scaffolding input). These different patterns for Hagar compared with the other two girls in both the development of CC constructions and the nature of parental support are a suggestive and, as far as we know, novel finding, since prior studies that consider the role of adult input in developing complex syntax have dealt either with specific facets of clause-combining such as connectives (Braunwald, 1985), with particular types

of complex sentences (Bloom & Capatides, 1987) or with older children (Huttenlocher, Vasilyeva, Cymerman, & Levine, 2002; Vasilyeva et al, 2008). Further study with more children and/or additional linguistic domains is needed to support the impression that emerges from the present study for the importance of supportive contexts in co-construction of CC, as providing children with a training ground for when and how to mark grammatical relations between clauses.

Other desiderata to flesh out in this exploratory study include the need for *a priori* more carefully controlled samples in terms of size, amount, and duration of the corpus to be analyzed, although longitudinal recordings of several children that extend over more than a few months are known to encounter numerous practical obstacles. It would also be helpful to have a (semi-)automatic procedure for evaluating children's relative level of grammatical development analogous to MLU, which is adjusted to a highly morphologically fused language like Hebrew (Dromi & Berman, 1982). Note, however, that the relative differences in quality and rate of acquisition of CC of the three girls is reflected in *other* domains of their language acquisition as well, as shown in earlier studies cited in note 4. Thus, in acquisition of morphological inflection, verb-argument structure and ellipsis, as well as use of the connective 'and', these children reveal developmental paths similar to the area of CC: Smadar – an extremely rapid and precocious learner, Lior – a fairly average learner with a clear and consistent developmental pattern, and Hagar – showing a rather more prolonged and restricted process of acquisition. These observations suggest that the domain of CC complex syntax could usefully be applied as a developmental yardstick for level of grammatical development beyond age three, analogous to use of MLU in early stages of acquisition.

Finally, to return to the assumptions delineated in the introduction, examination of the recordings prior to the start of analysis (at age 1;8 for Smadar and 2;0 for Lior and Hagar) revealed that, indeed, simple-clause structure was well-established by the onset of CC production for all three girls. In Hebrew, this meant a combination of at least four clausal constituents and rich use of inflectional markings for number, gender, person, and tense. On the other hand, again, as expected, the task of combining together two and especially several clauses by appropriate lexico-syntactic means proved to present a challenge to these young children, whose communicative needs might be met just as well by relying on simple clauses – even in a language like Hebrew, with its relatively straightforward processes of clause combining. This is clearly shown by the Hagar's prolonged path in mastering even non-conventionally marked CC constructions. The other two girls, who produced far more grammatical CC constructions at an earlier developmental phase, were aided by a marked increase in supportive input in the critical intermediate second phase of their CC construction, suggesting that, indeed, adult contextual support may play an important role in fostering early clause-combining

abilities. A fourth assumption that was well supported by our findings was that throughout the process, two of the three girls proved able to produce both interactively supported and independently self-initiated “autonomous” CC constructions. That is, in development of CC and, presumably, in other domains of language acquisition, supportive interactive contexts are a critical facet of conversational discourse for toddlers and adults alike.

## References

- Armon-Lotem, S. (2004). The acquisition of subordination: From preconjunctionals to later use. In D.D. Ravid & H. Bat-Zeev Shyldkrot (Eds.), *Perspectives on language and language development: Essays in honor of Ruth A. Berman* (pp. 191–202). Dordrecht: Kluwer.
- Arnon, I. (2011). Relative clause acquisition in Hebrew and the learning of constructions. In E. Kidd (Ed.), *The acquisition of relative clauses: Processing, typology, and function* (pp. 81–106). Amsterdam: John Benjamins.
- Arnon, I., & Clark, E.V. (2011). Introduction. In I. Arnon & E.V. Clark (Eds.), *Experience, variation, and generalization: Learning a first language* (pp. 1–14). Amsterdam: John Benjamins. DOI: 10.1075/tilar.7
- Berman, R.A. (1986). A step-by-step model of language learning. In I. Levin (Ed.), *Stage and structure: Re-opening the debate* (pp. 191–241). Norwood, NJ: Ablex.
- Berman, R.A. (1990). Acquiring an (S)VO language: Subjectless sentences in children's Hebrew. *Linguistics*, 28, 1135–1166. DOI: 10.1515/ling.1990.28.6.1135
- Berman, R.A. (1996). Form and function in developing narrative abilities: The case of 'and'. In D.I. Slobin, J. Gerhardt, A. Kyrtzis, & J. Guo (Eds.), *Social interaction, context, and language* (pp. 243–268). Mahwah, NJ: Lawrence Erlbaum Associates.
- Berman, R.A. (1997). Theory and research in the acquisition of Hebrew as a first language. In Y. Shimron (Ed.), *Studies in the psychology of language* (pp. 37–69). Jerusalem: Magnes.
- Berman, R.A. (2009). Beyond the sentence: Language development in narrative contexts. In E. Bavin (Ed.), *Handbook of child language* (pp. 354–375). Cambridge: CUP.
- Berman, R.A., & Lustigman, L. (2012). HARSP: A developmental language profile for Hebrew. In M.J. Ball, D. Crystal, & P. Fletcher (Eds.), *Assessing grammar: The languages of LARSP* (pp. 43–76). Bristol: Multilingual Matters.
- Berman, R.A., & Neeman, Y. (1994). Development of linguistic forms: Hebrew. In R.A. Berman & D.I. Slobin (Eds.), *Relating events in narrative: A crosslinguistic developmental study* (pp. 285–328). Hillsdale NJ: Lawrence Erlbaum Associates.
- Berman, R.A., & Nir, B. (2009). Clause-packaging in narratives: A crosslinguistic developmental study. In J. Guo, E. Lieven, S. Ervin-Tripp, N. Budwig, S. Özçalışkan, & K. Nakamura (Eds.), *Cross-linguistic approaches to the psychology of language: Research in the tradition of Dan I. Slobin* (pp. 149–162). Mahwah, NJ: Lawrence Erlbaum Associates.
- Berman, R.A., & Nir, B. (2012). Clause-combining in Hebrew from early childhood across adolescence. Interim Scientific Reports to Israel Science Foundation [Grant No. 190/10], Jerusalem, Israel.
- Berman, R.A., & Slobin, D.I. (1994). *Relating events in narrative: A crosslinguistic developmental study*. Hillsdale, NJ: Lawrence Erlbaum Associates.

- Bloom, L. (1970). *Language development: Form and function in emerging grammars*. Cambridge, MA: The MIT Press.
- Bloom, L., & Capatides, J.B. (1987). Sources of meaning in the acquisition of complex syntax: The sample case of causality. *Journal of Experimental Child Psychology*, 43, 112–128. DOI: 10.1016/0022-0965(87)90054-3
- Bowerman, M. (1979). The acquisition of complex sentences. In P. Fletcher & M. Garman (Eds.), *Language acquisition, 1st ed.* (pp. 283–305). Cambridge: CUP.
- Bowerman, M., & Perdue, C. (Eds.). (1990). *The structure of the simple clause in language acquisition* [Special Issue]. *Linguistics*, 28. DOI: 10.1515/ling.1990.28.6.1253
- Braunwald, S.R. (1985). The development of connectives. *Journal of Pragmatics* 9, 513–525. DOI: 10.1016/0378-2166(85)90019-0
- Brown, R. (1973). *A first language: The early stages*. Cambridge, MA: Harvard University Press. DOI: 10.4159/harvard.9780674732469
- Bybee, J., & Noonan, M. (Eds.) (2001). *Complex sentences in grammar and discourse: Essays in honor of Sandra A. Thompson*. Amsterdam: John Benjamins.
- Clark, E.V. (2009). *First language acquisition, 2nd ed.* Cambridge: CUP.
- Clark, E.V., & de Marneffe, M.-C. (2012). Constructing verb paradigms in French: Adult construals and emerging grammatical contrasts. *Morphology*, 22, 89–120.
- De Villiers, J. (1995). Empty categories and complex sentences: The case of wh-questions. In P. Fletcher & B. MacWhinney (Eds.), *The handbook of child language* (pp. 508–540). Oxford: Blackwell.
- Diessel, H. (2004). *The acquisition of complex sentences*. Cambridge: CUP.
- Dromi, E., & Berman, R.A. (1982). A morphemic measure of early language development. *Journal of Child Language*, 9, 403–424. DOI: 10.1017/S0305000900004785
- Dromi, E., & Berman, R.A. (1986). Language-general and language-specific in developing syntax. *Journal of Child Language*, 14, 371–387.
- Givón, T. (2009). *The genesis of syntactic complexity*. Amsterdam: John Benjamins.
- Haiman, J., & Thompson, S.A. (Eds.). (1988). *Clause combining in grammar and discourse*. Amsterdam: John Benjamins. DOI: 10.1075/tsl.18
- Halliday, M.A.K. (1989). *Spoken and written language*. Oxford: OUP.
- Hickmann, M. (2003). *Children's discourse: Person, space, and time across languages*. Cambridge: CUP.
- Huttenlocher, J., Vasilyeva, M., Cymerman, E., & Levine, S. (2002). Language input and child syntax. *Cognitive Psychology*, 45, 337–374. DOI: 10.1016/S0010-0285(02)00500-5
- Karmiloff-Smith, A. (1986). Some fundamental aspects of language development after age five. In P. Fletcher & M. Garman (Eds.), *Language acquisition, 2nd ed.* (pp. 455–474). Cambridge: CUP. DOI: 10.1017/CBO9780511620683.026
- Limber, J. (1973). The genesis of complex sentences. In T. Moore (Ed.), *Cognitive development and the acquisition of language* (pp. 169–186). New York, NY: Academic Press.
- Lust, B., Foley, C., & Dye, C. (2009). The first language acquisition of complex sentences. In E. Bavin, (Ed.), *Handbook of child language* (pp. 463–505). Cambridge: CUP.
- Lustigman, L. (2012). Non-finiteness in early Hebrew verbs. *Brill's Annual of Afroasiatic Languages and Linguistics (BAALL)*, 4, 213–231.
- Lustigman, L. (2013). Developing structural specification: Productivity in early Hebrew verb usage. *First Language*, 33, 47–67. DOI: 10.1177/0142723711426828
- MacWhinney, B. (2000). *The CHILDES project: Tools for analyzing talk, 3rd ed., Vol. 1: The format and programs*. Mahwah NJ: Lawrence Erlbaum Associates.

- Mann, W.C., & Thompson, S.A. (1986). Relational propositions in discourse. *Discourse Processes*, 8, 57–90. DOI: 10.1080/01638538609544632
- Nir, B. (2008). *Clause packages as constructions in developing narrative discourse*. Unpublished doctoral dissertation. Tel Aviv University.
- Nir, B., & Berman, R.A. (2010a). Complex syntax as a window on contrastive rhetoric. *Journal of Pragmatics*, 42, 744–765. DOI: 10.1016/j.pragma.2009.07.006
- Nir, B., & Berman, R.A. (2010b). Parts of speech as constructions: The case of Hebrew “adverbs”. *Constructions and Frames*, 2, 242–274. DOI: 10.1075/cf.2.2.05nir
- Ravid, D. (1997). Morphological development a duo: Pre- and proto-morphology in the language of Hebrew-speaking twins. *Papers and Studies in Contrastive Linguistics*, 33, 79–102.
- Ravid, D. (2012). *Spelling morphology: The psycholinguistics of Hebrew spelling*. Dordrecht: Springer. DOI: 10.1007/978-1-4419-0588-8
- Reilly, J.S. (1986). The acquisition of temporals and conditions. In E.C. Traugott, A. ter Meulen, J.S. Reilly, & C.A. Ferguson (Eds.), *On conditionals* (pp. 309–331). Cambridge: CUP. DOI: 10.1017/CBO9780511753466.017
- Scollon, R. (1976). *Conversations with a one-year-old: A case study of the developmental foundations of syntax*. Honolulu, HI: University Press of Hawaii.
- Silva, M.L., Sánchez Abchi, V., & Borzone, A. (2010). Subordinated clause usage and assessment of syntactic maturity. *Journal of Writing Research*, 2, 47–64.
- Uziel-Karl, S., & Berman, R.A. (2000). Where's ellipsis: Whether and why there are missing arguments in Hebrew child language. *Linguistics*, 38, 457–482. DOI: 10.1515/ling.38.3.457
- Vasilyeva, M., Waterfall, H., & Huttenlocher, J. (2008). Emergence of syntax: Commonalities and differences across children. *Developmental Science*, 11, 84–97. DOI: 10.1111/j.1467-7687.2007.00656.x
- Veneziano, E. (1999). Early lexical, morphological and syntactic development in French: Some complex relations. *International Journal of Bilingualism*, 3, 183–217. DOI: 10.1177/1367006990030020501



PART 4

**Interactional effects on language structure  
and use (three papers)**



# Analytic and holistic processing in the development of constructions

Joan Bybee  
University of New Mexico

A corpus analysis over several decades of American English shows the development of the minor construction from an idiom: *not have two Xs to rub together*. The sequence of developments indicates that the loss of compositionality and analyzability of the construction is heavily influenced by context of use. When the idiom is used metaphorically for intellectual impoverishment, indicators of loss of compositionality begin to arise.

**Keywords:** speech processing, constructions, compositionality, analyzability, usage-based

## Introduction

The development and use of constructions requires an interesting interplay between analytic and holistic processing. The construction as a whole has meaning due to the tendency to assign meaning to the highest level of chunked material (Ellis, 1996) but the words within the construction maintain to some degree their network connections to the same words in other contexts, supplying a degree of analyzability and compositionality (Beckner & Bybee, 2009; Bybee, 2010). Hay (2001) argues that holistic processing of morphologically complex derived words is more likely if they are more frequent than their base forms. That is, they are less likely to be analyzed into parts if they are frequent. This line of thinking is more difficult to apply to constructions, though it is true that for a construction to be established, it must be repeated, at least a few times. In addition to repetition, I argue that contexts of use that emphasize the holistic meaning over the analytic one lead to greater loss of analyzability.

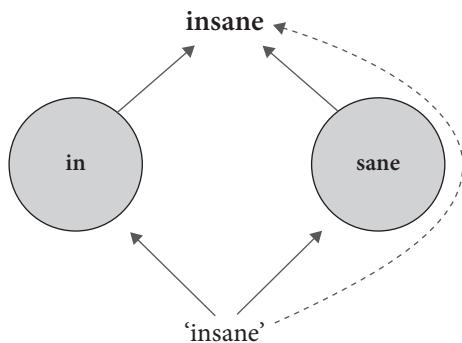
### 1. Defining analyzability and compositionality

Langacker (1987, 2008) finds it useful to distinguish analyzability of complex expressions from compositionality. Both deal with the identification of component parts of expressions and in most cases compositionality implies analyzability. Compositionality refers to the degree of predictability of the meaning of the whole from the meaning of the parts. Analyzability according to Langacker is the ‘recognition of the contribution each component makes to the composite conceptualization’ (1987: 292), even if the conceptualization is not fully predictable from the meanings of the parts, as in metaphorical expressions.

Compositionality and analyzability are both gradient notions and can be modeled with an exemplar cum network model by supposing that when a complex expression is processed, both the whole expression (if it has been experienced earlier) is activated, as well as the component parts. If the expression is fully compositional, the parts and their meanings are activated. If the expression is not, then there is partial activation of the parts (e.g. their form but not their meaning). Of course activation levels are gradient so that either the whole or the parts can have stronger or weaker activation. In a usage-based grammar, each instance of activation has an impact on the strength of connections within the network. Thus an expression used frequently and activated as a whole will lose connections to its component parts, resulting in a loss of compositionality and analyzability (Beckner & Bybee, 2009; Bybee, 2010).

There is some indication that relative frequency of the whole vs. the parts may lead to the loss of compositionality and analyzability. Hay (2001) argues for morphologically complex words that, if the derived form is more frequent than the base, are likely to be less analytic and less compositional than in cases where the base is more frequent than the derived form. Thus *insane*, which is more frequent than *sane* is less analyzable than *uncommon*, which is less frequent than *common*. Her hypothesis is that there are two routes to a complex word: one creates the word by combining the parts and the other by accessing the complex word directly, as shown in Figure 1 adapted from her article. As a word grows more frequent, its representation is strengthened and it is therefore easier to access directly. If the base from which it was originally derived is less frequent, then activation of the base is skipped in access to the complex form. Failure to access the component morphemes leads to a loss in compositionality and analyzability. In the example, the complex word *insane* becomes autonomous from related words (Bybee, 1985, 2010).

While this approach addresses the question of how analyzability is lost, it does not provide an account of how compositionality decreases. As compositionality has to do with semantic transparency, and, according to Hay, is lost when the meaning of the complex word changes, we still need an explanation for why the



**Figure 1.** Schematized dual-route model (Hay, 2001:1045). Solid lines indicate the decomposed route and broken lines indicate the direct route. Font size indicates the resting activation level: *insane* is more frequent than *sane*.

meaning changes. This problem becomes more acute when we try to apply Hay's hypothesis to the loss of analyzability and compositionality in constructions. One problem is that constructions can lose their compositionality without actually being terribly frequent (Hoffman, 2004) and without being more frequent than their component words. In fact, the verb in the particular idiomatic construction I will discuss here occurs in COCA (the Corpus of Contemporary American English) 32 times within the construction in which it is losing its compositionality, but 14,725 in the corpus as a whole. Clearly relative frequency is not the answer to loss of compositionality. Instead, we have to look to contexts of use to find the reasons for semantic change and the resulting loss of compositionality.

## 2. Holistic meaning overrides compositional meaning

It is important to remember that language use takes place in context and listeners are not just decoding words or morphemes as they come along in sequence, but are rather trying to grasp the whole message the speaker is attempting to convey. To do this, the listener must rely not just on the meanings of the words, but also on the information in the context and his or her ability to make inferences. This is true both for adults and children (Clark & Amaral, 2010). As pointed out in a landmark paper in 1979 by Clark and Clark, this fact has a major impact on semantic change as meanings taken in context become attached to the word, chunk or construction and are re-used and conventionalized. To take just one example, Clark and Clark (1979) demonstrate that new verbs created by zero-conversion from nouns require extensive context for interpretation. The interpretations that

apply in the early contexts can become the conventional meaning for the new verb. Consider the verb formed from the noun *man*. Of all the situations that could be captured by making a verb from *man* the one that is conventionalized means ‘to supply or furnish with men for support, defense or service; to be stationed at in order to defend, care for or operate’ according to the *American Heritage Dictionary* 1976. To imagine other possible meanings for *to man*, consider a jocular usage by a friend of mine, who referred to single women as not being *manned*, or being *unmanned* (a derivative of the participle). How did the specific conventional meaning of *to man* arise? This verb has been around for many centuries and most often used in a military or nautical context; it presumably arose in such a context, where its meaning was filled in by the interlocutors’ knowledge of the general and specific situation.

In analyzing the way children and adults assign meanings to words, particularly to verbs, Clark (1993: 56) proposes the ‘whole-action assumption’, the ‘speakers use words to pick out whole actions, not just parts of actions’. In a similar discussion, Ellis (1996) draws from the psychology literature to remind us that we humans tend to assign meaning to chunks of experience at the highest level possible. For instance, if we observe a person throwing a ball, a dog chasing after it and catching and bringing it back to the person who threw it, we see an activity that we chunk and label *playing fetch*. A dog who has not learned this game sees only unconnected separate activities. Our strong tendency to chunk and assign meaning at the highest level applies also to linguistic chunks, including constructions, which tend to take on holistic meaning at times overriding the meaning of the parts.

The study of the way new constructions develop over time is in its infancy, though Israel (1996) and Wilson (2009) study the way an idiom expands over time through semantic change and by increasing its schematicity. In fact, the creation of new constructions has been called ‘schematization’ by Noël (2007). In the following discussion we examine a construction that increases its schematicity but we will be less concerned with schematicity and more concerned with the way the general, holistic meaning of the construction and the contexts in which it is used interacts with the meanings of the words that constitute it – that is, the issue of how analytic and holistic meaning interact. The example to be discussed in the next two sections demonstrates how the general meaning of the idiom which expanded to become a partially filled construction gave rise to substitutions of the original words in the idiom. In a later development, expansion to new contexts via metaphor resulted in even more variation within the words of the construction as their contribution to the general meaning grows ever more tenuous.

### 3. An expanding construction: Stage 1

The 2010 film *The Social Network* is about the founding of *Facebook* by the then Harvard student Mark Zuckerman. In the film Zuckerman is portrayed as a computer whiz and a bit of a nerd. Two other Harvard students, Tyler and Cameron Winklevoss, handsome, athletic, socially competent twins, talk to Zuckerman about a web-based dating network for Harvard students. Six weeks later, Zuckerman launches what he called *Thefacebook* which is an instant success. The Camerons and their business partner Divya naturally think that he stole their idea. Here is their reaction.

- (1) CAMERON (on the phone with his father's lawyer while Tyler and Divya are present): Yeah, Divya was just reading that 650 students signed up for it on the first day.  
 TYLER: If I were a drug dealer I couldn't give free drugs to 650 people in one day.  
 DIVYA: *And this guy doesn't have three friends to rub together to make a fourth.*

The last line here is what is of interest: what does it mean not to have three friends to rub together to make a fourth? Our understanding of this statement requires understanding of a construction that indicates extreme poverty, most commonly an expression such as that found in COHA 1971: *He hasn't got two nickels to rub together*. The intended meaning is that Zuckerman is so socially isolated that it cannot be his own friends that signed up the first day. Of course, Divya's statement was the creation of the screenwriter, but it does show how manipulable what appears to be a fixed expression can be (Barlow, 2000). Just how flexible is this particular idiomatic construction? That can be determined by following its use in corpora from the 19th to the 21st century – the Time Magazine Corpus, the BNC, COHA, COCA and the OED (Davies, 2004, 2007, 2008, 2010, 2012).

Though it is difficult if not impossible to document exactly where this construction came from, it very probably derived from an expression such as *X is so poor s/he does not have two sticks to rub together to make a fire*. There are three arguments for this source: first, some contemporary speakers cite this as the source (though many are puzzled by the expression, see below); second, the expression of having two sticks to rub together to make a fire is still found in modern corpora; and third, one example of *sticks* used in the construction was found: *You won't have two sticks to rub together when my husband finds out how I've been treated here* (Soap Opera Corpus, 2006).

Given this source, the first step in the development of this construction is the establishment of the holistic meaning of the idiom *not have two sticks to rub*

*together*, as an expression of the extreme poverty of the subject. The evidence of the meaning of ‘extreme poverty’ overriding the compositional meaning is the substitution of nouns denoting coins for *sticks*. The use of coins in this expression occurs early; the first example in the OED is from 1827. The purpose clause, *to make a fire* has been lost and with it the compositionality and even the analyzability of *sticks*. The meaning of the whole motivates the use of coin words in the construction. Table 1 shows the nouns indicating coin-like entities found in the combined corpora. While this slot in the construction is schematic, it is highly constrained semantically.

However, compositionality and analyzability are not lost when holistic meaning arises as shown by variations on the construction. For instance, the compositional meaning of *two* is retained as shown by the substitution of *a couple of*, *a few*, and *two, maybe three* in the corpus examples. In addition, no *two* occurs with *tuppence* ‘two pence’ showing both the compositionality of *two* in the construction and of *tuppence* itself. The analyzability of the noun is retained as shown by the occasional appearance of adjectives to modify what is being rubbed together, as shown in example (2). Also variations on the negative marker show that it is the meaning of negative, not the particular form that expresses it that is conventionalized; *not*, *never* and *without* can be used to indicate the negative (see example (3)).

Table 1. Nouns denoting coin-like entities and their numbers of tokens in the Time Magazine, COCA, COHA, BNC and OED.

nouns (coin-like)	tokens		
nickels	11		
coins	06		
farthings	04		
pennies	03		
ha'pennies	01		
tuppence	01		
shillings	01		
sous	01		
guineas	01		
rubles	01		
cents	01	<u>less coin-like</u>	
dimes	01	pieces	01
quarters	01	beans	03
dollars	01	tesseracts	01
euros	01		
food stamps	01		total: 41

- (2) Every sport in Harlem who had two *white quarters* to rub together opened a policy house (COHA 1969)  
 ‘Every (black) man in Harlem who had two white-man’s quarters (coins) to rub together opened a policy (gambling) bank’
- (3) For a no-good gambling addict who left you *without* two nickels to rub together? (COCA 2012)

The expression is also used in the affirmative to indicate a minimal level of solvency, something just above the poverty line. This occurs with indefinite subjects, as in (4) and with modifiers such as *only* and *at least* as in example (5), which also shows that the construction can be made into a relative clause. When used with *a few* rather than *two* the expression seems to indicate a slightly higher level of prosperity, as in example (6). These uses indicate that the holistic meaning of rubbing coins together has to do with a position below or just above the poverty line is fully operative, but at the same time as the analyzability of the component parts – the negative, the numeral and the type of coin – is fully accessible. However, as we see in example (7) some speakers are puzzled by the fixed phrase *rub together*.

- (4) And anybody with two nickels to *rub together* can buy a debt portfolio and set up and start collecting the debts. (COCA 2009)
- (5) Out all night at the Argosy, drinking and gambling away the only two cents we got to *rub together*. (COCA 2010)
- (6) We get a few years under our waders and a few coins to *rub together*, and we buy a bigger boat, a club membership, a camp. (COCA 2011)
- (7) Babe said. “I mean, not to put too neat a point on it, but she’s loaded and you’ve got *two, maybe three nickels to rub together*.” And where does that expression come from? Why would anyone want to rub two nickels together? Are they supposed to make babies if you rub them together? (COCA 2010)

These facts demonstrate that analyzability and compositionality of elements are gradient when these elements are tucked into a construction with a holistic meaning within a pragmatic situation. The construction in its basic form – negative with a coin or something very similar in the schematic slot – occurs in 35 of the examples. These could be processed as chunks with holistic meaning. But the variations on that with affirmatives and other permutations indicate both analyzability and compositionality are still viable to some degree.

In an exemplar model where networks of connections are formed among words and phrases, instances of the whole construction with the meaning gleaned from the context are stored in cognitive representation and activate exemplars of the construction as well as the exemplar clusters for the component words (Bybee, 2010). However, the extent to which the components contribute to the meaning in

the particular context will determine the degree to which they are activated and their relation to the whole construction strengthened. For example, if the usual meanings of *rub* and *together* are not important to the interpretation of the construction, they will not be activated to any significant degree. Each such instance will decrease the analyzability and compositionality of the construction.

#### 4. Further expansion: Stage 2

As might be expected, metaphorical extensions of the construction to impoverishment in other domains have also occurred. An isolated example occurred in COCA 2009 said of a baseball team, “they’re having trouble finding *two hits per game to rub together*”. Another metaphorical extension beginning in 1957 gives us examples with *words*, *sentences* and *ideas* in the schematic slot, as shown in (8).

- (8) But Tex was a man who never had *two sentences to rub together* in his whole life. (COHA 1957)

Perhaps building on this metaphor, or as an independent development, the domain of impoverishment is moved to the brain itself with seven examples starting in 1966 similar to (9).

- (9) Given all the compelling evidence that's out there, I can not believe, in my heart of hearts, that anyone with *two brain cells to rub together*, regardless of what their position is in the tobacco industry, believes that, because it's counterintuitive and it's a lie. (COCA 1993)

It is at this point that further deterioration of the analyzability and compositionality of the construction becomes evident. The examples in this domain include a second one like (9), which uses *two brain cells*, one with *two brain neurons* and one with *two grey cells*. These examples would fall under normal expansions of schematicity, having very similar meanings. However, consider the following two examples, where the NP in the schematic slot is *two brains*.

- (10) In fact, I sometimes think that if Orrin would say things like that he'd generate less dislike than he does with that way he has of acting as though he doesn't think you have *two brains to rub together* if you disagree with him. (COHA 1966)
- (11) And believe me, this is just the way management likes it, convinced as they are that not a single person whod work as a tech at their company actually has *two brains to rub together*. (COCA 2005)

In normal usage, human beings are usually described as having one brain each, so these instances of the construction show a loss of compositionality for *two* and for

the plural, though these items are necessary to agree with *together*. A further instance shows a speaker giving up on *two* and the plural, which means loss of compositionality for *together*.

- (12) we're talking Terry, Terry who hasn't got *a brain cell to rub together*  
 (BNC)

It is important to note that these breaches of compositionality and analyzability are only found when the construction is used in the domain of intellectual capacity and the NP in the schematic slot is related to the brain. Thus it is the use in the new context that precipitates changes in the construction. The reason is most likely because physical entities such as coins can indeed be rubbed together, even though that makes less sense than when the expression applied to sticks, but the application of *rub together* to brain cells or neurons is more difficult to interpret. The extra opacity introduced by the context causes the holistic meaning to dominate even more, obscuring the way that the component parts contribute to the meaning and opening the door to creative expressions such as (10) – (12). Thus in addition to frequency of use conditioning loss of analyzability, establishment of the construction in novel contexts also leads to this loss. This finding applies not just to the development of constructions, but also to their grammaticalization when that further step occurs.

Let us return now to the instance that first attracted my attention – to the construction shown in (1), *And this guy doesn't have three friends to rub together to make a fourth*. Here we see several innovations. First, there is the extension to the social domain; second, *three* is used instead of *two*. One reason the scriptwriter may have avoided *two* is because when two people rub together, things do happen. Third, the purpose clause is present, surprisingly, since it has been absent in all the corpus examples. It was pointed out to me by an audience member attending a presentation of these ideas that in colleges such as Harvard, it is common for students to play the card game, Bridge, where four people are needed. It could be that this instance is referring to finding a 'fourth' for Bridge. This purpose clause harkens back to what we have supposed to be the original expression about rubbing two sticks together to make a fire. So this particular instance has innovative and conservative features. It is accessing the holistic meaning of impoverishment to the detriment of the meaning of *rub together*, but it reinforces the analytic meaning by adding in the purpose clause.

## 5. Conclusion: Word meaning and constructional meaning

Both children and adults analyze the parts of holistic expressions. This is particularly evident in child language acquisition when the child makes overt his/her

analysis. Clark (1993: 40) provides a long list of utterances from ages 2;4 – 3.2 showing analysis taking place. Consider her examples reproduced in (13) and (14):

- (13) D (2;4,3, looking at toy car) *That a motor-car. It got a motor.*
- (14) D (2;11,28 looking at flowering ice-plant) *What's that called?*  
Mo That's ice-plant.  
D *Does it grow ice?*

In adult language such analyses sometimes become apparent in folk etymologies, in which speakers change a complex expression to reflect what they regard as the meaning of the whole. For example, the common name for a chair with an extension that elevates the feet is a *chaise lounge*. This came from the French expression *chaise longue* ‘long chair’. Lack of familiarity with the French word *chaise* and the spelling *longue* apparently led English-speakers (primarily Americans) to analyze the second word as *lounge* in keeping with the activity associated with such a piece of furniture. Such a replacement based on the context and meaning of the whole parallels what happened when names for coins replaced *sticks* in the *rub together* construction.

As mentioned earlier, holistic meaning arises because of the propensity for both children and adults to assign meaning to the largest chunk of experience (grouping several actions into *playing fetch*) and the largest linguistic chunk (the whole word, compound or construction). Analytic meaning comes about because the words or morphemes in a complex construction activate the stored representations of these words or morphemes. As argued in Bybee 2010, repetition of the whole chunk will strengthen the holistic meaning at the expense of the identity of the individual words. In this paper I have argued that contexts of use, in addition to repetition, can reinforce the holistic meaning while detracting from the compositionality and analyzability of the components of the construction.

The meanings of both words and constructions require mappings across domains – the experiential domain and the linguistic representation. Acquiring the meaning of words and constructions requires that language users pick out certain aspects of the full context for retention as contributions to the meaning of the linguistic form. Once a chunk of linguistic material is assigned a meaning, development does not come to a halt. Rather, the context will continue to shape the meaning throughout the life of the individual and the life of the expression. As we have seen in this brief demonstration, contexts of use have led to the alteration of the meaning of the original idiomatic expression and these changes in meaning have led to the development of a schematic construction and to changes in its form.

## References

- Barlow, M. (2000). Usage, blends and grammar. In S. Kemmer & M. Barlow (Eds.), *Usage-based models of language* (pp. 315–345). Stanford, CA: CSLI.
- Beckner, C., & Bybee, J. (2009). A usage-based account of constituency and reanalysis. *Language Learning*, 59(Suppl.1, Dec.), 29–48. DOI: 10.1111/j.1467-9922.2009.00534.x
- Bybee, J. (1985). *Morphology: A study of the relation between meaning and form*. Amsterdam: John Benjamins. DOI: 10.1075/tsl.9
- Bybee, J. (2010). *Language, usage and cognition*. Cambridge: CUP. DOI: 10.1017/CBO9780511 750526
- Clark, E.V. (1993). *The lexicon in acquisition*. Cambridge: CUP. DOI: 10.1017/CBO9780511 554377
- Clark, E.V., & Amaral, P.M. (2010). Children build on pragmatic information in language acquisition. *Language and Linguistic Compass*, 4(7), 445–457. DOI: 10.1111/j.1749-818X. 2010.00214.x
- Clark, E.V., & Clark, H.H. (1979). When nouns surface as verbs. *Language*, 55(4), 767–811. DOI: 10.2307/412745
- Davies, M. (2004) BYU-BNC (*Based on the British National Corpus from Oxford University Press*). Retrieved from: <<http://corpus.byu.edu/bnc/>>.
- Davies, M. (2007). *Time Magazine Corpus (100 million words, 1923–2006)*. Retrieved from: <<http://corpus.byu.edu/time/>>.
- Davies, M. (2008) *The Corpus of Contemporary American English: 450 million words, 1990-present*. Retrieved from: <<http://corpus.byu.edu/coca/>>.
- Davies, M. (2010) *The Corpus of Historical American English: 400 million words, 1810–2009*. Retrieved from: <<http://corpus.byu.edu/coha/>>.
- Davies, M. (2012). *Corpus of American Soap Operas, 2001–2012 (100 million words)*. Retrieved from: <<http://corpus2.byu.edu/soap/>>.
- Ellis, N.C. (1996). Sequencing in SLA: Phonological memory, chunking and points of order. *Studies in Second Language Acquisition*, 18, 91–126. DOI: 10.1017/S0272263100014698
- Hay, J. (2001). Lexical frequency in morphology: Is everything relative? *Linguistics*, 39, 1041–1070. DOI: 10.1515/ling.2001.041
- Hoffman, S. (2004). Are low-frequency complex prepositions grammaticalized? On the limits of corpus data — and the importance of intuition. In H. Lindquist & C. Mair (Eds.), *Corpus approaches to grammaticalization in English* (pp. 171–210). Amsterdam: John Benjamins.
- Israel, M. (1996). The way constructions grow. In A.E. Goldberg (Ed.), *Conceptual structure, discourse and language* (pp. 217–230). Stanford, CA: CSLI.
- Langacker, R. (1987). *Foundations of cognitive grammar, Vol. 1: Theoretical prerequisites*. Stanford, CA: Stanford University Press.
- Langacker, R. (2008). *Cognitive grammar: A basic introduction*. Oxford: OUP. DOI: 10.1093/acprof:oso/9780195331967.001.0001
- Noël, D. (2007). Diachronic construction grammar and grammaticalization theory. *Functions of Language*, 14(2), 177–202. DOI: 10.1075/fol.14.2.04noe
- Oxford English Dictionary, 3rd ed.* (n.d.). Retrieved from: <<http://dictionary.oed.com/>>.
- Wilson, D.V. (2009). From ‘remaining’ to ‘becoming’ in Spanish: The role of prefabs in the development of the construction Quedar(se) + ADJECTIVE. In R. Corrigan, E. Moravcsik, H. Ouali, & K. Wheatley (Eds.), *Formulaic language, Vol. I*. (pp. 273–295). Amsterdam: John Benjamins.



# From speech with others to speech for self

## A case study of “externalized drama”

Dan I. Slobin  
University of California, Berkeley

The chapter is a case study of a particular kind of speech-for-self produced by a preschool-aged girl, characterized as “externalized dramas.” Unlike most such records of vocalized thought, this speech is not involved with guiding ongoing behavior, but rather with acting out problems of interpersonal relations with peers. Using two or more voices in dialog, the speech is full of insults and denials, claims and counter-claims, promises, excuses – all of the continuing struggles to define social roles and one’s own position. Externalized dramas practice and refine pragmatic devices of prosody, lexicon, and speech acts, while dealing with underlying problems of emotional states, violence, fantasy and reality, and other minds. It is suggested that audible inner speech goes inward to become silent speech that continues to be concerned with social dynamics and individual status and roles.

*Speakers use language to present themselves.*

Eve V. Clark (2003, p. 352)

*[C]hildren can and do affect one another’s talk in complex ways, away from adult supervision, models, or intrusion.*

Marjorie Harness Goodwin and Charles Goodwin (1987, p. 227)

As Eve Clark has often pointed out, “social interaction is essential to the process of acquisition” (Clark, 2004: 472). Her studies of early language development have illuminated the role of caregiver-child interaction in the development of grammar and lexicon. The invitation to contribute a chapter in Eve’s honor in a book entitled *Language in interaction* has led me back to the Berkeley-Stanford work of the 1960s on the development of “communicative competence” (Slobin, 1967) and to some unusual data that I’ve wanted to work on for a very long time. I present these data to Eve, who has been an intellectual and personal friend ever since that era, helping to guide the way toward understanding of how we learn to talk and how language works.

The data I present here reflect another kind of social interaction – discussions and disputes between children. The skills to be acquired here are pragmatic, in addition to grammatical and semantic. Consider two examples of how young children confront each other through speech. The first is from Brenneis and Lein (1977, p. 59) and is between two first-grade boys. The second is from my own data, in the voices of two 4-year-old girls.

- (1) CHILD A: Give that ball, you little. . .
  - CHILD B: No.
  - CHILD A: Give me that ball. [grabs ball]
  - CHILD B: You give me that ball back.
  - CHILD A: No.
  - CHILD B: You give me it. Give me that ball. [grabs ball]
  - CHILD A: Give it.
  - CHILD B: No.
  - CHILD A: Yes, I got it first.
  - CHILD B: I got the ball.
  - CHILD A: Give me it.
  - CHILD B: No, I got it.
  - CHILD A: Yes.
  - CHILD B: No.
  
- (2) CHILD A: I just dropped something.
  - CHILD B: What did you drop? Gimme it. It's mine.
  - CHILD A: No it's not.
  - CHILD B: Yes it *is*. It's *mine, mine*.
  - CHILD A: Remember I had it? Remember that time? I had it? I loved it so much. Remember?
  - CHILD B: [sarcastic laugh]
  - CHILD A: Remember that time? I loved it *so* much. *Remember?* That time?
  - CHILD B: Mm-mm [negative intonation], I don't.
  - CHILD A: Well, do!

I said that the second example is “in the voices” of two girls. That is because there was only one child in the room, my 4-year-old daughter.<sup>1</sup> Yet the structure of the

1. The data are from my daughter, Heida Shoemaker, recorded at age 4;3 in 1971. She and Eve knew each other at that time, actually meeting in Prague in 1969; Heida has generously allowed me to use these data, in 2012, as a presentation to Eve and our colleagues. Heida has approved this paper, which presents personal information about her early childhood. (Her emerging meta-linguistic awareness is discussed in Slobin (1978).)

two dialogs is strikingly similar. My study was next to her room, and one afternoon, when she was age 4;3, I heard two voices coming from her room. She frequently played there with a girl friend who lived across the street, but I hadn't heard anyone come into the house. So I went quietly to her door and looked in, and there was only one child there, sitting at her table, engrossed in drawing, and talking out loud in two different voices. I asked her if there was someone there and she was startled. She had been completely lost in drawing and had no idea that she had been speaking out loud. The speaking and the drawing were unrelated in content. Being a psycholinguist parent of that era, it was 1971, I placed a microphone in her room and recorded a number of such episodes, as well as interactions with her actual girl friend and with her mother and me. In addition, I had been keeping a daily diary since before she was 2. The diary indicates the beginning of "a flowering of imaginary play" about a year earlier, first noted at 3;3. And by 3;5 I noted: "Beginning of explicit role-playing games. These games consist of talking to herself, often for 5-10 minutes." At 3;6: "Much dialog play – taking two parts – especially in moral discussions with self. Concerned with 'good' and 'bad'." Thus it was clear that by the time I started to make recordings, these sorts of audible inner dialogs – I'll call them "externalized dramas" – had been well established. They had apparently begun a few months after she started attending preschool, where she was thrust into daily interactions with other children.

These externalized dramas provide a rare window into outer voices that may well become inner voices. Eventually the audible "speech for self" becomes silent inner speech. The goal of this little case study is to explore functions of inner speech that have not received much attention in the literature.

### Functions of inner speech

Speech for self has been examined at two periods of language development: early crib speech (Kuczaj, 1983; Nelson, 1989; Weir, 1962) and later private speech (Piaget, 1923, 1962; Vygotsky, 1934; Winsler, et al., 2009). Crib speech, as the designation indicates, reflects early attempts to control the phonology and morphosyntax of the language heard during the day, along with attempts to make sense of behavioral and moral contexts structured by adult caregivers. Private ("egocentric") speech has been studied in preschool and kindergarten settings, where children speak out loud in the presence of others to regulate their own activities. Both research traditions are concerned with the internalization of speech that the child hears – primarily the speech of caregivers and teachers. In our adult heads we talk to ourselves using distillations of the messages we received from adult guides and critics.

Externalized dramas, as I will attempt to show, reflect other voices – the voices that come from the child's peers. This reflects the world of social interaction that is central in the development of identity. Social dramas continue to be played out on the silent stage of the mind. This theme appears again and again in George Herbert Mead's seminal work, *Mind, self and society*" (1934): "He [the child] becomes a self in so far as he can take the attitude of another and act toward himself as others act" (p. 171). The two epigraphs to this chapter, from Eve Clark and from the Goodwins, underline this social and socializing role of language. Peer speech deals with themes of acceptance and rejection, hierarchy, praise and criticism, social control. It is full of insults and denials, claims and counter-claims, promises, excuses – all of the continuing struggles to define social roles and one's own position. In this workshop the child learns the speech pragmatics of the culture, using the basic tools of phonology, morphosyntax, and prosody developed earlier. Interactive discourse gives the child the tools to construct both a social language and a self.

Katherine Nelson (2012), in a recent paper that looks back on her detailed study of the crib speech of a much younger girl, Emily, provides a characterization that can also be applied to Heida's verbal enactments: "I am proposing here that the main purpose served in Emily's use of private speech for self was *making sense of the new views of reality that language use* offered through exchange of speech with others, especially realities of the social world" (p. 1). This is what Heida was concerned about in this period: making sense of the social world, but in her case it was the new social world of peers rather than family. Nelson notes that this kind of private speech is "*private in a double sense*, (1) being addressed to the self (2) in a private situation with no one else present. It is not unusual for such speech to also occur when children are playing alone, a situation that has rarely been studied as such, although it is known to occur frequently" (p. 2).

The data from Heida allow us to examine both the forms of imaginary dialog and their content. Both form and content co-constitute emerging pragmatic competence – its tools and dominant themes. And, as we will see, the patterns of solitary dialogs mirror what is reported in the literature of child discourse, going back to the classic volumes of Ervin-Tripp and Mitchell-Kernan (1977) and Ochs and Schieffelin (1979). The forms are conversational acts and narrative forms: assertion and denial, threats and insults, reconciliations and reciprocity. There are also speech acts directed at the self as "the other." The dominant contents for Heida were explorations of causes of emotional states and dealing with "the dark side": violence, killing and death. Let us first examine form, and then content.

## Forms of conversational acts in the externalized dramas

A convenient unit of analysis is John Dore's "Conversational Act Unit," defined as "an utterance in a person's turn at speaking in conversation which conveys information and expresses an attitude (intention, expectation, belief, etc.) relative to that information" (Dore, 1979, p. 342). The conversational acts of the externalized dramas tend to be moves in arguments. The turn-taking is rapid, marked by alternations of voice quality. Changes in pitch and timbre give the illusion of hearing two, and sometimes three different speakers. There is rarely any uncertainty, in transcription, in assigning conversational acts to separate speakers. And unlike the difficulty of transcribing child interaction, there can, of course, be no overlap, since there is only one person speaking the different parts. Arguments escalate, with chains of assertions and denials, peppered with insults and threats, as described for children's disputes by Brenneis and Lein (1977; Lein & Brenneis, 1978) and many others. Moral judgments are frequent. Narratives are rare, making this speech different from the crib narratives reported by Weir (1962) and Nelson (1989). These are not reminiscences of the day's events, but rather generalized "scripts" of preschool interaction, augmented by fantasy drama that exceeds everyday events.

### *Assertion and denial*

Examples (1) and (2) include alternations of assertions and denials. Brenneis and Lein (1977, p. 56) call such sequences "inversion": "Successive statements may be drawn in turn from a category and its inverse, or ... from a category and one other category which represents a denial or negation." They give the following as an example:

- (3) a. DAVE: I am, you dumb-dumb.  
LARRY: I'm not no dumb-dumb, dodo.  
DAVE: Yes, you are  
LARRY: No, I'm not.  
DAVE: Yes, you are.  
LARRY: No, I'm not.

The following example, in two of Heida's voices at 4;3, shows the same structure. Second-person charges are answered by first-person denials, with predicate ellipsis and appropriate matching of affirmative and negative auxiliary forms – what Goodwin and Goodwin call "format tying": "ways in which return moves tie to the detailed structure of the talk that they are opposing" (1987: 205). And as is typical of both enacted and real disputes, the interchanges escalate, with contrasting stress

on the appropriate lexical items and eventual reduction to a repeated series of negative interjections.

- (3) b. VOICE A: You're my friend? Are you?  
 VOICE B: Nope.  
 VOICE A: You *were*.  
 VOICE B: No I wasn't.  
 VOICE A: Yes you *were*.  
 VOICE B: No I wasn't.  
 VOICE A: *Were*.  
 VOICE B: No I *wasn't*.  
 VOICE A: Yes you *were*.  
 VOICE B: No I *wasn't*. I said, no I wasn't.  
 VOICE A: You *were*.  
 VOICE B: Nuh-uh.  
 VOICE A: Uh-huh.  
 VOICE B: Mm-mm.  
 VOICE A: Uh-huh.  
 VOICE B: Uh-uh.

### *Threats*

Voice A terminates this fruitless interchange with a threat: "You do then you'll be dead! Dead!" Lein and Brenneis (1978) make use of the same category in their study of children's disputes in three cultures, giving the example: "I'll kill you." Threats are frequent as a conversational act of last resort in Heida's externalized dramas. In the following example, the threat follows a prolonged bout of liking and hating. As soon as mutual liking is established, Voice B turns on Voice A with threats of violence.

- (4) a. VOICE A: Do you like me?  
 VOICE B: Yes.  
 VOICE A: OK, then I like you.  
 VOICE B: Heh-heh. OK. Um, you know what I will do to you? Kill you.  
 VOICE A: Well then, *I* don't like *you*.  
 VOICE B: Then *I* don't like *you*. I will kill you.  
 VOICE A: Ohhh. I like you.  
 VOICE B: Then I like you. I won't kill you.

Having achieved the desired goal of mutual liking, Voice B goes on to a different play bout, saying: "Good. C'mere, tuck in my little girl."

Threats are also noted in the diary, though they do not continue as an escalation between parent and child, as they do between peers. At 3;5 she said to her father:

- (4) b. You be nice to me or I won't be nice to you. If you don't be nice to me every time I will hit you every time.

### *Insults*

Example (3a), from Brenneis & Lein, is built on insults: "dumb-dumb," "dodo." Heida's externalized dramas also contain insults of various sorts: "dumb-dumb," "punkin' head," "crybaby." When she was first in preschool, at an earlier age (3;4), the diary notes her concern with insults between children: "Why Nicky says it to Jess: 'you idiot'?" And she reports her own use of insults, along with threats: "I call Nicky 'idiot' sometimes when I make him dead." And she gives her parents this succinct summary of preschool interactions: "Then what happened was somebody came and pinched me. This whole thing goes again and again. The boys hit the girls then the girls hit the boys then the boys hit the girls." Clearly, the insults and threats in the externalized dramas are modeled on preschool interchanges rather than interactions with adult caregivers.

### *Agreement, reconciliation, reciprocity*

Unlike reports in the literature of children's disputes, Heida's imaginary characters are not always irreconcilable antagonists and rivals. Conflicts are sometimes resolved in various ways, as in the following examples.

- (5) a. VOICE A: I don't want you to.  
VOICE B: OK, I won't.
- b. VOICE A: I can't do this one. I'm gonna do this one later. OK?  
VOICE B: OK, do it later.

There is an underlying normative sense of reciprocity and its violation. The following two voices each rely on an unstated assumption that affective states between two individuals should be in balance: you feel something towards *me* and I feel the same towards *you* (note the appropriate contrastive stress on pronouns in this example).

- (5) c. VOICE A: I hate you.  
VOICE B: I don't hate *you*. Why d'you hate *me* if I don't hate *you*?  
VOICE A: Because I don't like you and you don't like *me*.

VOICE B: I like you.  
 VOICE A: Do you?  
 VOICE B: Yeah.  
 VOICE A: Yeah? Then I like you too.

Conversational acts of reciprocity are based on an *if...then* structure. The diary notes “beginning of reciprocity” at 3;3, with a prototypical conversational act of this type, addressed to her father: “If I give you a bite of mine, you give me a bite of yours.” At 3;4 it was generalized to affective states, at the same time that Heida began to speak of preschool experiences: “If you be nice to me, I will be nice to you.” “If you get mad at me, I will get mad at you.” A month later, at 3;5, she turned the *if...then* construction into a corresponding *X...so Y*. She reported a preschool interaction, with an appropriate double negative: “I like him, so he can’t not be nice to me.” And at 4;3, the time of the recordings of externalized dramas, she made the entire structure evident in a conversational act directed to her father: *if not X, then not Y; if X then Y* – “If you don’t be nice to me, I won’t be nice to you. If you be nice to me, I’ll be nice to you.”

### *Self as referent*

Various types of conversational act place Heida in the grammatical role of third-person referent. Sometimes this is done by explicit naming, in which both voices refer to “Heida,” one as her antagonist and one as her supporter

- (6) a. VOICE A: Don’t ever play with Heida then.  
 VOICE B: Uh, sometimes play with Heida.

Another type of sequence introduces a third voice who seems to intervene on Heida’s behalf. This interchange follows directly on the dialog in (5c), repeating the last utterance here. The three voices are audibly distinct in voice quality, with the Voice C speaking at a lower pitch. Both Voice B and Voice C refer to the speaker of Voice A in the third person.

- (6) b. VOICE A: Yeah? Then I like you too.  
 VOICE B: Hey, but *I* don’t like *you*.  
 VOICE C: Well, that’s my friend. I’m a people and this is my house.  
 VOICE B: Oh yeah, but I still don’t like her. Sorry, but I still don’t like her. Anyway, OK?  
 VOICE C: OK, you never like [unclear].  
 VOICE B: I sometimes like her. But not right now.

Heida can also be the object of praise. First the voice seems to be guiding an activity – a commonly reported function of speech for self. Finally she has completed a project and is evaluating her work, using a single voice in a sequence without an interlocutor. Note the switch from first-person agent to an oblique self-reference as *me*.

- (6) c. I did this one already. It's nice of me. This one is *ready*. There, I did that one already.

#### *Outside observer as critical voice*

Sometimes an apparent outside observer criticizes a participant. In the following bout, the criticism is voiced in the third person, and is responded to with a challenge to the observer, addressed in the second person. And then the dispute can descend into first-person statements. These interchanges may well be modeled on preschool scenes in which children make third-person critical comments about others. The following segment, from 3;4, suggests that the challenger, Voice A, is Heida and the third person being commented on is someone else, or perhaps another part of Heida.

- (7) a. VOICE A: She doesn't know all those things. She doesn't know anything.  
 VOICE B: You're wrong. She doesn't know anything and *you* don't know anything.  
 VOICE A: I know lots. I know *some* things. Mama knows lots of things.  
 Papa doesn't know any.

In a long sequence at the same age, Heida is making something that breaks and the critic gives guidance and then criticism. Here the externalized drama has similarities to Vygotsky's (1934) discussions of inner speech, in which the outside voice of a teacher or parent becomes the child's inner voice, moving, as he put it, from interpersonal to intrapersonal. (However, the externalized speech recorded here does not have the elliptical quality that, according to Vygotsky, characterizes inner speech.) The following gives the highlights (Heida is Voice A and the guide/critic is Voice B).

- (7) b. VOICE A: Broke  
 VOICE B: But, if you do it *slowly* [drawn out] it won't break, but if you do it *that way* it will break.

VOICE A: I'mna do it slowly and it won't break. I'mna try it. [It apparently breaks again and Voice B gives the same critique, with more emphasis. This cycle repeats several times. After a while, Heida apparently does it quickly, and it breaks again – but now the advice seems have shifted to fast rather than slow action.]

[...]

Did it fast, then it *breaked*. *Pretty* fast anyway. Did it *pretty fast*.

VOICE B: Not at all fast. Not at *all* fast. Sorry, but not at all fast. Not at all, not at all fast. Sor-ry. *Not*.

VOICE A: I am really sorry, really, really sorry. But just can't.

### *Fantasy narratives*

There are no narratives that are simply reports of past experience, but there are occasional fantasy narratives, with word play and free association on the sorts of themes that also occur in the externalized dramas. The following is the most elaborate. It was part of a play telephone conversation with a monkey, who only says "hello," "OK," and "bye." The narrative was apparently based on a real memory, with some reported speech and ending with a moral.

- (7) VOICE A: Telephone, it goes zero, three, two, one, five, four three, two one, zero. Hi.
- VOICE B: Hello.
- VOICE A: Hi, monkey.
- VOICE B: Hi.
- VOICE A: Remember you could do lots of things? And, um, and you did lots of things for us? Remember that time?
- VOICE B: Yeah
- VOICE A: But now, our ducks, our dogs. Mickey, Nickey, Blickey, Bickey, Nickey, and, and your chickies are lost. Here's friend horse. Know what? Don't never pick this up cause, um, never pick this up cause it might chip and break. And Jenny gave it to me and it's *very* fragile. Once a little kid wanted to play with this and I said "no" and he cried. And my papa came up and said, "What's wrong?" And I said, "Well, he was gonna play with this and this is very- careful, have to be very careful with this." And um, so, and um, so, when little people come over, don't let them come over. Cause they might break this.

OK? Cause they wooshly come from- from in the monkey.  
 So when you let- when somebody comes from that crack,  
 you say, "Don't hit that!" And then I'll be happy. OK?

VOICE B: I will do it.

VOICE A: Always do that when somebody comes up and I'm not here  
 or else when I am here. OK?

VOICE B: OK. Bye.

### **Contents of conversational acts: Dominant themes**

The preceding examples have shown that many forms of dialog were established in this period and that Heida was actively engaged in contrasting fantasy and reality, using a range of grammatical constructions and rhetorical devices. At the same time, the examples cited from externalized dramas strikingly reveal major concerns of this 3–5 age range: coping with positive and negative emotions, affiliation and rejection, promises and reciprocity, fears of loss and injury, killing and death. These themes are also found in diary notes of the same period. Conversations with parents anticipate and echo the inner dialogs that go on in the child's mind.

#### *Exploration of causes of emotional states*

She was concerned with why people have the feelings they do, sometimes adding her own moral judgments. The diary records often indicate an emerging self-awareness. And there are striking early generalizations over these reflections, reflected in generic "when" statements and modifications with "sometimes."

- (10) a. When I'm sad I do bad things because I don't like it, and sometimes I get tired. (3;3)
- b. [discussing bad people, I point out: "Nobody's bad all the way through."]  
     Everybody's bad I know. [=Everybody that I know is bad.]
- c. When people doesn't come in the door, I'm not shy and when people does come I'm shy. (3;4)
- d. Maybe she's sad because she's been hit very hard by another person. (3;5)
- e. [to mother] I was happy but I am now sad because papa's been shouting at me and you been shouting at me. (3;6) [note the contrast of generic past *was*, present *am now*, and extended recent past *been shouting*]

- f. I feeled like tearing up something. I really feeled like tearing up it. (3;4)
- g. [to father] You're not promising to be nice to me [=not keeping your promise]. (4;5)

### *Concerns, with violence, killing, and death*

We have already seen these concerns acted out as threats in internalized dramas. At home, she reports that she is afraid of physical attacks from other children. At 3;4, commenting on a boy who was a neighbor and preschool friend:

- (11) a. FATHER: Do you want to see Jess tomorrow?  
HEIDA: No.  
FATHER: Why?  
HEIDA: Cause I'm afraid of Jess. He hits me.

Threats of violence were part of play at home. For example, at the same age, she made pictures and then tore them up, and shot at magazine pictures with imaginary guns.

- (11) b. Let's make mama dead. Let's make somebody dead. Let's kill them because they're sad and they want to be dead.

At 3;5 the diary notes: "for about the past two months, frequent preoccupation with death and fear of being killed." In a discussion about killing, she says she's afraid of being killed by a cop:

- (11) c. I don't want to shoot him because then he will shoot me.

### **Precursors: Foundations of fantasy play**

The fantasy of externalized drama is well-grounded in developments from early in symbolic life, as documented in detail by Piaget (1945) and many other investigators of cognitive, linguistic, and social development. In the year before the first recordings of externalized drama at 4;3, there are numerous diary notes documenting concern with pretense and counterfactuals, as well as evidence for awareness of inner speech and the contents of other minds.

### *Pretense: Fantasy, and reality*

As already mentioned, fantasy and role play were first noted in the diary at age 3;3. At 3;4 she began to play around with relations between words and things.

- (8) a. I wanna be called “Jess” [name of a friend]. Sometimes Jess can be called Jess and I can be called Jess. I will have two names: Heida and Jess.

And at 3;5 she explicitly referred to the play of changing objects by changing their names as “pretend.” She told her mother that she couldn’t dance because she wasn’t wearing a dress:

- (8) b. HEIDA: Call it a dress, please.  
MOTHER: It’s not a dress.  
HEIDA: Tend! [=pretend]

Later, at 4;3, she created a verb for pretending: “You’re make-believing.” And around that time she showed an explicit interest in pretense. At 4;6 she said to me:

- (8) c. When I drink orange juice or apple juice like wine I think I’m a grown-up – I pretend. [laughs] Funny thing, pretend. Did you do that when you were little?

At 4;8, I entered her room in the midst of a play scenario involving a queen doll who was overseeing piles of things. When I asked what the doll was doing, she clearly explained the play scenario she was setting up, referring to two real friends, Amy and David, who also figured in many of her externalized dramas at that time.

- (8) d. See, this is Amy’s pile. I’m pretending David and Amy are here. And this is Amy’s pile. This is my pile. And this is David’s pile. (4;8)

Long before, however, beginning at 3;3, there are indications of a concern with fantasy and reality. The diary records numerous “what if” questions, role-playing games, and even identity shifts. Here are some examples:

- (8) e. If I would eat this plate, what would happen? (3;3)  
f. What happens if I eat not any food? What happens if I wouldn’t eat any food? (3;4)

And role-playing games were carried out with and without props:

- (8) g. This little person wants to go home. This little person doesn’t want to go home. [role-playing game, with props] (3;5)  
h. I’m a baby kitty what doesn’t know how to talk. I’m a kitty who only can talk like this: meow, meow. (3;6)

In the midst of this buzzing period of pretense, role-play, and externalized drama, she was trying to work out the borders between fantasy and reality. There is a

telling diary note to this effect, from age 4;8: "She enjoys attempts to deceive us, and comes up with questions beginning with 'Do you think..', expecting us to take the wrong option. But she isn't good at it, and can't understand how we can tell which option is true."

### *Awareness of inner voice and other minds*

During this period there were indications that she was explicitly aware of mental processes. At 3;4 she talked about her own inner voice:

- (9) a. I said to myself, "I want my mama and my papa to play with me." (3;4)

At 3;10 inner speech apparently includes self-directives and assertions of autonomy:

- (9) b. I wanna wear those things that I tell myself to do. (3;10)

She even seemed to show an awareness of her mind and other minds, as reflected in this rather surrealistic assertion at 4;8:

- (9) c. I'm the bosser in my room. I'm the bosser in your head. I'm the bosser in everyone's head – everywhere – except in my ogubohagu.

## Discussion

The literature on children's "private speech" or "speech for self" in this age range is generally concerned with the executive function, as realized through verbal self-regulation, focusing on the planning and imperative functions of speech as instrumental in cognitive development. This focus is clear in the title of a recent book by Winsler et al. (2009): *Private speech, executive function, and the development of verbal self-regulation*. As Vygotsky saw it, these functions begin interpersonally, becoming intrapersonal – "from the social to the individual" (1962, p. 20). When Piaget first encountered Vygotsky's posthumous work, in 1962, he reported that he shared Vygotsky's view with regard to the cognitive functions of speech for self, summarizing that "this interiorized language can serve ... logical thinking" (Piaget, 1962, p. 7). Heida's externalized dramas, however, do not seem to be on their way inward in the service of logical thinking or the learning of scientific concepts in school, though they may play a role in some kinds of self-regulation.

Piaget had another interest in children's quarrels – namely, as functional in bringing the child to decenter: "It may well be through quarreling that children first come to feel the need for making themselves understood" (1955, p. 83). And, in his response to Vygotsky: "...it is precisely co-operation with others ... that

teaches us to speak ‘according’ to others and not simply from our own point of view” (1962, p. 8). This is, in fact, close to the position of George Herbert Mead, quoted above: “The taking or feeling of the attitude of the other toward yourself is what constitutes self-consciousness” (1934, p. 171).

All of this may be true – exercising executive function, taking the point of view of the other, re-narrating experiences in order to grasp their temporal and causal structure – but these do not seem to be the salient functions of externalized drama. In the practice of externalized drama, Heida was learning, and practicing, types of conversational moves. When these become skilled, she will be able to run through them silently. She will be able to experience a challenge and try out counter-challenges; she will be able to imagine threats and insults and responses to them. The expertise developed in the dramas provides a flexibility for eventually trying out moves internally. The dialogs are different from regulatory speech for self, in which one part of the self talks to another. What emerges here is the ability to mentally create “story-boards” and see where they might lead. Heida’s brief period of externalized dramatic interactions have let us listen in on the work of the script writer. The analogy to stage drama provides a conception of this variety of speech for self as part of the development of social interaction. At the same time, it shows emerging control of a range of sociolinguistic and pragmatic skills.

Where do these externalized dramas go as the child grows older? In addition to the ways in which the speech of others moves inward and becomes speech for self, inner speech in several voices provides the child with a private stage on which to re-enact and attempt to master complex worlds of social interaction. The child not only speaks to herself in different voices, but echoes the interplay of voices of daily life with other children. There is no “input” here in the classic sense of caregiver speech – that is, there is no internalization of structures and routines received from adults – but there is the input of “sociolinguistic texts” co-produced with peers. On the stage of inner conversation, the child works to reconcile competing voices, to overcome her fears, to establish her own social identity. Do we continue to act out internalized dramas as adults? To some extent this must be true, but – except, perhaps, for playwrights and novelists – our dramas have less fantasy, less make-believe. Or, at least, this seems to be true of conscious listening to inner voices. The voices of dreams, and the voices that are evoked by some kinds of therapy and altered states, show that the dramas of the 4-year-old have not been totally replaced by logical reasoning, planning, and self-regulation. But here we would have to leave psycholinguistics and sociolinguists – the expertise of this volume – and move to the unfamiliar territory of Bakhtin’s (1981) “polyphonic novels,” Hubert Hermans’ (2011) Dialogical Self Theory, Jungian psychology, and no doubt more. Such a trip is not on our itinerary, but it is valuable to at least look at the map and think about different functions of adult inner speech. Hermans

(2003, p. 94) gives us a glimpse of the journey not taken in this paper: “The Bakhtinian speaker is not the origin of abstract thinking or analytical thought, so typical of the Cartesian *Cogito*, but is involved in exterior and interior dialogues in which emotions are expressed in intonations and virtual gestures. The words of other people, invested with indignation, anger, doubt, anxiety, or pleasure, enter interior dialogues and create an ‘inner society of voices’ that, in its oppositions, agreements, disagreements, negotiations, and integrations, does not, in essence, differ from the communications in the outside world.”

In concluding, it is time to bring all of these considerations, all of these intriguing bits of data, back to Eve Clark’s lifelong mission to explore how children acquire language. In her monumental guide to the field, with its straightforward title, *First language acquisition* (2003), she lays out the issues of “social dimensions” that are addressed in the sorts of real and imagined conversations discussed here:

Conversation provides a forum for using language. It displays language embedded in larger systems for communication and so should present children with critical material for making sense of language as they try to understand others and make themselves understood. Conversational exchanges ... should also be a forum for learning to become a member of the society and the culture (p. 7).

## References

- Bakhtin, M.M. (1981). *The dialogic imagination: Four essays* [written during the 1930s]. Austin, TX: University of Texas Press.
- Brenneis, D., & Lein, L. (1977). “You fruithead”: A sociolinguistic approach to children’s dispute settlement. In S. Ervin-Tripp & C. Mitchell-Kernan (Eds.), *Child discourse*, (pp. 49–65). New York, NY: Academic Press.
- Clark, E.V. (2003). How language acquisition builds on cognitive development. *First language acquisition*, 8, Cambridge: Cambridge University Press. 472–478.
- Clark, E.V. (2004). How language acquisition builds on cognitive development. *Trends in Cognitive Sciences*, 8, 472–478.
- Dore, J. (1979). Conversational acts and the acquisition of language. In E. Ochs & B.B. Schieffelin (Eds.), *Developmental pragmatics* (pp. 339–363). New York, NY: Academic Press.
- Ervin-Tripp, S., & Mitchell-Kernan, C. (Eds.). (1977). *Child discourse*. New York, NY: Academic Press.
- Goodwin, M.H., & Goodwin, C. (1987). Children’s arguing. In S. Philips, S. Steel, & C. Tanz (Eds.), *Language, gender, and sex in comparative perspective* (pp. 200–248). Cambridge: Cambridge University Press.
- Hermans, H.J. (2003). The construction and reconstruction of a dialogical self. *Journal of Constructivist Psychology*, 16, 89–130.
- Hermans, H.J.M. (2011). *Handbook of dialogic self theory*. Cambridge: Cambridge University Press.

- Kuczaj, S.A. (1983). *Crib speech and language play*. New York, NY: Springer.
- Lein, L., & Brenneis, D. (1978). Children's disputes in three speech communities. *Language in Society*, 7, 299–324.
- Mead, G. H. (1934). *Mind, self & society from the standpoint of a social behaviorist*. Chicago, IL: University of Chicago Press.
- Nelson, K. (Ed.). (1989). *Narratives from the crib*. Cambridge, MA: Harvard University Press.
- Nelson, K. (2012). Making sense with private speech. Presented at "Monologuer: formes et pratiques/Self-talk: Forms and practices." Sorbonne Nouvelle, Institut des humanités: Paris, France.
- Ochs, E., & Schieffelin, B.B. (Eds.). (1979). *Developmental pragmatics*. New York, NY: Academic Press.
- Piaget, J. (1923). *Le langage et la pensée chez l'enfant*. Paris, France: Delachaux et Niestlé [translated 1926. *The language and thought of the child*. London: Kegan Paul, Trench, Trubner & Co.]
- Piaget, J. (1945). *La formation du symbole chez l'enfant: Imitation, jeu et rêve, image et représentation*. Paris, France: Delachaux et Niestlé [translated 1962. *Play, dreams and imitation in childhood*. New York, NY: Norton].
- Piaget, J. (1955). *The language and thought of the child*. New York, NY: Meridian Books.
- Piaget, J. (1962). *Comments on Vygotsky's critical remarks concerning The language and thought of the child, and Judgment and reasoning in the child*. Cambridge, MA: The MIT Press.
- Slobin, D.I. (Ed.). (1967). *A field manual for cross-cultural study of the acquisition of communicative competence*. Berkeley, CA: Language-Behavior Research Laboratory.
- Slobin, D.I. (1978). A case study of early language awareness. In A. Sinclair, R. Jarvella, & W.J.M. Levelt (Eds.), *The child's conception of language* (pp. 45–54). Heidelberg: Springer.
- Vygotsky, L.S. (1934). *Myslenie i reč'*. Moscow, RU: Izd-vo Narkompros RSFSR. [translated 1962. *Thought and language*. Cambridge, MA: The MIT Press and New York, NY: Wiley.]
- Weir, R. (1962). *Language in the crib*. The Hague: Mouton.
- Winsler, A., Ferryhough, C., & Montero, I. (2009). *Private speech, executive function, and the development of verbal self-regulation*. Cambridge: CUP.



# How to talk with children

Herbert H. Clark

Department of Psychology, Stanford University

It is generally assumed that adults learn how to talk with children in a special style called child-directed speech. But this cannot be the whole story. Each child's ability to speak and understand is a moving target, changing yearly, weekly, even daily. How could adults adapt to these changes? Evidence shows that in conversation both adults and children try to establish, as they go along, the mutual belief that they have understood each other well enough for current purposes. It is this process, called grounding, that allows adults to infer the child's current abilities and to adapt their speech to that child. Adults, then, learn how to talk with children in the very act of talking with them.

## Introduction

Talking with children isn't easy. We may know the language a particular child is destined to speak – English, French, Japanese – but we cannot know which parts he or she already knows. We may know other aspects of communication the child will master – from taking turns to being polite – but here again we cannot know which aspects he or she has already mastered. Each child's ability to speak, understand, and converse is a moving target: it changes not just yearly, monthly, weekly, and daily, but moment to moment, and we cannot know where he or she is at any moment.

How, then, *do* we manage to talk with children? One view is that we switch to a style of speaking that has variously been called *baby talk*, *motherese*, and *infant- or child-directed speech*, which I will lump under the term *child-directed speech*. But how would we learn this style? If we acquired it by watching other adults talk to children, what about adults who have no such models? And if we acquired the style by talking with children, how would we do that? Even once we learned child-directed speech, how would we adapt the style to a child whose abilities were constantly changing? Child-directed speech by itself is no solution to the child-as-moving-target problem.

To discover the solution people actually use, we first need to distinguish (1) talking *for* others, (2) talking *to* others, and (3) talking *with* others.

1. *Talking-for.* When people talk *for* others, they design what they say to be understood by a *class* of people and not by any one individual or set of individuals. On late-night television talk shows, the hosts and their guests talk in ways they believe their unseen viewers will be able to understand. Even when they address viewers directly, they treat them not as individuals but as a class. The same goes for the adults, children, and puppets on the television program *Sesame Street*. They talk in a way they believe most viewers aged 2 to 6 should be able to understand.
2. *Talking-to.* When people talk *to* others, they design what they say for the individual people they are speaking to. When I leave a voice message for my sister on her telephone, I design what I say not merely for people *like* my sister, but for my sister herself. I refer to people, places, and things she and I both know because of our shared history. Talking-*to* is a more stringent form of communication than talking-*for*.
3. *Talking-with.* When people talk *with* others, they engage them in interaction. When my sister and I talk face to face, we do more than take turns talking *to* each other. I design what I say in *collaboration with* her, and she does the same with me, and that takes actions from us both (see H. Clark & Wilkes-Gibbs, 1986; H. Clark & Schaefer, 1989; H. Clark, 1996; Sacks, Schegloff, & Jefferson, 1974; Schegloff, 2007). Talking-*with* is even more stringent a form of communication than talking-*to*.

Child-directed speech is really a type of talking-*to*. Several of the ground-breaking papers on child-directed speech appeared in a 1973 book called *Talking to children*, in which it was tacitly assumed that adults talk *to* children, not *with* them. This, of course, cannot be right. In many societies, parents do more than talk to children. They engage them in interactive conversation. Although not all of their talk is interactive, much of it is.

In societies like this, I suggest, adults solve the child-as-moving-target problem simply by talking *with* children. When they do that, they do what they would do with anyone they talk with: they try to establish, as they go along, that they and their partner have understood what each other meant well enough for current purposes. This is a process my colleagues and I have called *grounding* (H. Clark, 1996; H. Clark & Brennan, 1991; H. Clark & Wilkes-Gibbs, 1986; H. Clark & Schaefer, 1989). It is grounding that enables adults to track a child's abilities moment by moment.

The account I will outline here was not pulled out of thin air. It was inspired by, perhaps even purloined from, Eve Clark's work on conversations between

adults and children. Still, it was also based on the work on grounding in adult conversations. Let us start with the accounts of child-directed speech.

### Adult speech to children

One day when I was talking with my niece, she turned to talk with her two-year-old son, and as she did so, her speech changed dramatically. She talked more slowly, used shorter sentences, and spoke with an exaggerated intonation, speaking in a way she would never have spoken with me. It is this style that has been identified as child-directed speech. In describing this style, I will rely Eve Clark's excellent review in *First Language Acquisition* (Clark, 2009).

#### *Clarity in speaking to children*

A striking feature of child-directed speech is its clarity: adults go to great lengths to make themselves clear. Adults talking to children use shorter sentences and fewer subordinate or conjoined clauses. They speak more slowly – lengthening important words and adding pauses both within and between sentences. And they almost never produce disfluencies such as *uh*s, *um*s, repeated words, and self-repairs. In one study (Newport, Gleitman, & Gleitman, 1977), adults produced only one disfluency in 1500 utterances in conversations with children, although they produced disfluencies in 5% of their utterances in conversations with adults.

The classic research on child-directed speech was carried out by Eve Clark's long-time colleague Charles Ferguson (Ferguson, 1964, 1977, 1982). He observed that adults talking to very young children try to avoid phonologically difficult consonants, such as "f" and "th" in English, and difficult clusters of consonants, such as "st" in English. So instead of *father*, *mother*, and *stomach*, they use *daddy*, *mommy*, and *tummy*. He argued that features like these defined a *register* he called baby talk.

If adults everywhere try to make themselves clear to children, then many features of baby talk should be common across languages, and Ferguson argued that they are. In the six languages he studied (Ferguson, 1964), he found that:

1. “baby talk items consist of simple, more basic kinds of consonant, stops and nasals in particular, and only a very small selection of vowels” (p. 109);
2. there is “the predominance of reduplication, both of parts of words and of whole words” (p. 109), as in English *mama*, *papa*, *poo-poo*, *pee-pee*, and *itty-bitty*;
3. there is “the absence of any inflectional affixes” as in “*baby go poo-poo*” and “the presence of a special baby-talk affix” (p. 110), such as the diminutive “*y*” in English, as in *doggie*, *kitty*, and *birdie*; and

4. there is a “greater use of nouns rather than pronouns and verbs” (p. 106), leading adults to say “dollie pretty” instead of “the doll is pretty” and “daddy wants” instead of “I want.”

Another goal of child-directed speech is simply to capture and keep children’s attention. One way to do this is with exaggerated intonation. Research also reviewed in *First Language Acquisition* shows that adults speaking to children raise the pitch of their voice and exaggerate the rises and falls in their pitch. They use special intonation even before infants can understand words. In a study by Fernald (1989), mothers were asked to produce utterances for five different functions, and in doing so, they used five identifiable patterns of prosody (roughly intonation plus rhythm):

<u>Function</u>	<u>Example</u>
1. Approval	“Good boy yeah” in large, smooth pitch swings on each word
2. Prohibition	“No. No. No” in three short, rhythmic, staccato bursts with small pitch swings
3. Attention-bid	“Look at the ball” in two large smooth pitch swings
4. Comfort	“Oh … yeah …” in one long, low, smooth pitch swing
5. Game	“Peek-a-boo” in two low plus one large rhythmic pitch swings

Adults used roughly the same prosody regardless of language. In a follow-up study by Fernald (1993), one-year-olds with English speaking mothers were found to respond differently to spoken approvals and disapprovals not only in English, but in languages they had never heard – German and Italian. What the infants responded to, apparently, was the intonation of the approval or disapproval and not its wording.

Ferguson’s most important insight was that child-directed speech is simply a “modified version of normal adult speech” (Ferguson, 1977). And why do adults modify their speech? It is to capture and keep children’s attention and to make it easy for them to understand. These features, as we will see, fall out naturally from the process of grounding.

### *Conversation with children*

There has long been evidence that child-directed speech is also designed to support conversation. In many cultures at least, mothers create exchanges with babies from as early as one month of age and elaborate on these exchanges as the babies grow older (Trevarthen, 1979). By the time babies are three months old in these cultures, many mothers talk to them in type of pseudo-conversation, as in this example recorded by Snow (1977, p. 12):

- (1) Ann: (smiles)
- Mother: oh, what a nice little smile  
Yes, isn't that nice?  
There  
There's a nice little smile
- Ann: (burps)
- Mother: what a nice little wind as well  
Yes, that's better, isn't it?  
Yes  
Yes
- Ann: (vocalizes)
- Mother: there's a nice noise

This is only an apparent exchange since it was the mother who created the turns. When Ann (aged 3 months) smiled, burped or vocalized, her mother responded as if Ann had taken a genuine turn. By the time Ann was 18 months old, she was taking genuine turns with single word utterances such as "mouth" and "face." As children get older, Snow argued, these exchanges look more and more like adult conversations.

In conversations like these, adults do more than try to get children to understand. They try to manage and engage them in continued talk. If adults were simply trying to choose easy sentences, they should prefer ones that are simple syntactically, such as declaratives like "You can sing a song." And yet in a study by Newport et al. (1977), adults talking with 12- to 24-month-olds were more likely to use imperatives ("Sing a song!"), wh-questions ("What can you sing?"), yes-no questions ("Can you sing a song?"), and other questions ("What is that?"). They used the complex constructions 62% of the time and declaratives only 30% of the time. Why did adults use the more complex constructions? To get children to take the next turn and respond with relevant information. One of their goals, pretty clearly, was to keep children engaged in conversation.

Adults have many ways of keeping children engaged. They can use the child's name ("Anne!"), direct children's attention with deictic expressions ("Look over there" "What's this?"), and simply point (Clark & Estigarribia, 2011). In conversations studied by Newport et al. (1977), adults used "this," "that," "here" or "there" in 16% of the utterances they addressed to 12- to 24-month-old children, but in only 2% of those they addressed to other adults. Adults can also repeat words until the child responds, as with this mother speaking to a two-year-old: "Pick up the red one. Find the red one. Not the green one. I want the red one. Can you find the red one?" (Snow, 1972).

Why, then, do adults speak to children the way they do? I will argue that it isn't just to help them understand or take part in conversation. It is about getting these conversations to succeed. That, in effect, is what grounding is all about.

### Grounding in conversation

To communicate is, etymologically, to "make common," to establish information as common or shared. But how do people make certain that the right information has become truly common or shared?

One approach has focused on failures in communication (see, e.g., Schegloff et al., 1977). The idea is that conversations are vulnerable to "intrinsic troubles," as when participants misspeak, mishear, or change their minds. It is up to the participants to monitor for troubles and, when they find them, repair them. If so, people need a system for identifying and repairing the troubles that arise.

A second approach has focused, instead, on success and failure together (Clark, 1996; Clark & Schaefer, 1989). The idea is that participants try to establish, as they go along, the *mutual belief that they have understood each other well enough for current purposes*. Not only should they repair things that have gone wrong, but they should display and acknowledge *positive* evidence of things that have gone right. It takes both to reach joint closure on what addressees have understood speakers to mean. The process of establishing these beliefs is called *grounding* (Clark, 1996; Clark & Brennan, 1991; Clark & Schaefer, 1989; Clark & Wilkes-Gibbs, 1986). People work at grounding as much as the circumstances allow.

Grounding, I will argue, is as essential to conversations with children as it is to conversations with adults. Grounding is part of children's earliest attempts to communicate with adults, and it is needed by adults for communicating with children. Effective grounding is not possible when people talk *to* or *for* each other. It is only possible when they talk *with* each other.

### *Evidence of understanding*

An important source of positive evidence of understanding in dialogues is *acknowledgments*. In listening to extended descriptions, people regularly add "uh-huh," "m-hm," "yeah," "yes," and nodding as positive claims of understanding-so-far (Jefferson, 1984, 2002; Schegloff, 1982). Here is a British description with four acknowledgments (with overlaps marked in adjacent pairs of square brackets):<sup>1</sup>

1. Most of the examples of grounding come from either the London-Lund corpus (Svartvik & Quirk, 1980) and are numbered by file, conversation, and line (e.g., 1.5.34), or from the Switchboard Corpus (Godfrey, Holliman, & McDaniel, 1992) and are numbered by conversation (e.g., 3476).

- (2) 1 Beth and I went to some second year seminars, where there are only about half a dozen people,  
 2 Ann [m,]  
 3 Beth [and] they discussed what a word was,  
 4 Ann [m,]  
 5 Beth [and -] what's a sentence, that's [ev] en more difficult, .  
 6 Ann [yeah,]  
 7 Ann yeah, -  
 8 Beth and so on, . (1.5.34)

With each “m” (the British equivalent to American “m-hm”) and “yeah,” Ann claimed to have understood the just previous clause well enough for Beth to go on, and she invited her to do so (Schegloff, 1982). When people want something stronger than “uh-huh” or “yeah,” they can use *assessments* such as “gosh,” “really,” “oh?” or “good God” (Goodwin, 1986).

Another source of positive evidence is second parts of so-called *adjacency pairs*. An adjacency pair is a pair of spoken turns, by different speakers, in which the first part projects the second part as the next turn (Schegloff & Sacks. 1973). Examples are question plus answer, and greeting plus greeting. Consider this question and answer:

- (3) 1 Kate who is it?  
 2 Jane oh, it's Professor Worth's secretary, from University College,  
 3 Kate m, (8.3d.230)

When Jane responded “oh, it's ...” she passed up the chance to ask for clarification (“Do you mean me?”) or a repeat (“What?”), and responded instead with what she believed was an appropriate answer. She was claiming, by implication, that she understood Kate's question well enough to answer it. But *did* she? She could only be sure once Kate acknowledged her answer, as Kate did with “m” in line 3. In this way, Jane and Kate each provided the other with positive evidence that Jane had understood Kate as intended. That, in turn, allowed them to consider the question to be grounded (Clark & Schaefer, 1989).

When an answer to a question yields negative evidence of understanding, that normally leads to a repair. Here is an example:

- (4) 1 Abe Do do do you know, where you are, do the schools emphasize the metric system?  
 2 Bill Yeah, in the engineering they all do pretty much.  
 3 Abe No, I I I meant I meant down, like, in the elementary schools.  
 4 Bill Oh, in the elementary schools. I don't know. (*continues*) (3476)

Although Bill thought he understood Abe's question (in line 1), his answer (in 2) showed Abe that he did not, so Abe clarified it (in 3) with "No, I meant ..." Bill, in turn, displayed his understanding of the clarification (in 4) "Oh, in the elementary schools ...," which Abe *did* accept. Like Kate and Jane, Abe and Bill worked together to get joint closure: *both* had positive evidence of Bill's understanding of Abe's question.

### *Patterns of repairs*

Repairing problems in dialogue is an interactive process. As Schegloff, Jefferson, and Sacks (1977) argued, speakers have two preferences in making repairs. The first is for speakers to make their own repairs, as in this example:

- (5) 1 Sam we must ha- we're . big enough to stand on our own feet now,  
     2 Reynard yes, (1.2.33)

Sam's replacement of "we must ha-" with "we're big enough ..." was a *self-repair*, and self-repairs are legion. Repairs can also be made by one's partner, as in this example:

- (6) 1 Maggie you fancy it yourself do you? -  
     2 Julia what, the men's doubles?  
     3 Maggie yeah,  
     4 Julia well more than the singles, yes, - (7.3e.278)

"The men's doubles" was Julia's repair, not Maggie's, so it was an *other-repair*. There is much evidence that people prefer self-repairs to other-repairs.

The second preference is for speakers to *initiate* their own repairs. When Sam replaced "we must ha-" with "we're big enough ..." he initiated the repair without being prompted – a *self-initiated repair*. For a contrast, consider the next example:

- (7) 1 Roger well there's no general agreement on it I should think,  
     2 George on what, .  
     3 Roger on uhm – on the uhm – the mixed up bits in the play, the  
               [uhm]  
     4 George [yes] (3.5a.283)

Here Roger clarified the word "it" only after Sam asked for clarification ("on what?"). Roger's self-repair was *other-initiated*. There is also much evidence for the preference of self- over other-initiation of repairs.

### *Grounding in full dialogues*

Dialogues are often treated as *bare* dialogues – as if they consisted of nothing but speech. But when people are face-to-face, they engage in *full* dialogues in which they use not only speech, but gestures, eye gaze, positioning of their bodies, and other signals. Full dialogues, in turn, are used in carrying out larger joint activities – planning, telling stories, buying and selling goods, playing basketball. A common problem in many studies is that dialogues are analyzed only after they have been surgically excised from the activities they were part of, which obscures *why* participants said what they said.

In full dialogues, grounding makes liberal use of gestures, nods, smiles, eye gaze, positioning of the body, and other unspoken signals. In one study (Clark & Krych, 2004), one participant (the “director”) was asked to guide a second participant (the “builder”) in assembling a small model of Lego blocks. The two participants sat at opposite ends of a table. In the following excerpt, David and Ben could not see each other or each other’s space:

- (8) 1 David And then you’re gonna take a blue block of four.
- 2 Ben M-hm.
- 3 David And you’re gonna put it on top of the four blocks – four yellow blocks farthest away from you.
- 4 Ben Which are the ones closest to the green.
- 5 David Yeah
- 6 Ben Okay. But the green’s still not attached.
- 7 David Yeah. And then ...

David made a series of self-repairs, one initiated by himself and the rest by Ben. Everything they did was with speech, 49 words in total. In the next excerpt, in contrast, Doris and Betty could see each other, and Doris could see Betty’s workspace and the model she was building:

- (9) 1 Doris Take a short blue.
- 2 Betty (*Retrieves a short blue block.*)
- 3 Doris (*Looks at Betty’s block.*) Put it at the end of the yellow close to the green.
- 4 Betty (*Places the blue block on the yellow block.*)
- 5 Doris (*Looks at result.*) Take a ...

Although Doris and Betty were at precisely the same point in assembling the Lego model as David and Ben, they did most of their grounding by displaying blocks and looking at the result. Doris used only 16 words in total, and Betty used none. This was typical. Assembling Lego models took half the time, and fewer than half

the words, when the director could see the builder's workspace – as Doris could. The goals in grounding were the same for Doris and Betty as for David and Ben, but the techniques were not.

Grounding also makes reference to the larger joint activity the talk is coordinating. In line 3, Doris told Betty, "Put it at the end of the yellow close to the green." Betty could have responded "Okay," but she responded instead by placing the blue block on the yellow block for Doris to check visually. The two of them grounded Doris' request by referring to the next step in their joint activity – Betty's placement of the blue block on the yellow block.

The gestures used in grounding take many forms. In a further example from the Lego study, Danny the director and Ed the builder could see each other, and Danny could see Ed's workspace:

- (10) 1 Danny and now get (.75) a-uh eight piece green, (*waits 1.5 sec while Ed rummages through the blocks and retrieves an "eight piece green"*)  
 2 Ed (*exhibits the block to Danny*) (= "I've got one")  
 3 Danny and join the two ... (continues)

In line 1, Danny asked Ed to get a particular type of block, and when Ed got one, he *exhibited* it to Danny, holding it out for Danny to see (line 2). In line 3, Danny accepted Ed's block as correct by continuing on. Line 1 was grounded by Ed's exhibiting the block in line 2. Gestures also arise in side sequences, as in the continuation of Danny and Ed's dialogue:

- (11) 3 Danny and join the two so it's all symmetric-  
 4 Ed (*poises the block over a location in the model-so-far*) (= "Does the block go here?")  
 5 Danny yeah, right in the center  
 6 Ed (*affixes the block to the model-so-far*)

In line 3, Danny told Ed where to put the block he had just retrieved. But because Ed was uncertain where to put it, he initiated a side-sequence in line 4 (an other-initiated repair) with a gesture, *poising* the block over the location he believed it should go. Danny took him as asking, "Does the block go here?" and answered, "Yeah, right in the center." Once the location had been cleared up, Ed affixed the block as instructed (line 6).

In conversation, then, participants do more than hope and pray they have understood each other. They work together to establish the mutual belief that they have understood each other well enough for current purposes. To establish this belief, they use speech, gestures, and any other evidence they can assume is shared. This is the process called grounding.

## Conversations with children

The idea that to communicate is to make common is one that children appear to appreciate from the beginning. Evidence suggests that children try to ground their communication from as early an age as they try to communicate at all. What is remarkable is how effective they are in grounding even before they can speak and even when they can produce only one word at a time.

### *Pursuit of mutual understanding*

When children want something – and want it badly enough – they will persist in trying to reach their goal until they have reached it (or run into insurmountable barriers or lose interest). Even before they are a year old, they realize that for certain goals they need the help of an adult. Consider a 12-month-old named Jordan, studied by Golinkoff (1983, p. 58), who wanted a sponge that was out of reach on a nearby counter. Because he wasn't yet able to speak, this is what he and his mother did:

- (12) 1 Jordan (vocalizes repeatedly until his mother turns around)
- 2 Mother (turns around to look at him)
- 3 Jordan (points at one of the objects on the counter)
- 4 Mother Do you want this? (holds up milk container)
- 5 Jordan (shakes his head no) (vocalizes, continues to point)
- 6 Mother Do you want this? (holds up jelly jar)
- 7 Jordan (shakes his head no) (continues to point)
- 8 [2 more offer-rejection pairs]
- 9 Mother This? (holds up sponge)
- 10 Jordan (leans back in high-chair, puts arms down, tension leaves body)
- 11 Mother (hands Jordan sponge)

Jordan's strategy was remarkable for a 12-month-old. He had to begin with a hierarchy of goals:

- a. I want the sponge on the counter
- b. I want my mother to get me the sponge on the counter
- c. I need to *make it common with my mother* that I want her to get me the sponge on the counter.

Jordan's top goal was to get the sponge, but achieving that depended on goal *b*, which depended on goal *c*. He was systematic in going after goal *c*. First, he got his mother's attention (line 1), which she grounded for him by looking at him (line 2). Then, and only then, did he point at the sponge (line 3). But because his mother

wasn't sure what he was pointing at, she initiated the side sequence in line 4 (an other-initiated other-repair), which failed in line 5, and so on, until she guessed the right object in line 9. Jordan grounded his reference in line 10 by discontinuing his pointing and relaxing his body. It took 15 moves for Jordan and his mother to reach closure on his reference to the sponge, but they persisted until the closure was complete (for similar examples, see Bates, Camaioni, & Volterra, 1975).

Children show the same persistence even after they have begun to produce words. In a report by Scollon (1976, 1979), Brenda (aged 19 months) was recorded as she faced an electric fan while speaking with her mother:

- (13) 1 Brenda fei  
      2 Brenda fae  
      3 Mother Hm?  
      4 Brenda fei  
      5 Mother Bathroom?  
      6 Brenda fani  
      7 Brenda fai  
      8 Brenda fei  
      9 Mother Fan! Yeah.  
 10 Brenda ku  
 11 Mother Cool, yeah. Fan makes you cool.

Brenda repeated the word *fan*, revising her pronunciation on each repetition, until her mother confirmed it saying, “Fan! Yeah,” which grounded the reference. Only then did Brenda proceed to say “ku,” which her mother grounded with “Cool, yeah.”

What Brenda produced was an utterance in two installments: “Fei” and “Ku.” Installment utterances are common in adults, as in this example (see Geluykens, 1992):

- (14) 1 Adam and that table tennis room of mine,  
      2 Ben mhm  
      3 Adam it's always cold in there (1.7.119.1)

Adam produced a first installment (“that table tennis room of mine”) and got Ben to confirm it (with “mhm”) before going on to the second installment. Brenda did much the same thing. She produced “fei”, which took her mother and her eight turns to ground, and only then did she go on to “ku.” These are what Scollon (1976, 1979) called *vertical constructions*. They require adult and child to work together, grounding the first part before going on to the second.

Not all of Brenda’s efforts at grounding were successful, as in this example at the same age:

- (15) 1 Brenda car [4 times, with revised pronunciations]  
 2 Father What?  
 3 Brenda go [2 times, with revised pronunciations]  
 4 Father xxx [untranscribable]  
 5 Brenda bus [9 times, with revised pronunciations]  
 6 Father What? Oh, bicycle? Is that what you said?  
 7 Brenda na'  
 8 Father No?  
 9 Brenda na'  
 10 Father No – I got it wrong

### *Varieties of grounding*

Adults, as I noted, prefer self-repairs to other-repairs, and self-initiated repairs to other-initiated repairs. Children have the same preferences – so far as one can tell (Laakso & Soininen, 2010; Solonen & Laakso, 2009). Brenda initiated and made five self-repairs of *fan*, three of *car*, one of *go*, and eight of *bus*. And even without words, Jordan made five self-repairs, although they were all initiated by his mother. Children appear to make or initiate other-repairs only when they are older (Forrester & Cherington, 2009). For a more complete picture, I will appeal to studies by Eve Clark and her colleagues (Chouinard & Clark, 2003; Clark, 2007; Estigarribia & Clark, 2007; Clark & Estigarribia, 2011; Clark & Bernicot, 2008), though I will describe only bits of those studies.

Recall that adults provide positive evidence of understanding with *uh-huh*, *yeah*, head nods and other acknowledgments. Another way is by repeating words from the previous turn, as in this exchange from a call to directory enquiries in Cambridge UK (Clark & Schaefer, 1987):

- (16) 1 Operator It's Cambridge 68947  
 2 Customer 68947  
 3 Operator That's right

In repeating 68947, the customer gave the operator the chance to correct it if it was wrong. Adults talking to children do much the same, as in the conversation with Brenda:

- (17) 10 Brenda ku  
 11 Mother Cool, yeah. Fan makes you cool.

The mother's repeat of "cool" gave Brenda a chance to correct it if it was wrong. Indeed, the mother's repeat of "bathroom" was wrong, and Brenda corrected it.

Children, too, use repeats to ground what is said. Here is an example from D aged 18 months:

- (18) 1 D (points at some ants on the floor) Ant. Ant.  
 2 Father (indicating a small beetle nearby) And that's a bug.  
 3 D Bug.

The next example is from Duncan aged 19 months:

- (19) 1 Mother (asking son about a shape) What does it look like?  
 2 Duncan A eight  
 3 Mother It looks like a square, doesn't it?  
 4 Duncan Square.

Both children were at roughly the one-word stage, and yet they used their limited capacities to ground the new words they had just heard (see Clark, 2007).

When adults have trouble understanding another adult, they can initiate a side sequence to clear up the trouble. Adults use the same technique in talking with children, as in this exchange with Abe, aged 2 years 5 months (Chouinard & Clark, 2003, p. 656):

- (20) 1 Abe the plant didn't cried  
 2 Father the plant cried?  
 3 Abe no  
 4 Father oh, the plant didn't cry  
 5 Abe uh-huh

Abe's father was apparently uncertain what Abe meant in line 1, so he initiated the side sequence in line 2. He offered first one interpretation, which Abe rejected, and then another, which Abe did accept. Like Jordan and his mother, the two of them persisted until they achieved joint closure on what was said.

Another way to make repairs is with what Jefferson (1982) called *embedded corrections*. Here is an example of Jefferson's:

- (21) Customer in a hardware store looking for a piece of piping:  
 1 Customer Mm, the wales are wider apart than that.  
 2 Salesman Okay, let me see if I can find one with wider threads.  
                   (Looks through stock) How's this?  
 3 Customer Nope, the threads are even wider than that.

In line 1, the customer used *wales* where he should have used *threads*. In line 2, the salesman corrected the term to *thread*, but he did so *en passant*, without remarking on the correction. In line 3 the customer took up the correction, again without comment. Jefferson called the pattern of *wales-threads-threads A-B-B*.

Adults correct children in the same way (Chouinard & Clark, 2003). Consider the father's exchange with Abe aged 2 years 5 months:

- (22) 1 Abe I want butter mine.  
 2 Father Okay give it here and I'll put butter on it.  
 3 Abe I need butter on it.

In line 1, Abe's "mine" was incorrect by adult standards, so in line 2, Abe's father did an embedded correction, reformulating "mine" as "on it." In line 3, Abe ratified his father's correction by repeating "on it." This is an *A-B-B* pattern. In other cases, the child took up the embedded correction with an acknowledgement such as "yeah" or "m-hm," as Abe did here:

- (23) 1 Abe my momma cry  
 2 Father Mommy cried  
 3 Abe uh-huh you yelling

This is an *A-B-yes* pattern.

Evidence shows that grounding in conversations with children is continuous and systematic. In a study by Chouinard and Clark (2003), five English- and French-speaking children were followed from age 2 to age 5 (with about 16,000 utterances). These children made observable mistakes in about 50% of their utterances, and when they did, adults corrected them between 50% and 70% of the time. There were errors in pronunciation, morphology, syntax, and meaning, which adults corrected about equally often. In making corrections, adults used side sequences about 65% of the time and embedded corrections 35% of the time. The children, for their part, were highly responsive. They explicitly took up the adult's reformulations between 20% and 50% of the time – in an *A-B-B* pattern, an *A-B-yes* pattern, or a rejection of the reformulation. Not only does grounding start at an early age, as with Jordan and Brenda, but it continues through to adulthood.

Most dialogues with the youngest children take place in daily routines such as eating, dressing, and going to bed or in other joint activities such as playing, looking at books, and going places. It is here that pointing, displaying, gazing, touching, and placement become particularly important.

Pointing is a good example (see Bates, Camaioni, & Volterra, 1975; Clark & Estigarribia, 2011; Tomasello, Carpenter, & Liszkowski, 2007). In (12), Jordan referred to the sponge by pointing at it, and his mother tried to ground his successive references by holding objects up. In studies by Clark and Estigarribia (2011; Estigarribia & Clark, 2007), parents talking with one-and-a-half- to three-year-olds often referred to novel objects (e.g., salad tongs) by pointing or holding them up; they continued their gestures until the children looked at the objects. Physical actions like these are characteristic of grounding with children.

### *Tracking a child's language abilities*

Adults are able to track children's language abilities precisely because these abilities are put on full display in the process of grounding.

Consider production. Jordan at 12 months was able to produce "repeated vocalizations," though not words, and could point at things communicatively. His mother could have inferred these abilities just from their attempts at grounding in this brief exchange.

Brenda, at 19 months, was able to produce single words such as "fei" and "ku" in vertical constructions such as "fei + ku" for "Fan cool." Her mother could infer all this from grounding their conversational moves. Brenda's father, in contrast, failed to ground even the words *car*, *go*, and *bus*, which kept him from discovering that she knew these words. Likewise, D and Duncan (at 18 and 19 months) were able to produce one- and two-word utterances, such as *ant*, *bug*, *a eight*, and *square*, and to point at things, abilities their parents could infer while grounding these few utterances. And although Abe, at 29 months, was able to produce utterances several words long, he had yet to learn past tense morphology, as in "didn't cried," and certain features of syntax, as in "I want butter mine." Abe's father could infer these limits *en passant* as they grounded what Abe said.

Tracking children's understanding is similar. At 12 months, Jordan was able to understand references when his mother pointed at things or held them up, knowledge she could infer from their exchanges. Likewise, at 19 months, Brenda was able to understand single word utterances like "Bathroom?" "Fan! Yeah" and "Cool, yeah," and perhaps even "Fan makes you cool." And at 18 and 19 months, D and Duncan showed that they understood simple introductions such as "That's a bug" and "It looks like a square" and could pick out the new terms in them – *bug* and *square*. And at 29 months, Abe showed he could understand utterances as complex as "the plant didn't cry," "I'll put butter on it," and "Mommy cried."

The conclusion seems clear. Adults infer children's current abilities to produce and understand particular words, constructions, and gestures in the very process of grounding those words, constructions, and gestures. And the parents of Jordan, Brenda, D, Duncan, and Adam all relied on these inferences in formulating their next utterances – just as they would have done in talking with adults. As Ferguson argued, talking with children is simply a "modified version of normal adult speech."

### **Learning to talk with children**

The optimal way for children to learn language is in conversations with adults. The evidence is well summarized in *First Language Acquisition*. The more time children spend in conversations with adults, the earlier and more thoroughly they

learn their first language. It isn't enough to see and hear adults talking with each other, or to see and hear adults talking *for* or *to* children. Learning is optimal when adults spend time talking *with* children.

The reason talking *with* children is optimal is that it requires grounding. Learning a language requires feedback, and the best feedback is focused, immediate, and frequent. Grounding has all three properties. It is focused on one utterance at a time. It is immediate – about the current or preceding utterance – which leads to immediate repairs when they are needed. And it provides explicit or implied feedback on virtually every utterance. Grounding provides just the type of feedback that is optimal for learning.

If interaction with feedback is the best way for children to learn how adults speak, then it should also be the best way for *adults* to learn how *children* speak. Because of the child-as-moving-target problem, adults need a way of assessing a child's abilities moment by moment. It is grounding that enables them to do that. Why? Because grounding yields focused, immediate, and frequent evidence for both (a) what a child can produce at that moment, and (b) what the child can understand.

Most features attributed to child-directed speech are really features of grounding. For example, if baby talk words such as *mommy*, *daddy*, *tummy*, and *peepee* are pronunciations that small children are able to produce, they are also pronunciations that adults can repeat in confirming what the children said. Likewise, imperatives and questions are first parts of adjacency pairs that lead children to reveal what they understand in the second parts. A child who understands "Sit down" will sit down. A child who understands "What is that?" will give an answer. But a child who understands "It looks like a square" need not respond at all. So reliable evidence of grounding is easier to establish with imperatives and questions than with assertions. Other features of child-directed speech have similar accounts.

To sum up, we learn how to talk with children in the very act of talking with children. The reasoning is straightforward. We cannot talk *with* a child without grounding what gets said, and in grounding what gets said, we automatically track that child's current abilities, both in speaking and in listening. As adults, we may pride ourselves on knowing so much more about language than the children we talk to. And yet, in face-to-face interaction, they are suddenly our peers, because grounding is a joint process. It is there we discover that we have as much to learn about a child's language as the child does about ours.

## Postscript

At the beginning of our careers, Eve Clark and I were advised one day by a senior colleague, "If you each want to be known in your own right, develop separate areas

of expertise, and publish separately." Eve and I took his advice seriously. It was natural for us to split the world into adults and children. Eve studied language in children, and I studied language in adults. It was also easy to write separately. From that day on, we published together only twice – a book in 1977 on psychology and language, in which we wrote separate chapters, and a paper in 1979 on how nouns surface as verbs.

And yet Eve and I have remained each other's most valued consultants. She has commented on most of my papers, and I have commented on most of hers. And thanks to our division of labor, I have never had to become expert in language acquisition. If there was something I needed to know, all I had to do was ask.

Why am I saying all this? Because it explains how hard it was to write this essay. Eve wasn't told about this volume, and that left me entirely on my own. I couldn't make use of her unlimited expertise, nor could I rely on her editorial comments. The experience brought home to me just how indebted I am to Eve for her intellectual support over all these years. And imagine how much this essay would have benefited from her advice.

## References

- Bates, E., Camaioni, L., & Volterra, V. (1975). The acquisition of performatives prior to speech. *Merrill-Palmer Quarterly of Behavior and Development*, 21(3), 205–226.
- Chouinard, M.M., & Clark, E.V. (2003). Adult reformulations of child errors as negative evidence. *Journal of Child Language*, 30(3), 637–670. DOI: 10.1017/S0305000903005701
- Clark, E.V. (2007). Young children's uptake of new words in conversation. *Language in Society*, 36(2), 157–182. DOI: 10.1017/S0047404507070091
- Clark, E.V. (2009). *First language acquisition*. Cambridge: CUP. DOI: 10.1017/CBO9780511806698
- Clark, E.V., & Bernicot, J. (2008). Repetition as ratification: How parents and children place information in common ground. *Journal of Child Language*, 35(2), 349–372. DOI: 10.1017/S0305000907008537
- Clark, E.V., & Estigarribia, B. (2011). Using speech and gesture to introduce new objects to young children. *Gesture*, 11(1), 1–23. DOI: 10.1075/gest.11.1.01cla
- Clark, H.H. (1996). *Using language*. Cambridge: CUP. DOI: 10.1017/CBO9780511620539
- Clark, H.H., & Brennan, S.A. (1991). Grounding in communication. In L.B. Resnick, J.M. Levine, & S.D. Teasley (Eds.), *Perspectives on socially shared cognition* (pp. 127–149). Washington, DC: APA Books. DOI: 10.1037/10096-006
- Clark, H.H., & Krych, M.A. (2004). Speaking while monitoring addressees for understanding. *Journal of Memory and Language*, 50(1), 62–81. DOI: 10.1016/j.jml.2003.08.004
- Clark, H.H., & Schaefer, E.F. (1987). Collaborating on contributions to conversations. *Language and Cognitive Processes*, 2(1), 19–41. DOI: 10.1080/01690968708406350
- Clark, H.H., & Schaefer, E.F. (1989). Contributing to discourse. *Cognitive Science*, 13(2), 259–294. DOI: 10.1207/s15516709cog1302\_7

- Clark, H.H., & Wilkes-Gibbs, D. (1986). Referring as a collaborative process. *Cognition*, 22(1), 1–39. DOI: 10.1016/0010-0277(86)90010-7
- Estigarribia, B., & Clark, E.V. (2007). Getting and maintaining attention in talk to young children. *Journal of Child Language*, 34(4), 799–814. DOI: 10.1017/S0305000907008161
- Ferguson, C.A. (1964). Baby talk in six languages. *American Anthropologist*, 66(6), 103–114. DOI: 10.1525/aa.1964.66.suppl\_3.02a00060
- Ferguson, C.A. (1977). Baby talk as a simplified register. In C.E. Snow & C.A. Ferguson (Eds.), *Talking to children: Language input and acquisition*, 209–235. Cambridge: CUP.
- Ferguson, C.A. (1982). Simplified registers and linguistic theory. In L.K. Obler (Ed.), *Exceptional language and linguistics* (pp. 49–66). New York, NY: Academic Press.
- Fernald, A. (1989). Intonation and communicative intent in mothers' speech to infants: Is the melody the message? *Child Development*, 60, 1497–1510. DOI: 10.2307/1130938
- Fernald, A. (1993). Approval and disapproval: Infant responsiveness to vocal affect in familiar and unfamiliar languages. *Child Development*, 64(3), 657–674. DOI: 10.2307/1131209
- Forrester, M.A., & Cherington, S.M. (2009). The development of other-related conversational skills: A case study of conversational repair during the early years. *First Language*, 29(2), 166–191. DOI: 10.1177/0142723708094452
- Geluykens, R. (1992). *From discourse process to grammatical construction: On left-dislocation in English*. John Benjamins. DOI: 10.1075/sidag.1
- Godfrey, J.J., Holliman, E.G., & McDaniel, J. (1992). SWITCHBOARD: Telephone speech corpus for research and development. *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing. (ICASSP-92)* (pp. 517–520). San Francisco, CA: IEEE.
- Golinkoff, R.M. (1983). The preverbal negotiation of failed messages: Insights into the transition period. In R.M. Golinkoff (Ed.), *The transition from prelinguistic to linguistic communication* (pp. 57–78). Mahwah, NJ: Lawrence Erlbaum Associates.
- Goodwin, C. (1986). Between and within: Alternative sequential treatments of continuers and assessments. *Human Studies*, 9(2–3), 205–217. DOI: 10.1007/BF00148127
- Jefferson, G. (1982). On exposed and embedded correction in conversation. *Studium Linguisticum*, 14, 58–68.
- Jefferson, G. (1984). Notes on a systematic deployment of the acknowledgement tokens "yeah" and "mm hm". *Papers in Linguistics*, 17, 197–216.
- Jefferson, G. (2002). Is "no" an acknowledgment token? Comparing American and British uses of (+)/(-) tokens. *Journal of Pragmatics*, 34(10), 1345–1383.
- Laakso, M., & Soininen, M. (2010). Mother-initiated repair sequences in interactions of 3-year-old children. *First Language*, 30(3–4), 329–353. DOI: 10.1177/0142723710370534
- Newport, E.L., Gleitman, H., & Gleitman, L.R. (1977). 'Mother, I'd rather do it myself': Some effects and non-effects of maternal speech style. In C.E. Snow & C.A. Ferguson (Eds.), *Talking to children: Language input and acquisition* (pp. 109–149). Cambridge: CUP.
- Sacks, H., Schegloff, E.A., & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *Language*, 50, 696–735. DOI: 10.2307/412243
- Salonen, T., & Laakso, M.L. (2009). Self-repair of speech by four-year-old Finnish children. *Journal of Child Language*, 36(4), 855–882. DOI: 10.1017/S0305000908009240
- Schegloff, E.A. (1982). Discourse as an interactional achievement: Some use of "uh-huh" and other things that come between sentences. In D. Tannen (Ed.), *Georgetown University Round Table on Languages and Linguistics, Analyzing discourse: Text and talk* (pp. 71–93). Washington DC: Georgetown University Press.

- Schegloff, E.A. (2007). *Sequence organization in interaction, Vol. 1: A primer in conversation analysis*. Cambridge: CUP. DOI: 10.1017/CBO9780511791208
- Schegloff, E.A., & Sacks, H. (1973). Opening up closings. *Semiotica*, 8(4), 289–327. DOI: 10.1515/semi.1973.8.4.289
- Schegloff, E.A., Jefferson, G., & Sacks, H. (1977). The preference for self-correction in the organization of repair in conversation. *Language*, 53, 361–382.
- Scollon, R. (1976). *Conversations with a one year old: A case study of the developmental foundation of syntax*. Honolulu, HI: University Press of Hawaii.
- Scollon, R. (1979). A real early stage: An unzipped condensation of a dissertation on child language. In E. Ochs (Ed.), *Developmental pragmatics* (pp. 215–227). New York, NY: Academic Press.
- Snow, C.E. (1972). Mothers' speech to children learning language. *Child Development*, 43(2), 549–565. DOI: 10.2307/1127555
- Snow, C.E. (1977). The development of conversation between mothers and babies. *Journal of Child Language*, 4(1), 1–22. DOI: 10.1017/S0305000900000453
- Svartvik, J., & Quirk, R. (1980). *A corpus of English conversation*. Lund: Gleerup.
- Tomasello, M., Carpenter, M., & Liszkowski, U. (2007). A new look at infant pointing. *Child Development*, 78(3), 705–722. DOI: 10.1111/j.1467-8624.2007.01025.x
- Trevarthen, C. (1979). Communication and cooperation in early infancy: A description of primary intersubjectivity. In M. Bullowa (Ed.), *Before speech: The beginning of interpersonal communication*, pp. (321–347). Cambridge: CUP.